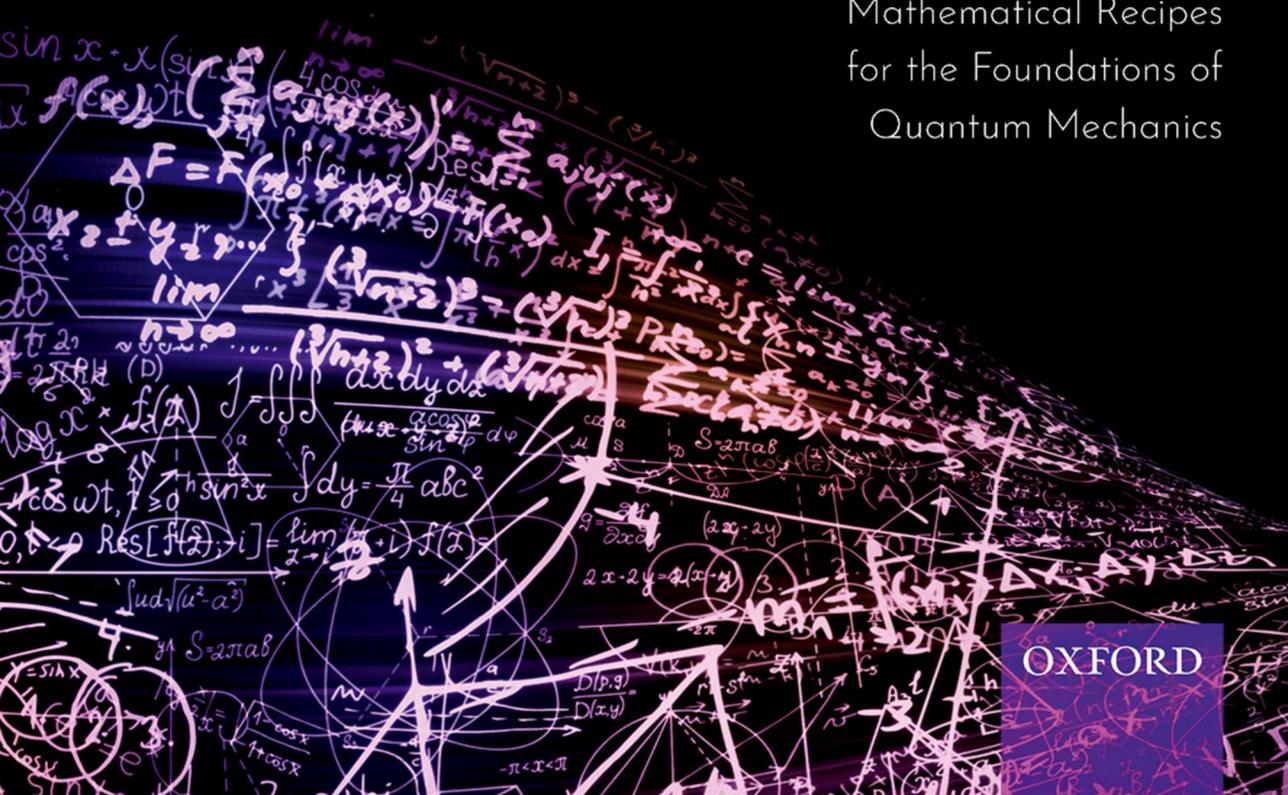


JIM BAGGOTT

THE QUANTUM COOKBOOK

Mathematical Recipes
for the Foundations of
Quantum Mechanics



OXFORD

THE QUANTUM COOKBOOK

The Quantum Cookbook

Mathematical Recipes for the Foundations
of Quantum Mechanics

Jim Baggott

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© Jim Baggott 2020

The moral rights of the author have been asserted

First Edition published in 2020

Impression: 1

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, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2019942517

ISBN 978-0-19-882785-6 (HBK)

ISBN 978-0-19-882786-3 (PBK)

DOI: 10.1093/oso/9780198827856.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

*To myself, aged 18,
when I took my first course on quantum
mechanics*

Preface

So this was the situation which I found at Cornell. Hans [Bethe] was using the old cookbook quantum mechanics that Dick [Feynman] couldn't understand. Dick was using his own private quantum mechanics that nobody else could understand. They were getting the same answers whenever they calculated the same problems.

Freeman Dyson*

There are a number of reasons why quantum mechanics is a difficult subject, both to teach and to learn. For sure, the subject is mathematically very challenging. But it is also *philosophically* challenging, forcing as it does a complete rethink of our naïve classical preconceptions concerning the ways in which we seek to represent physical reality in a scientific theory, and what we might expect such a representation to be telling us about it. The first challenge is recognized, and respected. The second perhaps less so.

I firmly believe that presentations of quantum mechanics that focus on formalism at the expense of all experimental, historical, and philosophical context run great risks of losing all but the most able students. Of course, science does not—it cannot—respect history. We make progress in science by moving on, by building on what we've learned without worrying overmuch precisely how we learned it. But the simple truth is that quantum mechanics did not suddenly materialize overnight in the minds of its creators, fully formed, complete with all its axioms and principles. It was instead tortured from much more familiar classical physical descriptions, such as thermodynamics, statistical mechanics, electromagnetic theory, special relativity, and atomic theory, over a period of decades, as physicists struggled to interpret a series of ever more baffling experimental results.

Only later was a much higher level of abstraction introduced into quantum mechanics, in an attempt to establish a secure mathematical foundation that would eradicate all the confusing classical misconceptions inherited from its birth and early childhood. This was a process begun by Paul Dirac in *The Principles of Quantum Mechanics* (1930) and John von Neumann in *Mathematical Foundations of Quantum Mechanics* (first published in German in 1932). Such was their success that we tend to overlook just how alien their approach was at the time. For example, in his review of *Principles*, Wolfgang Pauli warned that Dirac's rather abstract formalism and focus on mathematics at the expense

* From *Disturbing the Universe* by Freeman Dyson, copyright © 1981. Reprinted with permission of Basic Books, an imprint of Perseus Books, LLC, a subsidiary of Hachette Book Group, Inc.

of physics held ‘a certain danger that the theory will escape from reality’.* I fear he was right to be concerned.

Many students find the formalism completely baffling when they encounter it for the first time. Lectures and textbooks that dive straight into discussions of wavefunctions or vector spaces without any historical or philosophical context can leave students stranded, left to ponder: ‘Just how did they get that?’, and ‘Where did that come from?’† If the formalism is delivered to students as though the philosophical problems of its interpretation do not exist or are irrelevant, this can give the misleading impression that we really *do* understand what quantum mechanics is all about. Those students who then fail to penetrate the fog of confusion are left to brood on their own inadequacy. This is unfortunate, as the charismatic American physicist Richard Feynman was closer to the truth with his famous quote: ‘I think I can safely say that *nobody* understands quantum mechanics.’‡

To expose the real nature of the challenge, I believe it is helpful first to demonstrate that, despite appearances, mathematical complexity is not the principal problem. The second step is to provide some historical context, if only to explain that quantum mechanics was derived from real physics, not abstract mathematics. It also helps to explain how, *from the very beginning*, the physicists who helped to establish the theory were obliged to wrestle with its interpretation, arguing very energetically among themselves as they did so. Then we get the real insight. Nobody understands quantum mechanics because of its deep *philosophical* problems: we really don’t understand what it *means*, possibly because we’re not meant to.

The Quantum Cookbook is an attempt to provide a unique bridge between a popular exposition and a formal textbook presentation. The former tend to be necessarily extremely light on mathematical details, whereas the latter tend to be formalism-heavy, often paying little or no heed to problems of interpretation (though there are some notable exceptions). For curious readers with some background in physics and sufficient mathematical capability, neither popular exposition nor textbook provides them with what they need.

The book’s mission is to expose the real nature of the problems with quantum mechanics by walking readers step-by-step through the derivation of its most important foundational equations, including one result from special relativity ($E = mc^2$) because of its importance at key points in the story. It aims to provide sufficient context to enable readers to come to their own conclusions about its interpretation and meaning. In the process of demystifying the mathematics as much as possible, I hope also to demonstrate how *flexibly* mathematics is often applied in science, through simplified models,

* Wolfgang Pauli, *Die Naturwissenschaften*, 19 (1931), 188–9, quoted in Helge Kragh, *Dirac: A Scientific Biography*, Cambridge University Press, Cambridge, UK, 1990, p. 79.

† Especially those who, like me, were plunged into quantum mechanics without first being introduced (even superficially) to classical Hamiltonian mechanics and special relativity. My first encounter with quantum mechanics was as a student studying for a degree in chemistry, and these topics did not belong in a chemistry curriculum.

‡ Richard Feynman, *The Character of Physical Law*, MIT Press, Cambridge, MA, 1967, p. 129. The italics are mine.

and limiting assumptions and approximations. Despite its ‘unreasonable effectiveness’, mathematics is still a *language*, one that leaves plenty of room for interpretation (and doubt).

The first nine chapters build these results more or less chronologically, unfolding pretty much as they were presented by those physicists who left their fingerprints all over quantum mechanics. These are the quantization of energy (Planck); the equivalence of mass and energy (Einstein); quantum numbers and quantum jumps (Bohr); wave–particle duality (de Broglie); operators, eigenfunctions, and eigenvalues (wave mechanics—Schrödinger); quantum probability (Born); the uncertainty principle (Heisenberg and Robertson); the exclusion principle and electron spin (Pauli and Heisenberg); and relativistic quantum mechanics (electron spin and antimatter—Dirac). Chapter 10 will be a little different in structure as it deals with the establishment of the standard quantum formalism based on the concepts of state vectors in Hilbert space (Dirac and von Neumann).

The noted contemporary theorist Lee Smolin told me recently that as an undergraduate student he had been extraordinarily fortunate. In the spring semester of his first year at Hampshire College in Amherst, Massachusetts, he learned about quantum mechanics from Herbert Bernstein, by Smolin’s account a great physics teacher. The course concluded with detailed discussions of something called the EPR argument, named for Einstein, Boris Podolsky, and Nathan Rosen, and a famous theorem devised by John Bell. ‘Bell’s paper was not yet widely known and had by that time very few citations,’ Smolin explained to me. ‘That was probably the first and only quantum mechanics course for undergraduates that included EPR and Bell.’*

So, the final two chapters of *The Quantum Cookbook* cover topics that would not normally form part of an introductory course on quantum mechanics, though I would argue that they should (and Smolin would agree). These deal with the treatment of measurement in the quantum formalism (von Neumann) and the challenge posed by the interpretation of quantum entanglement and non-locality (Einstein, Bohm, and Bell).

Now, in setting out the book’s ambitions I need to be absolutely clear. It is *not* my intention to provide a detailed historical analysis of these physicists’ original publications, many of which are in any case intentionally obscure, as they sought to cover up underlying violence to the mathematics, unjustified assumptions, and occasional conceptual leaps of faith. After all, science doesn’t much care *how* a theory is arrived at: what’s important is how well the theory accommodates existing empirical facts and how well its predictions fare in the light of new examination.

The intention is rather to present the simplest possible derivations that are broadly consistent with the originals, which make use of current nomenclature, and which can be followed relatively easily. It’s important that readers can appreciate the logic, the nature of the challenges, and the occasional bit of mathematical sleight-of-hand.

Demystifying the mathematics means taking nothing for granted. Each derivation is presented as a ‘recipe’ with listed ingredients, including standard results from the

* Lee Smolin, personal communication, 7 September 2017.

mathematician's toolkit, such as the odd trigonometric identity, Stirling's formula, a standard integral or two, or a Taylor series expansion. Each recipe is then set out in a series of hopefully easy-to-follow steps, such that readers with limited ability in algebra and differential calculus and a background in physics should be able to cope. I've tried to write these recipes sympathetically, for readers who—like me—will often struggle to follow the logic of a derivation which misses out steps that are 'obvious' to the author, or which use techniques that readers are assumed to know. More mathematically competent readers who do not need everything spelled out in this way may therefore prefer to skip the intermediate steps.

Either way, I'm hopeful that readers will agree with my conclusion. There are obvious exceptions, but for the most part these derivations are triumphs of physical intuition over mathematical rigour and consistency.

The purpose of *The Quantum Cookbook* is not to teach readers how to *do* quantum mechanics, and it is not intended as a course textbook (it doesn't include any worked examples or problems). My hope is that this might prove to be a useful supplementary text for an introductory course, one that helps readers understand how to *think* about quantum mechanics.

My personal relationship with quantum mechanics now spans more than 40 years. Aside from making me feel quite old, this means that my debt of thanks by now extends to innumerable teachers, researchers, and authors whose efforts have helped bring light and inspiration in equal measure. I'd like to acknowledge personal debts to Peter Atkins, whose lectures and textbook *Molecular Quantum Mechanics* provided much technical clarity and insight, and Ian Mills, my erstwhile colleague at the University of Reading, who provided essential guidance as my understanding of the philosophical dimensions of the theory began its long, slow awakening. I must also give thanks to Lee Smolin and Carlo Rovelli, with whom more recent discussions helped to remind me why this is such an endlessly fascinating subject, and which encouraged me to return to it.

And, of course, I owe eternal gratitude to Sonke Adlung, Ania Wronski, Lucia Perez, and the production team at Oxford University Press for (once more) giving me the opportunity to get this lot off my chest.

Jim Baggott
September 2019

Contents

About the Author	xiii
Prologue: What's Wrong with This Picture? <i>The Description of Nature at the End of the Nineteenth Century</i>	1
1 Planck's Derivation of $E = h\nu$ <i>The Quantization of Energy</i>	19
2 Einstein's Derivation of $E = mc^2$ <i>The Equivalence of Mass and Energy</i>	35
3 Bohr's Derivation of the Rydberg Formula <i>Quantum Numbers and Quantum Jumps</i>	55
4 De Broglie's Derivation of $\lambda = h/p$ <i>Wave-Particle Duality</i>	73
5 Schrödinger's Derivation of the Wave Equation <i>Quantization as an Eigenvalue Problem</i>	89
6 Born's Interpretation of the Wavefunction <i>Quantum Probability</i>	111
7 Heisenberg, Bohr, Robertson, and the Uncertainty Principle <i>The Interpretation of Quantum Uncertainty</i>	133
8 Heisenberg's Derivation of the Pauli Exclusion Principle <i>The Stability of Matter and the Periodic Table</i>	157
9 Dirac's Derivation of the Relativistic Wave Equation <i>Electron Spin and Antimatter</i>	179
10 Dirac, Von Neumann, and the Derivation of the Quantum Formalism <i>State Vectors in Hilbert Space</i>	203
11 Von Neumann and the Problem of Quantum Measurement <i>The 'Collapse of the Wavefunction'</i>	219

12	Einstein, Bohm, Bell, and the Derivation of Bell's Inequality <i>Entanglement and Quantum Non-locality</i>	243
	Epilogue: A Game of Theories <i>The Quantum Representation of Reality</i>	265
	Appendix 1: Cavity Modes	269
	Appendix 2: Lorentz Transformation for Energy and Linear Momentum	272
	Appendix 3: Energy Levels of the Hydrogen Atom	275
	Appendix 4: Orthogonality of the Hydrogen Atom Wavefunctions	278
	Appendix 5: The Integral Cauchy-Schwarz Inequality	280
	Appendix 6: Orbital Angular Momentum in Quantum Mechanics	282
	Appendix 7: A Very Brief Introduction to Matrices	288
	Appendix 8: A Simple Local Hidden Variables Theory	291
	Index	295

About the Author

Jim Baggott is an award-winning science writer. A former academic scientist, he now works as an independent business consultant but maintains a broad interest in science, philosophy, and history, and continues to write on these subjects in his spare time. His books have been widely acclaimed and include the following:

Quantum Reality: The Quest for the Real Meaning of Quantum Mechanics - A Game of Theories (to be published in 2020)

Quantum Space: Loop Quantum Gravity and the Search for the Structure of Space, Time, and the Universe (2018)

Mass: The Quest to Understand Matter from Greek Atoms to Quantum Fields (2017)

Origins: The Scientific Story of Creation (2015)

Farewell to Reality: How Fairy-tale Physics Betrays the Search for Scientific Truth (2013)

Higgs: The Invention and Discovery of the 'God Particle' (2012)

The Quantum Story: A History in 40 Moments (2011, re-issued in 2015)

Atomic: The First War of Physics and the Secret History of the Atom Bomb 1939–49 (2009, re-issued in 2015), shortlisted for the Duke of Westminster Medal for Military Literature, 2010

A Beginner's Guide to Reality (2005)

Beyond Measure: Modern Physics, Philosophy and the Meaning of Quantum Theory (2004)

Perfect Symmetry: The Accidental Discovery of Buckminsterfullerene (1994)

The Meaning of Quantum Theory: A Guide for Students of Chemistry and Physics (1992)

Prologue

What's Wrong with This Picture?

The Description of Nature at the End of the Nineteenth Century

Anyone already familiar with some of the more bizarre implications of quantum mechanics—its phantoms of probability; particles that are waves and waves that are particles; cats that are at once both alive and dead; its uncertainty, non-locality, and seemingly ‘spooky’ goings-on—might look back rather wistfully on the structure of classical mechanics. We might be tempted to think that classical mechanics offers a much more appealing or comforting description of nature, one that is unambiguous, definite, and certain.

Such was the appeal of the classical structure that, towards the end of the nineteenth century, some in the physics community famously declared that all the most pressing problems had now been solved. Perhaps we can forgive the great British physicist Lord Kelvin (William Thomson) his sense of triumph, rendered rather hollow only through hindsight. In a lecture delivered to the British Association for the Advancement of Science in 1900 he is supposed to have declared: ‘There is nothing new to be discovered in physics now. All that remains is more and more precise measurement.’¹

Whilst it appears there is no evidence that Kelvin ever said this,² in *Light Waves and Their Uses*, a book based on a series of lectures delivered in 1899 to the Lowell Institute in Boston, Massachusetts, American physicist Albert Michelson wrote:³

Many other instances might be cited, but these will suffice to justify the statement that ‘our future discoveries must be looked for in the sixth place of decimals.’ It follows that every means which facilitates accuracy in measurement is a possible factor in a future discovery.

It is thought that Michelson was quoting Kelvin.

Of course, if this was indeed Kelvin’s assessment, he was quickly proved wrong.⁴ We now know that the classical structure breaks down in the microscopic realm of atoms and subatomic particles, and Isaac Newton’s laws of motion can’t handle objects moving at or near light speed. But, within its domain of applicability, classical mechanics is surely free of mystery and much less prone to endless bickering about what it’s all supposed to mean.

Except that it isn’t, really.

2 *The Interpretation of Space and Time*

Make no mistake, despite its intuitive appeal, classical mechanics is just as fraught with conceptual difficulties and problems of interpretation as its quantum replacement. The problems just happen to be rather less obvious, and so more easily overlooked (or, quite frankly, ignored). Quantum mechanics was born not only from the failure wrought by trying to extend classical physical principles into the microscopic world of atoms and molecules, but also from the failure of some of its most familiar and cherished concepts. To set the scene and prepare us for what follows, I thought it might be worth highlighting some of the worst offenders.

The Interpretation of Space and Time

The classical system of physics that Newton had helped to construct, by ‘standing on the shoulders of giants’,⁵ consists of three laws of motion and a law of universal gravitation. The *Mathematical Principles of Natural Philosophy*, first published in 1687, uses these laws to bring together aspects of the terrestrial physics of everyday objects and the ‘celestial’ mechanics of planetary motion, in what was nothing less than a monumental synthesis, fully deserving of its exalted status in science history. So closely did the resulting description agree with and explain observation and experiment that there could be little doubting its essential ‘truth’. By the end of the nineteenth century it had stood, unrivalled, for more than two hundred years.

Unrivalled, but by no means unquestioned. Newton’s mechanics might be intuitive but it demands a number of fairly substantial conceptual or philosophical trade-offs. Perhaps the most fundamental is that Newton’s physics is assumed to take place in an *absolute* space and time. This is a problem because, if it existed, an absolute space would form a curious kind of container, presumably of infinite dimensions, within which some sort of mysterious cosmic metronome marks absolute time. Actions impress forces on matter and things happen *within* the container and all motion is then referred to a fixed frame, thereby making all motion absolute.

If we could take all the matter out of Newton’s universe, then we would be obliged to presume that the empty container would remain, and the metronome would continue to tick. The existence of such a container implies a vantage point from which it would be possible to look down on the entire material universe, a ‘God’s-eye view’ of all creation.

But a moment’s reflection tells us that, despite superficial appearances, we only ever perceive objects to be moving towards or away from each other, changing their *relative* positions. This is relative motion, occurring in a space and time that are in principle defined only by the relationships between the objects themselves. If the motion is uniform, then there is in principle *no* observation we can make that will tell us if this object is moving relative to that object, or the other way around. In the *Mathematical Principles*, Newton acknowledged this in what he called our ‘vulgar’ experience.

If we can never perceive motion in an absolute space and time then we arguably have no good reason to accept that these exist. And if there is no absolute coordinate system of the universe; no absolute or ultimate inertial frame of reference against which all motion can be measured, then there can be no such thing as absolute motion. Newton’s

arch-rival, German philosopher Gottfried Wilhelm Leibniz, argued: ‘the fiction of a finite material universe, the whole of which moves about in an infinite empty space, cannot be admitted. It is altogether unreasonable and impracticable.’⁶ Now, any concept that is not accessible to observation or experiment in principle, a concept for which we can gather no empirical evidence, is typically considered to be *metaphysical* (meaning literally ‘beyond physics’).

Why, then, did Newton insist on a system of absolute space and time, one that we can never directly experience and which is therefore entirely metaphysical? Because by making this metaphysical *pre-commitment* he found that he could formulate some very highly successful laws of motion. Success breeds a certain degree of comfort, and a willingness to suspend disbelief in the grand but sometimes rather questionable foundations on which theoretical descriptions are constructed.

Classical Mechanics and the Concept of Force

Classical mechanics is the physics of the ordinary. Suppose we apply a force F for a short time interval, dt , to an object that is stationary or moving with constant velocity, v , in a straight line. In the *Mathematical Principles*, Newton explains that the force is simply an ‘action’, exerted or impressed upon the object, which effects a change in its linear momentum (p , given by the object’s mass m multiplied by v), by an amount dp . If we assume that mass is an intrinsic property of the object and does not change with time or with the application of the force, then dp is then simply the mass multiplied by the change in velocity: $dp = m dv$.

Applying the force may change the magnitude of the velocity (up or down) and/or it may change the direction in which the object is moving. Newton’s second law of motion is then expressed as $F dt = dp$ ($= m dv$). This equation may not look very familiar, but we can take a further step. Dividing both sides by dt gives

$$F = \frac{dp}{dt}. \quad (\text{P.1})$$

Logically, the greater the applied force, the greater the rate of change of linear momentum with time. But, as we’ve seen, $dp/dt = m dv/dt$. Obviously, dv/dt is the rate of change of velocity with time, or the object’s *acceleration*, usually given the symbol a . Hence Newton’s second law can be restated as the much more familiar

$$F = ma. \quad (\text{P.2})$$

Force equals inertial mass times acceleration, and we think of inertial mass as the measure of an object’s *resistance* to acceleration under an applied force. This is a statement of Newton’s second law *equation of motion*.

Though famous, this result actually does not appear in the *Mathematical Principles*, despite the fact that Newton must have been aware of this particular formulation, which

features in German mathematician Jakob Hermann's treatise *Phoronomia*, published in 1716.* It is sometimes referred to as the 'Euler formulation', after the eighteenth-century Swiss mathematician Leonhard Euler.

Newton's version of classical mechanics is expressed in terms of forces which result from the application of various mechanical 'actions'. Whilst it is certainly true to say that the notion of mechanical force still has much relevance today, the attentions of eighteenth- and nineteenth-century physicists switched from force to *energy* as the more fundamental concept. My foot connects with a stone, this action impressing a force on the stone. But a better way of thinking about this is to see the action as transferring energy to the stone.

Like force, the concept of energy also has its roots in seventeenth-century mechanical philosophy. Leibniz wrote about *vis viva*, a 'living force' expressed as mv^2 , and he speculated that this might be a *conserved* quantity, meaning that it can only be transferred between objects or transformed from one form to another—it can't be created or destroyed. The term 'energy' was first introduced in the early nineteenth century and it gradually became clear that kinetic energy—the energy of motion—is not in itself conserved. It was important to recognize that a system might also possess *potential* energy by virtue of its physical characteristics and situation. It was then possible to formulate a law of conservation of the *total* energy—kinetic plus potential—largely through the efforts of physicists concerned with the principles of thermodynamics, which we will go on to examine later in this Prologue.

If we denote the kinetic energy as T and the potential energy as V , then the total energy is simply $T + V$. The kinetic energy T is given by

$$T = \frac{1}{2}mv^2 = \frac{p^2}{2m}. \quad (\text{P.3})$$

It's helpful to understand how this relates to Newton's force, F . Differentiating (P.3) with respect to time gives

$$\frac{dT}{dt} = \frac{1}{2}m \frac{d(v^2)}{dt} = \frac{1}{2}m \left(v \frac{dv}{dt} + v \frac{dv}{dt} \right) = mv \frac{dv}{dt} = mva. \quad (\text{P.4})$$

In (P.4) we have assumed the mass m to be independent of time and we have applied the product rule $d(uv)/dx = v(du/dx) + u(dv/dx)$ to the evaluation of $d(v^2)/dt$. We can now make use of the second law $F = ma$ and the chain rule

* Newton published a third edition of the *Mathematical Principles* in 1726 and, if he had been so minded, could have incorporated this version of the second law.

$$\frac{dT}{dt} = Fv \quad \text{and so} \quad dT = Fvdt = F \frac{dx}{dt} dt = Fdx. \quad (\text{P.5})$$

Integrating then allows us to express the kinetic energy in terms of force as follows:

$$T = \int Fdx. \quad (\text{P.6})$$

We can now put Newton's conception of force on a much firmer basis. We *define* the potential energy V as

$$V = - \int Fdx. \quad (\text{P.7})$$

This shift in emphasis from force to energy in the eighteenth and nineteenth centuries meant that it made more sense to define the secondary property of force in terms of the primary property of potential energy:

$$F = - \frac{dV}{dx}. \quad (\text{P.8})$$

Equations (P.7) and (P.8) make perfect sense. Lifting a heavy weight from its initial position on the floor to shoulder height involves the application of a force which changes the potential energy of the weight. The force applied is negative (as it acts *against* gravity), and transfers energy from the gravitational field into the potential energy of the weight. Letting go of the weight exposes it to the force of gravity, converting the gravitational potential energy it contains into kinetic energy, and it falls back to its initial position on the floor. The force is directed in such a way as to reduce the potential energy—hence the negative sign in (P.7)—driving the system 'downhill'. And the 'steeper' the shape of the potential energy curve (the faster the potential energy changes with position), the greater the resulting force, (P.8).

Setting up the relationship between force and potential energy in this way means that in a closed system which cannot exchange energy with the outside world the rate of change of total energy with time balances to zero—energy can be moved back and forth between potential and kinetic forms but the *total energy is conserved*:

$$\frac{dT}{dt} + \frac{dV}{dt} = mva + \frac{dV}{dx} \frac{dx}{dt} = mva + v \frac{dV}{dx} = v \left(ma + \frac{dV}{dx} \right) = v(ma - F). \quad (\text{P.9})$$

We can see from this that the time derivatives of the expressions for kinetic and potential energy sum to zero—the total energy doesn't change with time.

This shift of emphasis led to a substantial and profound reformulation of classical mechanics, first by Italian mathematician and astronomer Joseph-Louis Lagrange (in 1764) and subsequently by Irish physicist William Rowan Hamilton (in 1835). This wasn't simply about recasting Newton's laws in terms of energy. Hamilton in particular

greatly elaborated and expanded the classical structure and the result, called *Hamiltonian mechanics*, extended the number of mechanical situations to which the theory could be applied.

Newton's equation of motion $F = ma$ is formulated in terms of position coordinates (such as Cartesian coordinates x, y, z) and time. This is fine in principle for very simple systems involving at most one or two objects, but it quickly becomes problematic for systems involving large numbers of objects. To define the physical 'state' of a system consisting of, say, N objects, such that we can predict how the system will evolve in time, we would need to specify the position *and* the velocity of *each* of the N objects in three-dimensional space, at specific moments in time. It's not enough just to specify the positions—Newton's second law applies to objects that are already in a state of rest or uniform motion, so to predict what happens next we also need to know how fast and in which directions the objects are moving as the force is applied. In other words we need a total of $6N$ coordinates *for each object*.

We can think of the motion of the system as a 'trajectory' in an abstract $6N$ -dimensional *configuration space*. Instead of positions and velocities, Hamilton's reformulation makes use of the positions of the objects and their *momenta*. If we keep things simple by restricting ourselves to a single object with inertial mass m moving along a single position coordinate x , then these *canonical coordinates* are (x, p) , where p is again the object's linear momentum. Hamilton's choice defines what would subsequently become known as *phase space*.

The motion of the object is then represented by the *trajectory of a point in the phase space coordinates*. This gives us an advantage in more complex systems because instead of specifying the initial positions and velocities of all the objects in a $6N$ -dimensional configuration space, in Hamiltonian mechanics we just need to specify the system's initial position in phase space. It then becomes possible to predict the future time evolution of the system from any starting point on its phase space diagram.

As we will draw on many of these concepts in what follows, it's worth taking the time here for a very brief and somewhat superficial look at Hamiltonian mechanics. The *Hamiltonian* of a classical system is simply the total energy, E , and is defined as

$$H (= E) = T + V. \quad (\text{P.10})$$

In Hamiltonian mechanics we're obviously interested to know the behaviour of the Hamiltonian H with respect to the canonical coordinates, which in a single dimension are given by (x, p) . This behaviour is summarized in Hamilton's equations of motion:

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x} \quad \text{and} \quad \frac{dx}{dt} = \frac{\partial H}{\partial p}. \quad (\text{P.11})$$

These equations may appear somewhat unfamiliar, but the second establishes a fairly straightforward connection between momentum and velocity. Remember, we assume

that the potential energy V is independent of p , and from (P.3) we know that $T = p^2/2m$:

$$\frac{\partial H}{\partial p} = \frac{\partial T}{\partial p} = \frac{\partial}{\partial p} \left(\frac{p^2}{2m} \right) = \frac{p}{m} = v = \frac{dx}{dt}. \quad (\text{P.12})$$

And the first is simply a restatement of Newton's second law:

$$-\frac{\partial H}{\partial x} = -\frac{\partial V}{\partial x} = F = \frac{dp}{dt} (= ma). \quad (\text{P.13})$$

It's worth noting in passing that we've traded Newton's single equation of motion, which is a *second-order* differential equation (remember, $a = d^2x/dt^2$), for Hamilton's two *first-order* partial differential equations, (P.11).

We can get some sense for how this works by considering a simple example. In one-dimensional simple harmonic motion (such as a low-amplitude pendulum or an object suspended on a spring), an object of mass m oscillates back and forth with an angular frequency ω under the action of a 'restoring' force $F = -m\omega^2x$. The Hamiltonian for this system is, therefore,

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2 \quad (\text{P.14})$$

(remember, $V = -\int Fdx$), and Hamilton's equations of motion are

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x} = -m\omega^2x \quad \text{and} \quad \frac{dx}{dt} = \frac{\partial H}{\partial p} = \frac{p}{m}. \quad (\text{P.15})$$

If we define the initial position (x_0) to be the origin at time $t = 0$ (i.e. $x_0 = 0$), then the solutions of these equations have the particularly simple form

$$p = p_0 \cos \omega t \quad \text{and} \quad x = \frac{p_0}{m\omega} \sin \omega t, \quad (\text{P.16})$$

where p_0 is the initial momentum. In a phase space with canonical coordinates (x, p) , the motion describes an elliptical trajectory:

$$\frac{x^2}{(p_0/m\omega)^2} + \frac{p^2}{p_0^2} = 1. \quad (\text{P.17})$$

Switching to a phase space description allows us to represent the mechanics in terms of the single trajectory of a point in a multidimensional space, summarizing the motion of the *entire system*, not the individual objects. This was a generalization discovered by French mathematician Henri Poincaré in 1888, from his study of the infamous

three-body problem (and which also led him to appreciate the sensitivity of dynamical systems to initial conditions, later to become an obsession of chaos theory).

A year later, Poincaré noted a rather curious phenomenon. In an ideal mechanical system with a finite upper bound on the volume of available phase space (one in which no objects can escape the system and in which energy is conserved), within a sufficiently long, but finite, time the phase space trajectory will return to its starting point.* This is called *Poincaré recurrence*. No matter how many objects are involved, if the dynamics unfold from some starting configuration and we have sufficient patience, the system *will* return to this configuration.

The Troublesome Concept of Mass

The development of our understanding of potential energy in the nineteenth century allowed us to put Newton's concept of force on a much firmer basis, as we've seen. There would appear to be no reason to question our understanding of any of the other concepts which appear in Hamilton's equations. We haven't forgotten the problems of absolute space and time but we surely know what we mean when we talk about acceleration, momentum, and mass.

But what, precisely, *is* inertial mass? Newton provides a handy definition very early in the *Mathematical Principles*:⁷

The quantity of matter is the measure of the same, arising from its density and bulk conjunctly . . . It is this that I mean hereafter everywhere under the name body or mass. And the same is known by the weight of each body; for it is proportional to the weight, as I have found by experiments on pendulums, very accurately made, which shall be shewn hereafter.

If we interpret Newton's use of the term 'bulk' to mean volume, then the mass of an object is simply its density multiplied by its volume. It doesn't take long to figure out that this definition is entirely circular, as Austrian physicist Ernst Mach pointed out many years later:⁸

With regard to the concept of "mass", it is to be observed that the formulation of Newton, which defines mass to be the quantity of matter of a body as measured by the product of its volume and density, is unfortunate. As we can only define density as the mass of a unit of volume, the circle is manifest.

We have to face up to the rather unwelcome conclusion that in classical mechanics we don't really know what inertial mass is.

* Poincaré's theorem also requires that phase volume is conserved as the system evolves, which is true for all Hamiltonian systems by virtue of Joseph Liouville's 1838 theorem.

The Force of Gravity

In Newton's law of universal gravitation, two objects with masses m_1 and m_2 experience a force of gravity that is proportional to the product of their masses (are these the same as inertial mass?) and inversely proportional to the square of the distance between them, r , or $F = Gm_1m_2/r^2$, where G is Newton's gravitational constant.

This was another great success, but it also came with another hefty price tag. Although the symbol F might be the same, Newton's force of gravity is distinctly different from the kinds of forces involved in his laws of motion. The latter forces are *impressed*; they are caused by actions such as kicking, shoving, pulling, or whirling. They require physical contact between the object at rest or moving uniformly and whatever it is we are doing to change the object's motion. Newton's gravity works very differently. It is presumed to pass instantaneously between the objects that exert it, through some kind of curious action at a distance. It was not at all clear how this was supposed to work. Leibniz was again dismissive: "This, in effect, is going back to qualities which are occult or, what is more, inexplicable."⁹

Newton himself had nothing to offer. In a general discussion (called a 'general scholium'), which he added to the 1713 second edition of the *Mathematical Principles*, he wrote:¹⁰

Hitherto we have explain'd the phaenomena of the heavens and of our sea, by the power of Gravity, but have not yet assign'd the cause of this power... I have not been able to discover the cause of those properties of gravity from phaenomena, and I frame no hypotheses.

Light Waves and the Ether

Newton sought to extend the scope of his mechanics to include light, and in his treatise *Opticks*, first published in 1704, he concluded that light is essentially 'atomic' in nature, consisting of tiny particles, or corpuscles. Two of his contemporaries, English natural philosopher and experimentalist Robert Hooke and Dutch physicist Christiaan Huygens, had argued compellingly in favour of a wave theory of light, and Newton's incendiary disputes with Hooke led him to postpone publication of *Opticks* until after Hooke's death in March 1703. Such was Newton's standing and authority that the corpuscular theory held sway for more than a hundred years.

But in a series of papers read to the Royal Society of London between 1801 and 1803, nearly eighty years after Newton's death, an English medical doctor (and part-time physicist) called Thomas Young revived the wave theory as the only logical explanation for the phenomena of light diffraction and interference. In one experiment, commonly attributed to Young (although historians are divided on whether he actually performed it), he showed that when passed through two narrow, closely spaced holes or slits, light produces a pattern of bright and dark fringes. These are readily explained in terms of a wave theory of light in which the peaks and troughs of the light waves from the two

slits start out in phase, spread out beyond, and overlap. Where a peak of one wave is coincident with a peak of another, the two waves add and reinforce to produce constructive interference, giving rise to a bright fringe. Where a peak of one wave is coincident with a trough of another, the two waves cancel to produce destructive interference, giving a dark fringe.

Today this logic seems inescapable, but Young's conclusions were roundly criticized, with some condemning his explanation as 'destitute of every species of merit'.¹¹ Nevertheless, as the nineteenth century progressed, the wave theory gained a slow, if somewhat grudging, acceptance. Then, as is so often the case in science, perhaps the most compelling arguments in favour of the wave theory emerged from a seemingly unrelated discipline.

The intimate connection between the phenomena of electricity and magnetism was established over a long period of study in the nineteenth century, most notably through the extraordinary experimental work of Michael Faraday at London's Royal Institution. Drawing on analogies with fluid mechanics, over a ten-year period from 1855 Scottish physicist James Clerk Maxwell developed a theory of *electromagnetic* fields whose properties are described by a set of complex differential equations. These equations can be manipulated to give expressions for the space and time dependences of the electric field \mathbf{E} and magnetic field \mathbf{B} in a vacuum, as follows (again simplified to one dimension):

$$\frac{\partial^2 \mathbf{E}}{\partial x^2} = \varepsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad \text{and} \quad \frac{\partial^2 \mathbf{B}}{\partial x^2} = \varepsilon_0 \mu_0 \frac{\partial^2 \mathbf{B}}{\partial t^2}. \quad (\text{P.18})$$

In Eq. (P.18), ε_0 and μ_0 are the relative permittivity and permeability of free space, respectively. The former is a measure of the resistance of a medium (in this case, the 'vacuum') to the formation of an electric field—a certain fixed electric charge will generate a greater electric flux in a medium with low permittivity. The latter is a measure of the ability of a medium to support a magnetic field—applying a certain fixed magnetic field strength will result in greater magnetisation in a medium with high permeability.

Maxwell had made no assumptions about how these fields are supposed to move through space. But his equations not only demonstrate rather nicely the symmetry of the interdependent electric and magnetic fields, they also rather obviously describe wave motion. For a wave travelling in one dimension with velocity v , a generalized wave equation can be written as

$$\frac{\partial^2}{\partial x^2} \Psi(x, t) = \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \Psi(x, t), \quad (\text{P.19})$$

where $\Psi(x, t)$ is a generalized 'wavefunction'. From (P.18) and (P.19) we can deduce that $v = 1/\sqrt{\varepsilon_0 \mu_0}$. The velocity of Maxwell's 'electromagnetic waves' could now be determined from the experimental values of the relative permittivity and permeability of free space, which had been reported by German physicists Wilhelm Weber and Rudolf Kohlrausch in 1856. Maxwell found that:¹²

This velocity is so nearly that of light, that it seems we have strong reason to conclude that light itself (including radiant heat, and other radiations if any) is an electromagnetic disturbance in the form of waves propagated through the electromagnetic field according to electromagnetic laws.

But an electromagnetic disturbance in what? If we throw a stone into a lake, and watch as the disturbance ripples across the surface of the water, we conclude that the waves travel in a 'medium'—the water in this case. There could be no escaping the conclusion: electromagnetic waves had to be waves in some kind of medium. Maxwell himself didn't doubt that electromagnetic waves must move through the ether, a purely hypothetical, tenuous form of matter thought to fill all of space.

And here's another price to be paid. All the evidence from experimental and observational physics suggested that if the ether really exists, then it couldn't be participating in the motions of observable objects. The ether must be stationary. If the ether is stationary, then it is also by definition absolute: it fills precisely the kind of container demanded by an absolute space. A stationary ether would define the ultimate inertial frame of reference.

Newton required an absolute space that sits passively in the background and which, by definition, we can never experience. Now we have an absolute space that is supposed to be filled with ether. That's a very different prospect.

If the Earth spins in a stationary ether, then we might expect there to be an ether wind at the surface (actually, an ether drag, but the consequences are the same). The ether is supposed to be very tenuous, so we wouldn't expect to feel this wind like we feel the wind in the air. But, just as a sound wave carried in a high wind reaches us faster than a sound wave travelling in still air, we might expect that light travelling in the direction of the ether wind should reach us faster than light travelling against this direction. A stationary ether suggests that the speed of light should be different when we look in different directions.

Any differences were expected to be very small, but nevertheless still measurable with late-nineteenth-century optical technology. In 1887, American physicists Albert Michelson and Edward Morley performed experiments to look for such differences using a device called an interferometer, in which a beam of light is split and sent off along two different paths. The beams along both paths set off in phase, and they are then brought back together and recombined. Now, if the total path taken by one beam is slightly longer than the total path taken by the other, then when the beams are recombined, peak may no longer coincide with peak and the result is destructive interference. Alternatively, if the total paths are equal but the speed of light is different along different paths, then the result will again be interference.

But they could detect no differences. Within the accuracy of the measurements, the speed of light was found to be constant, irrespective of direction, suggesting that there is no such thing as a stationary ether. This is one of the most important 'negative' results in the entire history of experimental science.

Newton's laws of motion demand an absolute space and time that we can't experience or gain any empirical evidence for. Maxwell's electromagnetic waves demand a stationary ether to move in, but we can't gain any evidence for this either.

Atoms and the Second Law

The second law in question here is that of thermodynamics, the science born from the study of engines, and particularly the relationship between heat and work. French physicist and engineer Sadi Carnot is credited with establishing the basis for thermodynamics with his 1824 publication *Reflections on the Motive Power of Fire*, although some ten years passed before the merits of Carnot's work were realized by his fellow countryman Émile Clapeyron, who helped rid Carnot's theory of the concept of heat as a fluid, called caloric. Nine years later English physicist James Joule identified the mechanical equivalent of heat—motion and heat are equivalent and interchangeable—and helped to establish the law of conservation of energy. When Kelvin coined the term 'thermodynamics' in 1854, the conservation of energy was summarized as its first law.

Carnot had imagined that useful work can be derived as heat 'falls' from a higher temperature to a lower temperature, just as falling water will turn a paddle wheel. But Carnot imagined that heat would be conserved, meaning that all the usable heat is transferred into work without loss, allowing the possibility of perpetual motion and obviously in conflict with the conservation of energy. In 1850, German physicist Rudolf Clausius resolved this problem by declaring as a principle that heat cannot spontaneously flow from a cold object to a hot object, with the rest of the universe remaining unchanged.* For a system undergoing a closed cyclic process in which heat is transformed into work which is then transformed back into heat, Clausius expressed this principle mathematically as an inequality:

$$\oint \frac{\delta Q}{T} \leq 0. \quad (\text{P.20})$$

In this equation the increments δQ represent the net amount of heat added to a system from an external reservoir at temperature T . For processes that are cyclical *and reversible*, meaning that infinitesimal changes that maintain thermodynamic equilibrium can in theory restore the initial state, the equality holds. But for processes that are irreversible the inequality holds. The logic here is fairly simple. In an irreversible process the (positive) heat input divided by the higher temperature *will always be smaller* than the (negative) heat output divided by the lower temperature. Summing (or integrating) over the cycle means $\delta Q/T < 0$.

Clausius was able to show that the ratio $\delta Q/T$ is a quantity which depends only on the physical state of the system, and not on the details of the path taken to produce it. Hence it is a *property* of the system, also called a function of state (or state function). In 1865 he went a little further, and identified this property as the *entropy* (symbol S) of the system, which he now defined for *reversible* open paths connecting some initial state i with a final state f , as

* Kelvin formulated a similar principle at around the same time.

$$\Delta S_{rev} = S_f - S_i = \int_i^f dS_{rev} = \int_i^f \left(\frac{\delta Q}{T} \right)_{rev}. \quad (\text{P.21})$$

The property of entropy accounts for the dissipative loss of heat (or energy) from the system, but to get a real sense for what this means we need to look at how Eqs. (P.20) and (P.21) can be combined. Equation (P.20) applies to a closed cycle which may involve paths that are reversible and/or irreversible, whereas (P.21) applies only to open paths that are reversible. So, imagine a closed cycle in which the path from initial to final state is irreversible, but the return path from final to initial state is reversible. From (P.20) we have

$$\oint \frac{\delta Q}{T} = \int_i^f \frac{\delta Q}{T} + \int_f^i \left(\frac{\delta Q}{T} \right)_{rev} \leq 0. \quad (\text{P.22})$$

But the return path is reversible, and so from (P.21) we know that

$$\int_f^i \left(\frac{\delta Q}{T} \right)_{rev} = S_i - S_f. \quad (\text{P.23})$$

Hence,

$$\int_i^f \frac{\delta Q}{T} + S_i - S_f \leq 0, \quad \text{or} \quad \Delta S_{irr} = S_f - S_i \geq \int_i^f \frac{\delta Q}{T}. \quad (\text{P.24})$$

We see that the change in entropy from initial to final state in an irreversible process is *always greater* than the corresponding change for a completely reversible process, which is a direct consequence of applying Clausius' inequality. Heat transfer to a system increases its entropy, and heat transfer from a system will decrease its entropy, but factors that result in irreversibility (such as friction and other loss mechanisms) *will always increase the entropy*. We can see this more clearly by generalizing (P.24) for any irreversible process in an *isolated* system (one which doesn't exchange energy with the external environment). In such a situation $\delta Q = 0$ and

$$\Delta S_{irr} \geq 0, \quad (\text{P.25})$$

which is a statement of the second law of thermodynamics.

This version of the second law was deduced by German physicist Max Planck in his 1879 doctoral thesis. He regarded it as a much more general statement, and so more fundamental and profound. For an isolated system energy will be conserved (first law) but entropy will inexorably increase to a maximum (second law) as the system achieves thermal equilibrium. Irreversibility and the increase in entropy are intimately linked, defining an 'arrow of time' such that any reverse process, spontaneously *decreasing*

entropy, implies running *backwards* in time, ‘so that a return of the world to a previously occupied state is impossible’.¹³

And therein lies another problem.

As the science of thermodynamics was being worked out in the nineteenth century, so too was an elaborate mechanical theory of atoms. Hard, impenetrable, indestructible atoms, no more sophisticated than those imagined by the atomist philosophers of ancient Greece, had been an accepted metaphysical pre-commitment of seventeenth-century mechanical philosophers such as Newton. This despite the fact that they were not really necessary and did not feature in the classical mechanics that these philosophers helped to establish. Newton’s atomism was quite influential in the eighteenth century, but as atoms appeared to lie well beyond the scope of any available experimental or observational technology, they remained firmly speculative.¹⁴

In 1738, the Swiss physicist Daniel Bernoulli had argued that the properties of gases could be understood to derive from the rapid motions of the innumerable atoms or molecules that constitute the gas (hereafter referred to simply as ‘atoms’). Gas pressure then results from the *impact* of these atoms on the surface of the vessel that contains them. Gas temperature is the result of the *motions* of the atoms. This *kinetic theory of gases* bounced around for a few decades before being refined by Clausius in 1857. Two years later Maxwell developed a mathematical formula for the distribution of the velocities of the atoms in a gas. As it is obviously impossible to keep track of the motions of large numbers of individual atoms, Maxwell was obliged to resort to probabilities and so derived a probability distribution. This was generalized in 1871 by Austrian physicist Ludwig Boltzmann, and is now known as the Maxwell–Boltzmann distribution.

Boltzmann built further on Maxwell’s ideas, applying probabilities to the distribution of *energy* instead of velocity, as he worked to derive all the most important thermodynamic quantities based on the underlying motions of the system’s constituent atoms. In 1877 he derived the expression for the entropy of an ideal gas which is carved on his gravestone,

$$S = k_B \ln(W), \quad (\text{P.26})$$

where k_B is Boltzmann’s constant and W is the number of microstates (the number of individual configurations of atomic positions and velocities or momenta that are possible). If it is assumed that all these microstates are equally probable, then the probability for each microstate is simply $1/W$. Bulk quantities such as pressure, temperature, and entropy summarize the macrostate of the system.

The second law can now be interpreted as the natural evolution of an isolated system towards the largest number of available microstates. If we pump a gas into one corner of an otherwise empty container, we anticipate that this system will evolve dynamically: the gas will expand and become diluted so that it fills all of the available space. The number of microstates (atomic positions and momenta) that are available in the final equilibrium situation is much greater than in the initial situation. Entropy increases.

We can now see how Hamiltonian mechanics is perfectly suited to the interpretation of thermodynamics in terms of complex systems involving the motions of large numbers of atoms. In his *Lectures on Gas Theory*, published in 1896, Boltzmann himself defined

'phase' to mean the collective state of a gas derived from the positions and momenta of all its constituent atoms, though he held back from calling it phase space.¹⁵

But towards the end of the nineteenth century the existence of atoms was still largely a matter for metaphysical speculation and many physicists were inclined to be rather stubborn about them. It's perhaps difficult for readers who have lived with the fallout from the 'atomic age' to understand why perfectly competent scientists should have been so reluctant to embrace atomic ideas, but we must remember that by 1900 there was very little evidence for their existence. Some physicists, such as the arch-empiricist Mach, rejected them completely. To make matters considerably worse, the statistical mechanical interpretation of thermodynamics produced conclusions which some physicists found extremely discomfoting.

Statistics have a dark side. They deal with *probabilities*, not certainties. What thermodynamics argues to be unquestionably irreversible and a matter of irresistible natural law, statistics argues that this is only the most probable of many different possible alternatives. The conflict was most stark in the interpretation of the second law and in 1895, with Planck's approval, his research assistant Ernst Zermelo took the argument directly to the atomists in the pages of the German scientific journal *Annalen der Physik*.

If we were to release two gases of different temperature in a closed container, the second law predicts that the gases will mix and the temperature will become uniform, with the entropy of the mixture increasing to a maximum. However, according to the atomists, the behaviour of the gases is a consequence of the underlying mechanical motions of the atoms of each gas, and the equilibrium state of the mixture is simply the most probable of many possible states. Furthermore, such dynamical systems could be expected to exhibit Poincaré recurrence, implying that, if we wait long enough, the system will eventually return to its initial far-from-equilibrium state, with the gases once more separated at different temperatures. Such a possibility runs directly counter to the second law, which insists that in an isolated system undergoing spontaneous change, entropy can never decrease, Eq. (P.25).

Boltzmann had no real alternative but to accept what statistical mechanics implied. Entropy does not always increase, he argued, in contradiction to the most common interpretation of the second law. It just almost always increases. Statistically speaking, there are many, many more states of higher entropy than there are of lower entropy, with the result that the system spends much more time in higher entropy states. In effect, Boltzmann was saying that if we do indeed wait long enough, we might eventually catch a system undergoing a spontaneous reduction in entropy. This is as miraculous an event as a smashed cocktail glass spontaneously reassembling itself, to the astonishment of party guests.

To Planck, this stretched the interpretation of his cherished second law to breaking point. It may have been that Planck was not averse to the atomic theory per se—he was certainly well aware of the theory's successes. But he judged that it was unlikely to offer a productive approach to a deeper understanding of thermodynamics. In a letter to Wilhelm Ostwald in 1893 he declared that the atomic theory was nothing less than a 'dangerous enemy of progress'.¹⁶ Matter is continuous, not atomic, he insisted. He had no doubt that atomic ideas would eventually have to be abandoned, despite their

success, ‘in favour of the assumption of continuous matter’.¹⁷ In his historical analysis, American philosopher Thomas Kuhn argues that Planck’s ‘continuous medium’ would subsequently become the ether.¹⁸

In seeking to find a way to refute Boltzmann’s statistical arguments, Planck chose as a battleground the physics of ‘black body’ radiation. And this is where our story really begins.

.....

NOTES

1. See, for example, Walter Isaacson, *Einstein: His Life and Universe*, Simon & Schuster, London, 2007, p. 90.
2. Although this statement is widely attributed to Kelvin, Isaacson could find no evidence to suggest that Kelvin had actually said it. See Isaacson, *Einstein*, p. 575.
3. A. A. Michelson, *Light Waves and Their Uses*, University of Chicago Press, Chicago, 1903, p. 24.
4. In fairness to Kelvin, we should also note that in April 1900 he delivered a lecture to the Royal Institution in London concerned with what he perceived to be several ‘clouds’ hanging over the dynamical theory of heat and light.
5. Isaac Newton, letter to Robert Hooke, 15 February 1676. See: <https://digitallibrary.hsp.org/index.php/Detail/objects/9792>
6. Gottfried Wilhelm Leibniz, from his correspondence with Samuel Clarke (1715–16), *Collected Writings*, edited by G. H. R. Parkinson, J. M. Dent & Sons, London, 1973, p. 226.
7. Isaac Newton, *Mathematical Principles of Natural Philosophy*, first American edition translated by Andrew Motte, published by Daniel Adee, New York, 1845, p. 73.
8. Ernst Mach, *The Science of Mechanics: A Critical and Historical Account of Its Development*, 4th edition, translated by Thomas J. McCormack, Open Court Publishing, Chicago, 1919, p. 194 (first published 1893). See also Max Jammer, *Concepts of Mass in Contemporary Physics and Philosophy*, Princeton University Press, Princeton, NJ, 2000, p. 11; and O. Bellkind, *Physical Systems*, Boston Studies in the Philosophy of Science, 264 (2012), 119–44.
9. Gottfried Wilhelm Leibniz, *New Essays on the Human Understanding*, reproduced in *Collected Writings*, ed. Parkinson, p. 167.
10. Newton, *Mathematical Principles*, p. 506.
11. This quote is taken from Henry Brougham, *The Edinburgh Review, or Critical Journal*, 1, January 1803, pp. 450–6. This was an anonymous review, but Young correctly identified the author as Brougham, then a barrister and later Lord Chancellor of England. For more details on the controversy, see Christine Simon, ‘Thomas Young’s Bakerian Lecture’, *The Fortnightly Review*, 2014: <http://fortnightlyreview.co.uk/2014/09/thomas-young/#fnref-13923-13>
12. J. Clerk Maxwell, ‘A Dynamical Theory of the Electromagnetic Field’, *Philosophical Transactions of the Royal Society of London*, 155 (1865), 466.
13. This statement appears in the very first introductory paragraph of Planck’s thesis, *Über den zweiten Hauptsatz der mechanischen Wärmetheorie* (On the Second Law of Thermodynamics), 1879. Quoted in Thomas S. Kuhn, *Black-body Theory and the Quantum Discontinuity 1894–1912*, University of Chicago Press, Chicago, 1978, p. 16.
14. See Alan Chalmers, *The Scientist’s Atom and the Philosopher’s Stone: How Science Succeeded and Philosophy Failed to Gain Knowledge of Atoms*, Springer, London, 2011, especially Chapters 6 and 7.

15. David D. Nolte, 'The Tangled Tale of Phase Space', *Physics Today*, April 2010, 33–8.
16. Max Planck, letter to Wilhelm Ostwald, 1 July 1893. Quoted in J. L. Heilbron, *The Dilemmas of an Upright Man: Max Planck and the Fortunes of German Science*, Harvard University Press, Cambridge, MA, 1996, p. 15.
17. Max Planck, *Physikalische Abhandlungen und Vorträge*, Vol. 1, Vieweg, Braunschweig, 1958, p. 163. Quoted in Heilbron, *Dilemmas*, p. 14.
18. Kuhn, *Black-body Theory*, p. 23.

.....

FURTHER READING

- Baggott, Jim, *Mass: The Quest to Understand Matter from Greek Atoms to Quantum Fields*, Oxford University Press, Oxford, 2017.
- Crease, Robert P., *A Brief Guide to the Great Equations: The Hunt for Cosmic Beauty in Numbers*, Robinson, London, 2009, see especially Chapters 2, 3, 5, and 6.
- Cushing, James T., *Philosophical Concepts in Physics: The Historical Relation Between Philosophy and Scientific Theories*, Cambridge University Press, Cambridge, UK, 1998.
- Heilbron, J. L., *The Dilemmas of an Upright Man: Max Planck and the Fortunes of German Science*, Harvard University Press, Cambridge, MA, 1996.
- Jammer, Max, *Concepts of Mass in Contemporary Physics and Philosophy*, Princeton University Press, Princeton, NJ, 2000.
- Kennedy, Robert E., *A Student's Guide to Einstein's Major Papers*, Oxford University Press, Oxford, 2012.
- Kibble, Tom W. B., and Berkshire, Frank H., *Classical Mechanics*, 5th edn, Imperial College Press, London, 2004.
- Kuhn, Thomas S., *Black-body Theory and the Quantum Discontinuity 1894-1912*, University of Chicago Press, Chicago, 1978.
- Susskind, Leonard, and Hrabovsky, George, *Classical Mechanics: The Theoretical Minimum*, Penguin, London, 2014.

Planck's Derivation of $E = h\nu$

The Quantization of Energy

Planck was born in Kiel in April 1858, descended from a line of pastors and professors of theology and jurisprudence. At school Planck was diligent and personable but not especially gifted. Physics was a subject for which Planck himself felt he had no particular talent, and he had once been counselled against a career in theoretical physics. His professor at the University of Munich had advised him that, with the discovery of the principles of thermodynamics, physics as a subject had been largely completed. There was, quite simply, nothing more to be discovered.¹ This was fine with Planck, who was quite content with the rather less heroic task of deepening the foundations of science. He had no real interest in making new discoveries.

He preferred the stability and predictability of a science which reflected the character of the bourgeois German society of which he was a part. He had risen through the academic ranks and established a solid international reputation as a master of classical thermodynamics, and especially the second law. Now in his early forties, he worked at a slow, steady, and conservative pace. By his own subsequent admission, he was 'peacefully inclined', and rejected 'all doubtful adventures'.²

Planck is thus a good candidate for the history of science's most unlikely revolutionary.

Black Body Radiation

As we saw in the Prologue, Planck was unwilling to accept Boltzmann's statistical interpretation of the second law. He therefore needed to find a way to show how irreversible processes could result from matter that forms a continuum. Such a continuum would exhibit some kind of collective, ordered, or correlated motion, in contrast to the disordered motions characteristic of the atoms of Maxwell and Boltzmann. Irreversibility would then be associated with changes in this collective motion; changes that are not described by resorting to arguments based on probabilities, which opens the door to unacceptable, entropy-reducing processes, no matter how improbable they may be. Although Planck had chosen to reject atoms, he held firm to the theory of mechanics, and in this way he hoped eventually to reconcile mechanics with thermodynamics.

That Planck should turn his attention from the thermodynamics of gases to the physics of black body radiation as a battleground might seem puzzling at first. But Planck saw no contradiction. To understand why, we first need to know a little more about it.

Heat any object to a high temperature and it will glow, emitting light of different colours. We say that the object is ‘red hot’ or ‘white hot’. Increasing the temperature of the object increases the intensity of the light and shifts it to a higher range of frequencies (shorter wavelengths). As it gets hotter, an object glows first red, then orange-yellow, then bright yellow, then brilliant white.

Theoreticians had sought to model the physics based on the notion of a black body, a completely non-reflecting object that is presumed to absorb and emit light radiation perfectly, without favouring any particular range of radiation frequencies or wavelengths (or colours). The density or intensity of the radiation that a black body emits, measured over a range of frequencies, is then directly related to the amount of energy it contains.

The properties of black body radiation could be studied in the laboratory using specialized cavities, vessels made of porcelain and platinum with almost perfectly absorbing walls. Such cavities could be heated, and the radiation released and trapped inside could be observed with the aid of a small pinhole, a bit like peeking into the glowing interior of an industrial furnace. Such studies provided more than just an interesting test of theoretical principles. Cavity radiation was also useful to the German Bureau of Standards as a reference for rating electric lamps.

Planck imagined that the source of the (continuous) electromagnetic radiation released into the cavity is a continuum of ideal mechanical vibrators, or ‘resonators’. These resonators were entirely imaginary, their sole purpose being to absorb and emit radiation and so bring the system—cavity and radiation—to a dynamic equilibrium. The radiation would have an entropy—just like a gas—and equilibrium would be characterized by maximum entropy.

Consequently, Planck wasn’t specific on where these resonators might be physically located, but if it helps, we can suppose they reside in the cavity material. We probably wouldn’t hesitate today to identify these with the electrons in the atoms of the material, but remember that the electron was only discovered in 1895 and in 1900 Planck was strenuously opposed to the idea of atoms. For now, let’s not worry overmuch about what these resonators might actually represent.

Planck subsequently acknowledged that the use of the term ‘resonators’ was inappropriate (a resonator oscillates only at specific—resonant—frequencies). These imaginary objects are actually so-called ‘linear Hertzian oscillators’, which we can think of as massless springs with electric charge at each end. Planck’s task was to show how irreversible processes (and the second law) could arise from the collective motions of the oscillators and the dynamic exchange of energy between the oscillators and the trapped radiation. No atoms to be seen, anywhere.

This must have seemed like a perfectly safe choice.

In the winter of 1859–60, the German physicist Gustav Kirchhoff had demonstrated that the ratio of emitted to absorbed radiation energy depends only on the frequency of the radiation and the temperature inside the cavity. This means that the density of radiation inside the cavity at equilibrium is a function only of frequency, ν , and

temperature, T , designated $\rho(\nu, T)$. The density does not depend in any way on the shape of the cavity, the shape of its walls, or the nature of the material from which the cavity is made. This implied that something quite fundamental concerning the physics of the radiation itself was being observed, and Kirchhoff challenged the scientific community to discover the origin of this behaviour.

Planck's Radiation Law

Much progress had been made. Studies of infrared (heat) radiation had in 1896 led German physicist Wilhelm Wien to devise Wien's law, which can be summarized (in modern notation) as follows:

$$\rho(\nu, T) = \frac{8\pi h\nu^3}{c^3} e^{-h\nu/k_B T}, \quad (1.1)$$

where c is the speed of light and h and k_B would later become known as Planck's constant and Boltzmann's constant, respectively. The real significance of these physical constants was not immediately apparent.

Wien's law seemed to be quite acceptable, and was supported by further experiments carried out by German physicist Friedrich Paschen at the Technical Academy in Hanover in 1897. But new experimental results reported in 1900 by Otto Lummer and Ernst Pringsheim at the Reich Physical-Technical Institute in Berlin showed that Wien's law failed at lower frequencies. Wien's law was clearly not the answer.

In June 1900, English physicist Lord Rayleigh (William Strutt) published details of a new theoretical model based on the 'modes of ethereal vibration' in the cavity. Each mode was supposed to possess a specific frequency, and could take up and give out energy continuously. Rayleigh assumed a classical distribution of energy over these modes. At equilibrium, each mode of vibration should then possess an energy directly proportional to the cavity temperature.

It's instructive to interrupt this historical narrative and fast-forward a few years to May 1905. Rayleigh obtained an expression for the constant of proportionality, but made an error in his calculation which was put right by James Jeans the following July. The result is now known as the Rayleigh-Jeans law:

$$\rho(\nu, T) = \frac{8\pi\nu^2}{c^3} k_B T. \quad (1.2)$$

Rayleigh's reasoning and use of thermodynamic principles was both logical and convincing, but the result was disastrous. The Rayleigh-Jeans law implies that $\rho(\nu, T)$ increases with the square of the radiation frequency without limit, and so the total emitted energy quickly mushrooms to infinity at high frequencies. In 1911 the Austrian physicist Paul Ehrenfest called this problem the 'Rayleigh-Jeans catastrophe in the ultraviolet', now commonly known as the *ultraviolet catastrophe*. Rayleigh's approach might have been perfectly logical, but the result was totally illogical.

But both Wien's law and the Rayleigh–Jeans law were glimpses of the complete picture, approximations of $\rho(\nu, T)$ applicable only at the extremes of high and low radiation frequency.

Planck had succeeded Kirchhoff at the University of Berlin in 1889, rising to full professor in 1892. He was unaware of Rayleigh's work when, on 7 October 1900, German physicist Heinrich Rubens visited him at his villa in the Berlin suburb of the Gr unewald. Rubens told him about some new experimental results he had recently obtained with his associate Ferdinand Kurlbaum.

Rubens and Kurlbaum had studied cavity radiation at even lower frequencies, and the behaviour they had observed set Planck thinking. After Rubens had left, Planck continued to work alone in his study. After some reflection, he found that he could now replace Wien's law with one of his own, 'a result of inspired guesswork, scientific tact, sober compromise, in short, of tinkering':³

$$\rho(\nu, T) = \frac{8\pi h\nu^3}{c^3} \frac{e^{-h\nu/k_B T}}{1 - e^{-h\nu/k_B T}}. \tag{1.3}$$

This result is shown graphically in Fig. 1.1. We can now see what happens. For very *high* ν (or short wavelengths), the term $e^{-h\nu/k_B T}$ becomes very small compared with 1 and Planck's radiation law reduces to Wien's law. If we multiply the exponential term in Planck's law top and bottom by $e^{h\nu/k_B T}$, we can re-write this term as $1/(e^{h\nu/k_B T} - 1)$.

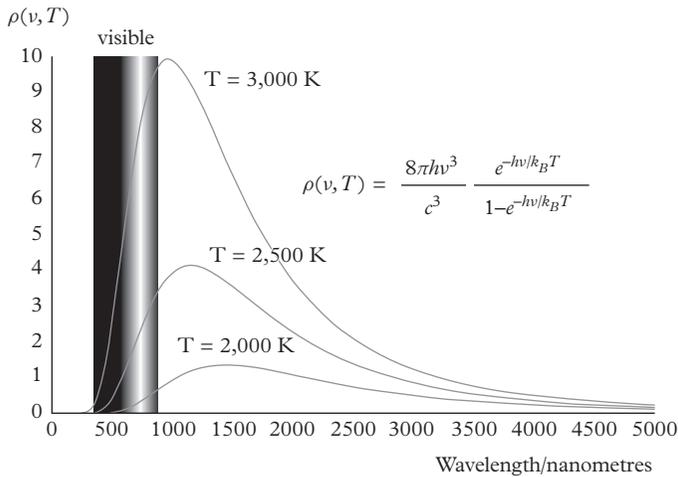


Figure 1.1 Planck's radiation law predicts the variation of radiation density with frequency or wavelength at different cavity temperatures. As the temperature increases, the peak wavelength shifts to shorter and shorter values.

For very *low* ν (long wavelengths), $e^{h\nu/k_B T}$ can be approximated as $1 + h\nu/k_B T$, and Planck's law reduces to the Rayleigh–Jeans law.

Planck sent Rubens a postcard summarizing details of his new radiation law, and he presented a crude derivation at a meeting of the German Physical Society on 19 October. He declared: 'I therefore feel justified in directing attention to this new formula, which, from the standpoint of electromagnetic radiation theory, I take to be the simplest excepting Wien's'.⁴ The next day, Rubens advised Planck that he had compared the experimental results with the new law and found 'completely satisfactory agreement in all cases'.⁵

But, although Planck's new law was satisfactory, it was in truth no more than a mathematical 'fit' to the data. Planck's challenge now was to find a deeper theoretical interpretation for it, and of course to make use of this interpretation to pursue his principal objective, which was to reconcile mechanics and thermodynamics and reassert the irreversible nature of the second law.

The Oscillator Energy

To follow Planck's logic we need to understand the nature of the relationships between $\rho(\nu, T)$ and thermodynamic quantities such as the internal energy U and entropy S of the cavity radiation. These latter quantities are, of course, inter-related.

According to the first law of thermodynamics the microscopic change in the internal energy of the radiation dU is equal to the amount of heat absorbed (δQ) less any work done. But the cavity radiation does no work (it's not used to drive a piston, for example), so $dU = \delta Q$; the change in internal energy logically derives only from the heat absorbed.

We know from Eq. (P.21) that $dS = \delta Q/T$, so in this situation $dU = TdS$, which is a version of the so-called fundamental thermodynamic relation. We can therefore get to the entropy from

$$dS = \frac{1}{T}dU \quad (1.4)$$

and integrating.

We can suppose that the frequency of vibration of the imaginary oscillators increases with increasing temperature and they exchange energy with the radiation *at the same frequency* with which they are vibrating. If we assume that *all* the energy of the oscillators is released into the cavity radiation, then at thermodynamic equilibrium we can further suppose that the internal energy (and entropy) of the oscillators is identical to that of the radiation. Consequently, Planck focused on the relationship between $\rho(\nu, T)$ and the average internal energy of the oscillators themselves (which we will refer to here as $U(\nu, T)$).

Between 1897 and 1899, Planck published a series of five papers titled 'On Irreversible Radiation Processes', in which he analysed a model system consisting of electromagnetic

waves interacting with a set of damped oscillators. In the final paper published in May 1899, he deduced the following important result:

$$\rho(\nu, T) = \frac{8\pi\nu^2}{c^3} U(\nu, T), \quad (1.5)$$

which states that the density of cavity radiation is directly related to the average internal energy of the oscillators. ‘The significant aspects of this equation,’ Planck wrote, ‘which was indispensable to me, is that according to it the energy of the resonant oscillator depends only on radiation intensity and frequency ν but not on any of its other properties.’⁶

There are a number of different ways to derive Eq. (1.5), but perhaps the simplest involves the analysis of the radiation in terms of cavity modes, as explained in Appendix 1. Interestingly, if we assume that the oscillators are all identical and follow a simple harmonic motion, then we know that the total average thermal energy of each oscillator consists of contributions of $\frac{1}{2}k_B T$ from translational motion (kinetic energy) and $\frac{1}{2}k_B T$ from its potential energy. This is the *equipartition theorem*, developed in the 1840s, which relates the temperature of a system to its average energies. Under this assumption, we see that $U(\nu, T)$ is independent of ν and equal to $k_B T$, which on substitution in Eq. (1.5) gives the Rayleigh–Jeans law.

So, in May 1899 Planck had access to a really rather simple and straightforward route to the Rayleigh–Jeans law, more than a year before Rayleigh himself published his (erroneous) version of it. However, it seems that, at this time, Planck was simply unaware of the equipartition theorem.

In any case, we know that the Rayleigh–Jeans law is physically unrealistic except as a low-frequency limit. Comparing Eq. (1.5) with Planck’s radiation law (1.3) shows that $U(\nu, T)$ is much more complicated than the equipartition theorem would suggest, and is indeed dependent on both ν and T :

$$U(\nu, T) = \frac{h\nu e^{-h\nu/k_B T}}{1 - e^{-h\nu/k_B T}} = \frac{h\nu}{e^{h\nu/k_B T} - 1}. \quad (1.6)$$

Planck’s task was now to make use of Eq. (1.6) to derive an expression for the average entropy of the oscillators, one that would be entirely consistent with his radiation law. This was to lead to ‘some weeks of the most strenuous work of my life’.⁷ Planck tried several different approaches, but he found that he was compelled to return to the statistical methods of his arch-rival Boltzmann. The mathematics led him in a direction he really did not want to go. He eventually succumbed, in a final act of desperation. As he later admitted: ‘A theoretical interpretation therefore *had* to be found at any cost, no matter how high’.⁸

Although the approach Planck took was subtly different from that of Boltzmann, as we will see, he found that black body radiation is absorbed and emitted *as though* it is composed of discrete ‘atoms’, which Planck called *quanta*. Moreover, he found that each

quantum of radiation has a fixed energy given by $E = h\nu$. Though much less familiar, this is an expression that is every bit as profound as Einstein's $E = mc^2$.

So, what did Planck do?

The Ingredients

1. Integration by substitution.
2. The standard integral of $\ln(x)$.
3. Boltzmann's equation for the entropy: $S = k_B \ln(W)$.
4. Combinatorics: partition theory.
5. Stirling's formula for the factorials of large numbers: $N! = (N/e)^N$.

The Recipe

It is perhaps a little unfortunate that this first recipe—the recipe that launched the quantum revolution—is somewhat convoluted and rather more difficult to follow than some of the other iconic equations of quantum mechanics. But this really shouldn't come as too much of a surprise. Thermodynamics is not the most obvious place to look for evidence of the quantum nature of radiation, and Planck had to torture the theory in a way that would eventually allow this conclusion to emerge from an entirely classical structure. It was always going to be a difficult birth.

We begin in Step (1) by manipulating the expression for the average internal energy of an oscillator, Eq. (1.6), such that it is in a form that can be more easily integrated. On integration we will have an expression for the entropy of an oscillator which is consistent with Planck's radiation law and which we can then generalize for a large collection of N oscillators. We can think of this as a derivation of the entropy based on thermodynamics which we know to be consistent with experimental data (as summarized by Planck's law).

As Planck soon realized, the result looks to all the world like a version of Boltzmann's equation for the entropy, based on the logarithm of the number of microstates, or the number of different possible configurations, W , as given in Eq. (P.26).

But the possible configurations of what, exactly? After all, there are no atoms or atomic motions in this system. So, in Step (2) we take another route, calculating the entropy using Boltzmann's methods but with a not-so-subtle difference. Boltzmann estimated the number of possible microstates W as the number of different ways that the available energy can be distributed over a large number of *distinguishable* atoms. But this couldn't give Planck the mathematical form demanded by the expression for the entropy which he had deduced in Step (1). So he did something different. He instead estimated W as the number of ways in which a series of *indistinguishable* energy elements can be distributed over a large number of oscillators.

We can think of this as a derivation of the entropy based on statistics. In his biography of Einstein, the American physicist Abraham Pais wrote: 'From the point of view of

physics in 1900 the logic of Planck's electromagnetic and thermodynamics steps was impeccable, but his statistical step was wild.⁹

Our final Step (3) is simply to compare the two derivations of the oscillator entropy and draw conclusions.

Step (1): Derive the Oscillator Entropy from Thermodynamics

Our starting point is the expression for the average internal energy of an oscillator which we know to be consistent with Planck's radiation law, Eq. (1.6). As this next bit is going to get complicated, let's simplify the notation, replacing $U(\nu, T)$ with U , but remembering that U is a function of both frequency and temperature. We can rearrange the expression we got for U in Eq. (1.6) as follows:

$$Ue^{h\nu/k_B T} - U = h\nu \quad \text{or} \quad Ue^{h\nu/k_B T} = h\nu + U. \quad (1.7)$$

Taking natural logarithms of both sides of this last expression gives

$$\ln(U) + \frac{h\nu}{k_B T} = \ln(h\nu + U), \quad (1.8)$$

which we can rearrange to give an expression for $1/T$:

$$\frac{1}{T} = \frac{k_B}{h\nu} [\ln(h\nu + U) - \ln(U)]. \quad (1.9)$$

We can tidy this up a bit by recognizing that

$$\ln(h\nu + U) = \ln\left(1 + \frac{U}{h\nu}\right) + \ln(h\nu) \quad (1.10)$$

and

$$\ln(U) = \ln\left(\frac{U}{h\nu}\right) + \ln(h\nu). \quad (1.11)$$

When we put these into the expression for $1/T$ the terms in $\ln(h\nu)$ cancel and we get

$$\frac{1}{T} = \frac{k_B}{h\nu} \left[\ln\left(1 + \frac{U}{h\nu}\right) - \ln\left(\frac{U}{h\nu}\right) \right]. \quad (1.12)$$

We can now substitute this expression for $1/T$ directly into the expression for dS , Eq. (1.4):

$$dS = \frac{k_B}{h\nu} \left[\ln\left(1 + \frac{U}{h\nu}\right) - \ln\left(\frac{U}{h\nu}\right) \right] dU. \quad (1.13)$$

To obtain an expression for the average oscillator entropy we need to integrate

$$S = \frac{k_B}{h\nu} \int \left[\ln\left(1 + \frac{U}{h\nu}\right) - \ln\left(\frac{U}{h\nu}\right) \right] dU. \quad (1.14)$$

This looks pretty complicated, but we can simplify it by making a couple of substitutions:

$$x = 1 + \frac{U}{h\nu} \quad \text{for which} \quad \frac{dx}{dU} = \frac{1}{h\nu} \quad \text{and} \quad dU = h\nu dx \quad (1.15)$$

and

$$y = \frac{U}{h\nu} \quad \text{for which} \quad \frac{dy}{dU} = \frac{1}{h\nu} \quad \text{and} \quad dU = h\nu dy. \quad (1.16)$$

Making these substitutions transforms the expression for S into

$$S = k_B \left[\int (\ln(x)dx - \ln(y)dy) \right]. \quad (1.17)$$

We can now use the standard integral $\int \ln(x)dx = x \ln(x) - x$ to give

$$S = k_B \left[\left(1 + \frac{U}{h\nu}\right) \ln\left(1 + \frac{U}{h\nu}\right) - \left(1 + \frac{U}{h\nu}\right) - \frac{U}{h\nu} \ln\left(\frac{U}{h\nu}\right) + \frac{U}{h\nu} \right] + C, \quad (1.18)$$

where C is a constant of integration.

The free terms $U/h\nu$ cancel, and the extra term $-k_B$ can be absorbed into the constant, leaving us with

$$S = k_B \left[\left(1 + \frac{U}{h\nu}\right) \ln\left(1 + \frac{U}{h\nu}\right) - \frac{U}{h\nu} \ln\left(\frac{U}{h\nu}\right) \right] + C'. \quad (1.19)$$

The final step involves one last bit of rearranging, to give

$$S = k_B \ln \left[\frac{\left(1 + \frac{U}{h\nu}\right)^{1+U/h\nu}}{\left(\frac{U}{h\nu}\right)^{U/h\nu}} \right] + C'. \quad (1.20)$$

The entropy of N oscillators is obviously N times the average entropy of one oscillator, and we quietly set aside the constant C' , as measures of entropy are based on differences in which any constant contributions will subtract out:

$$S_N = Nk_B \ln \left[\frac{\left(1 + \frac{U}{h\nu}\right)^{1+U/h\nu}}{\left(\frac{U}{h\nu}\right)^{U/h\nu}} \right]. \quad (1.21)$$

This is almost certainly *not* where Planck had hoped to get to. The term in square brackets in Eq. (1.21) is, as we will soon see, strongly reminiscent of a combinatorial expression, implying that this is nothing less than a version of Boltzmann's equation for the entropy, $S = k_B \ln(W)$. The mathematics had taken Planck in a direction he really had not wanted to go.

Planck had been fighting a losing battle against Boltzmann's logic for at least three years. He now succumbed to the inevitable. As he later explained: 'I busied myself, from then on, that is, from the day of its establishment, with the task of elucidating a true physical character for the [new radiation law], and this problem led me automatically to a consideration of the connection between entropy and probability, that is, Boltzmann's trend of ideas.'¹⁰

Step (2): Derive the Oscillator Entropy from Statistics

Boltzmann reasoned that the most probable state of a gas at thermal equilibrium is the one with the highest number of different ways to distribute (or *partition*) the available energy over the atoms or molecules of the gas, representing the maximum entropy at that energy. In essence, the second law of thermodynamics ensures that energy is distributed 'fairly' over all the particles that can carry it. In other words, it doesn't accumulate in a subset of these, equivalent to a small quantity of the gas (let's say the air in one corner of the room where you're sitting) spontaneously becoming hotter than the rest. As we've discussed already in the Prologue, this isn't completely ruled out by Boltzmann's logic; it's just that such a situation is very highly improbable. By working out the maximum number of possible ways to partition the energy (which Boltzmann called the number of *complexions*), it is a relatively simple step to calculate the entropy.

Boltzmann's approach to such a calculation involves assuming that the total available (and continuously variable) energy in a system can be thought of as being organized into a series of 'buckets'. The lowest energy bucket is assigned an energy E , the next an energy $2E$, the next $3E$, and so on. The atoms of the gas are then distributed among the buckets and the number of different possible permutations of atoms in the buckets is calculated. In this analysis the use of buckets is simply a calculation tool, with no real physical significance intended. All that Boltzmann had done was parcel up the energy so that he could count the number of atoms *in the energy range* zero to E , the range E to $2E$, the range $2E$ to $3E$, and so on, and thus calculate the number of different possible permutations.

For example, consider a gas consisting of just three atoms, which we assume are distinguishable and which we label a , b , and c . Let's assume this gas has a total energy of $4E$. We can't put all three atoms in the lowest energy bucket, as this doesn't account

for all the available energy. But the requirement is satisfied by putting two atoms in the lowest energy bucket E , and one in the $2E$ bucket. How many permutations are possible? There are just three. We can put atoms a and b into the lowest energy bucket and c in the next, a permutation we can label as $[ab, c]$. We can also put atoms a and c in the lowest energy bucket, and b in the next, labelled $[ac, b]$. The third possible permutation is $[bc, a]$ —see Fig. 1.2.

In general, the number of different possible permutations W is given by

$$W = N! \prod_{i=1,2,\dots} \frac{1}{P_i!} \tag{1.22}$$

where N is the number of particles and P_i is the ‘occupation number’ (the number of atoms in the i th energy bucket). In the simple example given above, $N = 3$ and $N! = 6$. The number of particles in energy bucket 1 is 2 (so $P_1 = 2, 1/P_1! = 1/2! = 1/2$) and the number of particles in energy bucket 2 is 1 (so $P_2 = 1, 1/P_2! = 1/1! = 1$). W is then

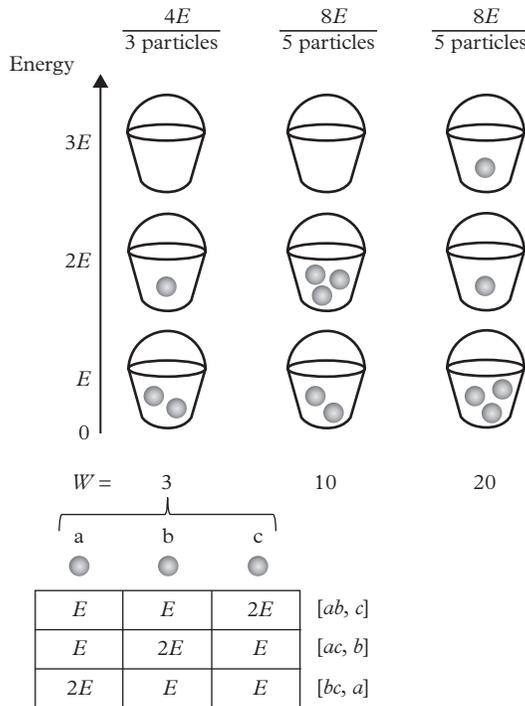


Figure 1.2 In Boltzmann’s approach the total available energy in a system is assumed to be continuously variable but organized into a series of ‘buckets’. This makes it possible to count the number of atoms in the energy range zero to E , E to $2E$, $2E$ to $3E$, and so on. The entropy is then calculated from the maximum number of possible permutations, W .

equal to $6 \times \frac{1}{2} \times 1 = 3$. Incidentally, if we assume each of these permutations is equally probable, then the probability for each permutation is simply $1/W = 1/3$.

It obviously gets more interesting as the number of atoms and the amount of energy increases. In the case of $8E$ distributed over 5 atoms, we can see that there are two possibilities, also shown in Fig. 1.2. In one of these we put 2 atoms in energy bucket 1 and 3 in energy bucket 2, giving the number of permutations $W = 10$. However, putting 3 particles in bucket 1, 1 in bucket 2, and 1 in bucket 3 increases W to 20. Nature will favour the combination which maximizes the complexity, and hence the entropy.

But if we look back at the expression for entropy derived from thermodynamics and Planck's radiation law, Eq. (1.21), we see immediately that the term in square brackets doesn't bear much resemblance to Eq. (1.22). Something else is going on.

Planck realized he needed to invert Boltzmann's procedure. Instead of putting distinguishable atoms into a series of energy buckets, he needed to distribute a series of finite but *indistinguishable* energy elements over the oscillators of the cavity material. This is still a statistical approach—it still requires counting all the different possible ways in which the energy elements can be distributed—but, make no mistake, it is very different from the approach used by Boltzmann.

For simplicity we'll continue to label the energy elements as E . Suppose we need to distribute four energy elements over three oscillators. We now find there are fifteen possible ways we can do this—see Fig. 1.3. We can put all the elements in the first oscillator, and none in the other two, a permutation which we can write as $(4E, 0, 0)$. Other permutations are $(3E, E, 0)$, $(2E, E, E)$, $(E, 2E, E)$, and so on. In general, the number of different possible permutations is

$$W = \frac{(N - 1 + P)!}{(N - 1)!P!}, \quad (1.23)$$

where P is the number of energy elements. For the example in which we distribute four energy elements ($P = 4$) over three oscillators ($N = 3$), we get $W = 6!/2!4! = 15$.

We can see why Planck was obliged to reach for this method of counting. In a practical physical system N and P are very large numbers, much larger than 1. We can therefore simplify the expression for W and apply Stirling's rule to give:

$$W \cong \frac{(N + P)!}{N!P!} \cong \frac{(N + P)^{N+P}}{N^N P^P}. \quad (1.24)$$

Note that the terms e^{N+P} cancel. The resemblance between Eq. (1.24) with the term in square brackets in Eq. (1.21) is now starting to become clearer.

If we're distributing P indistinguishable energy elements E , then, logically, the total energy available is simply P times E . Likewise, N oscillators each with an average internal energy U will have a total energy N times U . We conclude that

$$PE = NU \quad \text{or} \quad P = \frac{NU}{E}. \quad (1.25)$$

$$\frac{4E}{3 \text{ oscillators}}$$


4E	0	0
3E	E	0
3E	0	E
2E	2E	0
2E	E	E
2E	0	2E
E	3E	0
E	2E	E
E	E	2E
E	0	3E
0	4E	0
0	3E	E
0	2E	2E
0	E	3E
0	0	4E

$$W = 15$$

Figure 1.3 Instead of putting distinguishable atoms into a series of energy buckets, Planck found that he needed to distribute a series of finite but indistinguishable energy elements over the oscillators of the cavity material. This is a very different counting procedure.

Substituting for P in Eq. (1.24) gives

$$W = \frac{\left(N + \frac{NU}{E}\right)^{(N+NU/E)}}{N^N \left(\frac{NU}{E}\right)^{NU/E}} = \frac{N^{N(1+U/E)} \left(1 + \frac{U}{E}\right)^{N(1+U/E)}}{N^{N(1+U/E)} \left(\frac{U}{E}\right)^{NU/E}} = \frac{\left(1 + \frac{U}{E}\right)^{N(1+U/E)}}{\left(\frac{U}{E}\right)^{NU/E}}. \quad (1.26)$$

We now insert this into Boltzmann's equation for the entropy, Eq. (P.26):

$$S_N = Nk_B \ln \left[\frac{\left(1 + \frac{U}{E}\right)^{1+U/E}}{\left(\frac{U}{E}\right)^{U/E}} \right]. \quad (1.27)$$

This is an expression for the entropy of the oscillators derived from statistics. All we need to do now is compare this with the equation we got from thermodynamics and Planck's radiation law, Eq. (1.21). This is the final step.

Step (3): Make the Comparison

This couldn't be simpler. Equations (1.21) and (1.27) are both expressions for the total entropy of the oscillators in the cavity material at thermal equilibrium. By making use of the combinatorial expression for indistinguishable energy elements, Planck was led inexorably to the conclusion that these elements must have fixed values determined by the relation

$$E = hv. \quad (1.28)$$

Planck called the energy elements *quanta*. 'The constant . . . which is independent of oscillator characteristics, I designated by h and since it has the dimensions of a product of energy and time, I called it the elementary quantum of action or element of action in contrast with the energy element hv .'¹¹

Enter Einstein

Planck's derivation heralded the very beginning of the quantum revolution, but only in promise, not in deed. It certainly didn't change the interpretation of physics overnight. Although Planck was now ready to embrace the idea of atoms and the implications of Boltzmann's statistical approach to calculating thermodynamic quantities (with all its discomfiting consequences), as far as he was concerned 'quantization' was a curious property of the oscillators in the cavity material, emitting or absorbing energy only in discrete quanta, hv . He persisted with the assumption that the energy of the cavity radiation is continuously variable.

Over the next few years he made several attempts to refine his derivation and make it more rigorous, but there was little he could do. A structure that was based entirely on classical, continuous, physics had sprung a profoundly non-classical, discontinuous result. It's not surprising that this was not recognized for what it was for many years.

The quantum revolution really began in earnest with the help of a 'technical expert, third class', based at the Swiss Patent Office in Bern. Einstein opened the door to a completely new interpretation in a paper he published during his 'miracle year' of 1905. Einstein distrusted Planck's derivation, and instead posed a seemingly absurd question. What if the cavity radiation is itself composed of discrete quanta—with energy hv —emitted and absorbed discontinuously by the cavity material? What if it is the radiation that is 'quantized'? This is Einstein's 'light-quantum' hypothesis:¹²

If monochromatic radiation (of sufficiently low density) behaves . . . as though the radiation were a discontinuous medium consisting of energy quanta of magnitude $[hv]$, then it seems reasonable to investigate whether the laws governing the emission and transformation of light are also constructed such as if light consisted of such energy quanta.

He went on in the same paper to predict the outcomes of experiments on the photoelectric effect. This results from shining light on the surfaces of certain metals. Light with

frequencies above a threshold characteristic of the metal (called the work function) will cause electrons to be ejected from the surface. This was a bit of a challenge for the wave theory of light, as the energy in a classical wave depends on its intensity (related to the wave amplitude, the height of its peaks, and depth of its troughs), not its frequency. The higher the wave, the larger its classical energy.

But Einstein figured that if light actually consists of self-contained bundles of energy, with the energy of each bundle proportional to frequency, then the puzzle is solved. Light-quanta with low frequencies don't have enough energy to overcome the work function and dislodge the electrons. As the frequency is increased, the threshold is crossed and the absorption of a light-quantum knocks an electron out of the lattice of metal ions at the surface. Increasing the intensity of the light simply increases the number (but not the energies) of the light-quanta incident on the surface. He went on to make some simple predictions that could be tested in future experiments.

These were highly speculative ideas, and physicists did not rush to embrace them. Fortunately for Einstein, his work on the special theory of relativity, published in the same year, was better regarded. It was greeted as a work of genius (and we'll look at some aspects of the special theory in Chapter 2).

When in 1913 Einstein was recommended for membership of the prestigious Prussian Academy of Sciences, its leading members—Planck among them—acknowledged his remarkable contributions to physics. They were prepared to forgive his lapse of judgement over light-quanta: 'That he may sometimes have missed the target in his speculations, as, for example, in his hypothesis of light-quanta, cannot be really held against him, for it is not possible to introduce really new ideas even in the most exact sciences without sometimes taking a risk.'¹³

Einstein's boldness was rewarded just two years later, when American physicist Robert Millikan reported the results of further experiments on the photoelectric effect. Einstein's predictions were all borne out, earning him the 1921 Nobel Prize for physics. It's a curious twist of history that Einstein should win the Nobel Prize for his pioneering (and, as some had argued, misjudged) work on the light-quantum hypothesis rather than his work on relativity, especially given Einstein's later misgivings about quantum mechanics. But this didn't stop him from presenting a prize lecture on relativity in Stockholm the following year.¹⁴

Despite all the evidence in favour of the wave nature of light, which was certainly not invalidated by these new developments, Einstein was right. The cavity radiation—and indeed all electromagnetic radiation—consists of discrete energy quanta for which the American chemist Gilbert Lewis coined the name 'photon' in 1926. For this reason, Eq. (1.28) is often referred to as the *Planck–Einstein relation*.

.....

NOTES

1. See J. L. Heilbron, *The Dilemmas of an Upright Man: Max Planck and the Fortunes of German Science*, Harvard University Press, Cambridge, MA, 1996, p. 10.
2. Max Planck, letter to Robert Williams Wood, 7 October 1931. Quoted in Armin Hermann, *The Genesis of Quantum Theory (1899–1913)*, MIT Press, Cambridge, MA, 1971, p. 1.

3. Heilbron, *Dilemmas of an Upright Man*, p. 9.
4. Max Planck, *Verhandl. Der Deutsche Physikalische Gesellschaft*, 2, 1900, p. 202. Quoted in Thomas S. Kuhn, *Black-Body Theory and the Quantum Discontinuity 1894–1912*. University of Chicago Press, Chicago, 1978, p. 97.
5. Max Planck, *Physikalische Abhandlungen und Vorträge*, Vol. 3, Vieweg, Braunschweig, 1958, p. 263. Quoted in Hermann, *Genesis of Quantum Theory*, p. 15.
6. Planck, *Physikalische Abhandlungen*, p. 260. Quoted in Hermann, *Genesis of Quantum Theory*, p. 7.
7. Max Planck, Nobel Lecture, 2 June 1920, in *Nobel Lectures: Physics 1901–1921*, Elsevier, Amsterdam, 1967.
8. Max Planck, letter to Robert Williams Wood, 7 October 1931. Quoted in Hermann, *Genesis of Quantum Theory*, p. 1.
9. Abraham Pais, *Subtle is the Lord: The Science and the Life of Albert Einstein*. Oxford University Press, Oxford 1982, p. 370.
10. Planck, Nobel Lecture, 2 June 1920.
11. Max Planck, *Physikalische Abhandlungen und Vorträge*, Vol. 3, Vieweg, Braunschweig, 1958, p. 266. Quoted in Hermann, *Genesis of Quantum Theory*, p. 19.
12. Albert Einstein, *Annalen der Physik*, 17(1905), 143–4. English translation quoted in John Stachel (ed.), *Einstein's Miraculous Year: Five Papers that Changed the Face of Physics*, Princeton University Press, Princeton, NJ, 2005, p. 191.
13. G. Kirsten and H. Körber, *Physiker über Physiker*, Akademie Verlag, Berlin, 1975, p. 201. Quoted in Pais, *Subtle is the Lord*, p. 382.
14. Pais, *Subtle is the Lord*, p. 114.

.....

FURTHER READING

- Cushing, James T., *Philosophical Concepts in Physics: The Historical Relation between Philosophy and Scientific Theories*, Cambridge University Press, Cambridge, UK, 1998, pp. 273–8.
- Farmelo, Graham, 'A Revolution with No Revolutionaries: The Planck–Einstein Equation for the Energy of a Quantum', in Graham Farmelo (ed.), *It Must Be Beautiful: Great Equations of Modern Science*, Granta, London, 2002, pp. 1–27.
- Gamow, George, *Thirty Years that Shook Physics*, Dover, New York, 1966, Chapter I.
- Heilbron, J. L., *The Dilemmas of an Upright Man: Max Planck and the Fortunes of German Science*, Harvard University Press, Cambridge, MA, 1996.
- Hermann, Armin, *The Genesis of Quantum Theory (1899–1913)*. MIT Press, Cambridge, MA., 1971.
- Kennedy, Robert E., *A Student's Guide to Einstein's Major Papers*, Oxford University Press, Oxford, 2012, Chapter 6.
- Kuhn, Thomas S., *Black-Body Theory and the Quantum Discontinuity 1894–1912*. University of Chicago Press, Chicago, 1978.
- Pais, Abraham, *Subtle is the Lord: The Science and the Life of Albert Einstein*. Oxford University Press, Oxford, 1982.
- Stachel, John (ed.), *Einstein's Miraculous Year: Five Papers that Changed the Face of Physics*, Princeton University Press, Princeton, NJ, 2005, pp. 165–98.

Einstein's Derivation of $E = mc^2$

The Equivalence of Mass and Energy

Quantum mechanics and relativity are often taught separately as distinct subjects, but they are closely interrelated, and Einstein's special theory of relativity was a critically important ingredient in the early development of quantum theory in the 1920s. Specifically, Louis de Broglie used it as a basis for his derivation of the relationship between (quantum) wavelength and linear momentum, $\lambda = h/p$, which we will go on to consider in Chapter 4. I would argue that it does no harm to be familiar with the logic behind one of the most important conclusions arising from special relativity: the equivalence of mass (m) and energy (E), expressed in Einstein's most famous equation $E = mc^2$, where c is the speed of light. So, I've chosen to include Einstein's derivation here.

As we saw in the Prologue, the physicists of the early twentieth century inherited a couple of rather splendid structures—classical mechanics and electrodynamics. These were (still are) like fine, five-star hotels, full of character and ornament. But they were also built on surprisingly unsteady foundations. Hamiltonian mechanics did not rid the structure of its dependence on Newton's absolute space and time. And Maxwell's electrodynamics—with all its electromagnetic waves—demanded an ether that could not be found, no matter how hard the physicists looked for it.

The negative results of the Michelson–Morley experiments were a real puzzle. Irish physicist George FitzGerald (in 1889) and Dutch physicist Hendrik Lorentz (in 1892) independently suggested that these could be explained if the interferometer was assumed to be *physically contracting* along its length in response to pressure from the ether wind. In order to return a value for the speed of light unchanged by a change in measurement direction, the contraction had to have a very specific magnitude. If the 'proper' length of the path in the interferometer is l_0 , it was possible to reproduce the results if this was assumed to contract to a length $l = l_0/\gamma$, where γ is now known as the *Lorentz factor*, given by $1/\sqrt{1 - v^2/c^2}$, in which v is the speed of the interferometer relative to the stationary ether and c is the speed of light. It then wouldn't matter if the speed of light was different along different paths in the interferometer. Any such difference would be compensated by a *Lorentz–FitzGerald contraction*, ensuring that the light waves remained in phase. There would be no interference.

But this assumes that physics conspires to prevent the experimental detection of the ether. Wouldn't it just be simpler to take the results at face value and conclude that the ether doesn't exist? Einstein certainly thought so, and he suspected that the non-existence of the ether was closely associated with the non-existence of absolute space and time. He wrote:¹

Examples of this sort [electromagnetic phenomena], together with the unsuccessful attempts to detect a motion of the earth relative to the 'light medium' [ether], lead to the conjecture that not only the phenomena of mechanics but also those of electrodynamics have no properties that correspond to the concept of absolute rest.

He judged that the solution to these thorny problems demanded a firmly practical and pragmatic approach in which the 'observer' takes centre stage. Here 'observer' doesn't necessarily mean a human observer. It means that, to understand the physics correctly, we must accept that this is physics as seen from the perspective of someone or something *inside* the reality that is being observed or measured, with the aid of a ruler and a clock. Such an observer is implicit in the physics of Newton. But Newton's laws are formulated as though the observer is somehow imagined to sit outside of the reality in which all the action is taking place.

Einstein began by stating two basic principles. The first, which he called the *principle of relativity*, says that observers who find themselves in relative motion at different (but constant) speeds *must* make observations or measurements that conform to the laws of physics. Put another way, the laws of physics must be the same for everyone, irrespective of how fast they're moving relative to what they're observing or measuring (or the other way around). This seems unremarkable. Surely, this is what it means for a relationship between physical properties to be a 'law'? But accepting this means accepting that there can be no privileged or absolute frame of the reference. And therefore there can be no ether.

The second principle relates to the speed of light. In Newton's mechanics, speeds are additive. An object rolling along the deck of a ship as the ship plows across the Atlantic Ocean is moving with a total speed given by the speed of the roll along the deck plus the speed of the ship.* But light doesn't seem to obey this rule. The conclusion drawn from the Michelson–Morley experiment is that light always travels at the same speed. The light emitted from a flashlight moves away at the speed of light, c . Light from the same flashlight lying on the deck of the ship still moves at the speed of light, not c plus the speed of the ship.

Instead of trying to figure out *why* the speed of light is constant, Einstein simply accepted this as an established fact and proceeded to work out the consequences. He discovered that one immediate consequence of a fixed speed of light is that there can be no such thing as absolute time.

* Plus the speed of rotation of the Earth, plus the speed of Earth's orbit around the Sun, plus the speed of the Solar System's rotation around the centre of the Milky Way galaxy, plus the speed of the motion of the Milky Way relative to the Local Group of galaxies, . . . You get the basic idea.

The Relativity of Simultaneity

Suppose you observe a remarkable phenomenon (Fig. 2.1). During a heavy thunderstorm you see two bolts of lightning strike the ground simultaneously, one to your left and one to your right. You're standing perfectly still, so the fact that it takes a small amount of time for the light to reach you is of no real consequence. Light travels very fast so, as far as you're concerned, you see both bolts at the instant they strike. But, unlike you, I'm moving from left to right at a substantial fraction of the speed of light. In the time it takes for the light from the right-hand bolt to reach me, I've now moved a few steps closer to it. In contrast, in the time it takes for the left-hand bolt to reach me, I've moved a few steps away from it. The upshot is that the light from the right-hand bolt reaches me first.

You see the lightning bolts strike simultaneously. I don't. Who is right?

We're both right. The principle of relativity demands that the laws of physics must be the same for everyone, irrespective of the relative motion of the observer, and we can't use physical measurements to tell whether it is you or me who is in motion.

We have no choice but to conclude that there is no such thing as absolute simultaneity. There is no definitive or privileged inertial frame of reference in which we can declare that these things happened at precisely the same time. They may happen simultaneously in this frame or they may happen at different times in a different frame, and all such frames

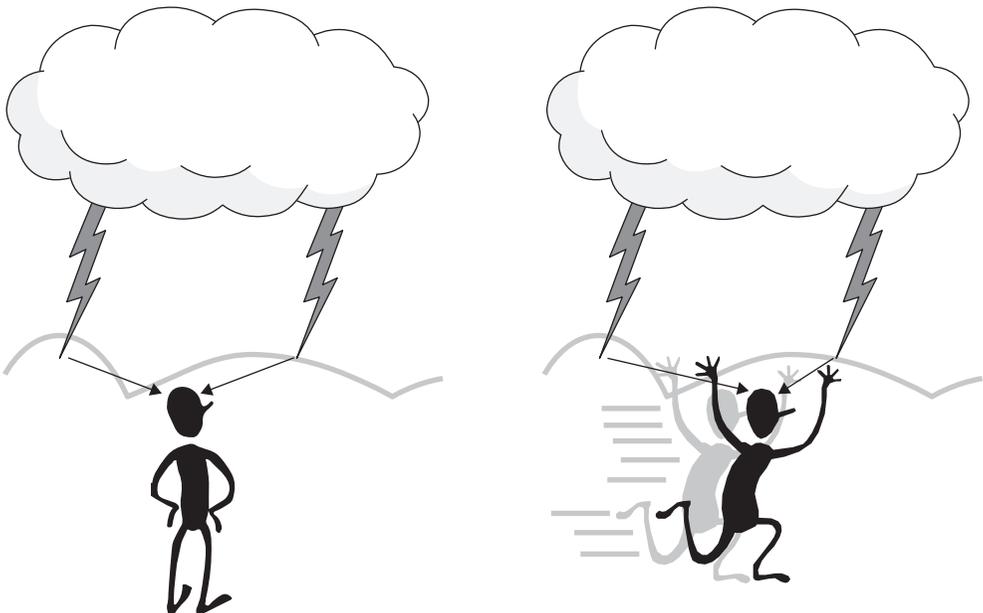


Figure 2.1 *The stationary observer (left) sees the lightning bolts strike simultaneously, but an observer moving at a considerable fraction of the speed of light (right), sees the right-hand bolt strike first.*

are equally valid. Consequently, there can be no ‘real’ or absolute time. We perceive events differently because time is relative.

The Relativity of Times and Distances

To get some sense for what this statement means it is helpful to take a quick look at a ‘thought experiment’. Imagine we’re travelling on a train together. It is night, and there is no light in the carriage. We fix a small flashlight to the floor of the carriage and a large mirror on the ceiling. The light flashes once, and the flash is reflected from the mirror and detected by a small light-sensitive cell or photodiode placed on the floor alongside the flashlight. Both flashlight and photodiode are connected to an electronic box of tricks that allows us to measure the time interval between the flash and its detection.

We make our first set of measurements whilst the train is stationary, alongside the station platform, and measure the time taken for the light to travel upwards from the flashlight, bounce off the mirror, and back down to the photodiode. Let’s call this time interval t_0 . The light travels a total distance of $2d_0$, where d_0 is the height of the carriage. We know that $2d_0 = ct_0$, where c is the speed of light, as the light travels up (d_0) and then down (another d_0) in the total time t_0 at the speed c (see Fig. 2.2).

You now step off the train onto the platform and repeat the measurement as the train moves past from left to right with velocity v , where v is a substantial fraction of the speed of light. Of course, trains can’t move this fast in real life, but that’s okay because this is only a thought experiment.

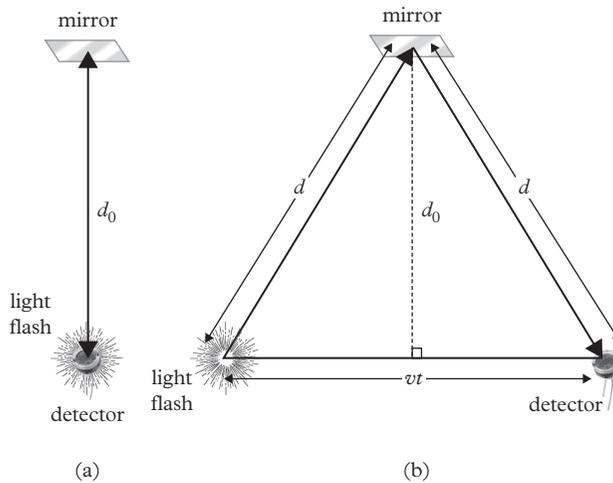


Figure 2.2 (a) When the train is stationary the light flash travels straight up and down, a total distance of $2d_0$. (b) When the train is moving with velocity v , from the perspective of a stationary observer on the platform the light flash travels upwards and downwards to the detector as the train is moving forwards, describing a triangular path with a total distance of $2d$.

Now from your vantage point on the platform you see something rather different. The light no longer travels straight up and down. At a certain moment the light flashes, and in the small (but finite) amount of time it takes for the light to travel upwards towards the ceiling, the train moves forward, from left to right. It continues to move forward as the light travels back down to the floor to be detected by the photodiode. From your perspective on the platform the light path now looks like a ' Λ ', a Greek capital lambda or an upside-down ' V '.

Let's assume that the total time required for the light to travel this longer path is t . In the time taken for the light to travel up a distance d and down another distance d , the train has moved forward a total distance vt , which forms the base of the equilateral triangle formed by the Λ path. If we now draw a perpendicular (which will obviously have length equal to d_0) from the apex of the triangle and which bisects the base we can use Pythagoras' theorem:

$$d^2 = d_0^2 + \left(\frac{1}{2}vt\right)^2. \quad (2.1)$$

But we know that $d = \frac{1}{2}ct$ (the speed of light is the same on the moving train) and from our earlier measurement $d_0 = \frac{1}{2}ct_0$ so we have

$$\left(\frac{1}{2}ct\right)^2 = \left(\frac{1}{2}ct_0\right)^2 + \left(\frac{1}{2}vt\right)^2 \quad (2.2)$$

We can now cancel all the factors of $\frac{1}{2}$ and multiply out the brackets to give

$$c^2t^2 = c^2t_0^2 + v^2t^2 \quad (2.3)$$

We gather the terms in t^2 on the left-hand side and divide through by c^2 to obtain:

$$t^2 \left(1 - \frac{v^2}{c^2}\right) = t_0^2 \quad \text{or} \quad t = \gamma t_0 \quad \text{where} \quad \gamma = \frac{1}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}. \quad (2.4)$$

There is only one possible conclusion. From your perspective as a stationary observer standing on the platform, *time is measured to slow down on the moving train*. If the train is travelling at about 86.6% of the speed of light, the factor γ is equal to 2. What took 1 second when the train was stationary now appears to take 2 seconds on the moving train when measured from the platform. In different moving frames of reference, time intervals are 'dilated'.

Similar thought experiments can be concocted to show that the length of an object moving relative to a stationary observer will appear to contract in the direction of travel compared with measurements of an observer 'riding' on the object (in which the object is

then judged to be stationary). It should be no surprise to learn that if the ‘proper’ length of the object is l_0 , then the measured length l is given by l_0/γ , precisely as FitzGerald and Lorentz had demanded. An object moving at 86.6% of the speed of light will be measured to contract to half its proper length. The difference is that the contraction is not physical, in the sense that the length is compressed by the ether wind. It is a consequence of the relativity of space.

This is all a bit disconcerting, and it’s tempting to slip back into older, more comfortable ways of thinking. If this is all about observation and measurement at speeds close to that of light, then surely this must be just a matter of perspective and perception? From the perspective of this inertial frame of reference, time *appears* to slow down and distances *appear* to contract. Surely, time doesn’t really slow down and distances don’t really contract?

But they do. Space and time are relative, not absolute, and there is therefore no unique or ‘correct’ perspective which will give us absolute measures of distance and time. In this stationary frame of reference the distance is this and the time interval is that. But in this moving frame of reference the distance is half this and the time interval is twice that. There is no absolute frame of reference.

The Lorentz Transformation

We don’t need to be constantly referring to thought experiments to work out the consequences of making time and distance measurements in different inertial frames of reference. All these consequences were worked out by Lorentz, FitzGerald, and Joseph Larmor towards the end of the nineteenth century. The recipe for transforming from one inertial frame of reference (S) to another moving relatively to it with velocity v (S') is called the *Lorentz transformations* (a name coined by Poincaré).

We should note in passing that changing perspective from one inertial frame of reference to another is just as much a feature of the mechanics of Galileo and Newton as it is special relativity, and it’s quite instructive to compare the two. Now it’s quite difficult to work simultaneously in three spatial dimensions (x, y, z) *and* a time dimension, so we simplify things a little by constraining S and S' to move relatively to each other in the x -direction only. This particular arrangement is sometimes called the ‘standard configuration’. The transformations then allow us to move from S (distance interval x , time interval t) to S' (x', t') and back from S' to S (see Table 2.1).

We can see from this that the Galilean transformation has no effect on time measurements and simply shifts distance measurements in the x coordinate by an amount that depends on the relative velocity v , as we would expect. But the Lorentz transformation is very different. By assuming the principle of relativity and a fixed speed of light, we introduce the speed of light c into the equations and the Lorentz factor γ . Also notice how the time interval t' now depends on the distance x .

We can check the authenticity of the Lorentz transformation by considering a couple of hypothetical examples of its application. Imagine a point-particle at rest in a frame S which is not moving, and which we call the ‘rest frame’. It experiences two ‘events’. It

Table 2.1 Galilean and Lorentz Transformations in the Standard Configuration

Galilean		Lorentz	
From S to S'	From S' to S	From S to S'	From S' to S
$t' = t$	$t = t'$	$t' = \gamma \left(t - \frac{v}{c^2} x \right)$	$t = \gamma \left(t' + \frac{v}{c^2} x' \right)$
$x' = x - vt$	$x = x' + vt$	$x' = \gamma (x - vt)$	$x = \gamma (x' + vt')$
$y' = y$	$y = y'$	$y' = y$	$y = y'$
$z' = z$	$z = z'$	$z' = z$	$z = z'$

doesn't really matter what these are but suppose it emits a burst of light at time t_1 and emits another burst at time t_2 , and $t = t_2 - t_1$. The particle is at rest in this frame, so $x = 0$ (it doesn't move in any spatial coordinate). An observer in the moving frame of reference S' will measure $t' = \gamma [t - (v/c^2)x]$, or $t' = \gamma t$ (remember, $x = 0$). Time is dilated in the moving frame.

Now imagine an extended object, such as a metal rod, with a length measured to be x in the rest frame S . Lights flash at either end of the rod, at times t_1 and t_2 , with $t = t_2 - t_1$. But this time we've forgotten to bring a clock, so we have no way of measuring the time interval between flashes, t . No matter, because we can arrange to observe the rod in a frame of reference moving with velocity v in the x -direction, such that the two flashes are measured to be simultaneous. This defines the moving frame S' . We know that in this frame $t' = 0$ (as there's no time difference between the flashes). We can then use the Lorentz transformation from S' to S to deduce that $x = \gamma (x' + vt')$, or $x = \gamma x'$ (remember, $t' = 0$). We conclude that $x' = x/\gamma$. In the moving frame of reference the measured length of the rod contracts.

On the Electrodynamics of Moving Bodies

Einstein published two papers on special relativity in 1905. The first, titled 'On the Electrodynamics of Moving Bodies', was published in June. This is the principal paper in which Einstein explores the implications of the principle of relativity and a fixed speed of light on the form of Maxwell's equations for electromagnetism in a vacuum. We don't need to be detained by the details, but there is one result that Einstein derived in this paper that will be important in what follows.

Suppose that a distant source emits electromagnetic radiation which, from our perspective, can be satisfactorily approximated as a plane wave. We further suppose that we can represent either or both of the electric or magnetic components of the radiation as a simple sine wave, such as $A \sin \Phi$, where A is the amplitude and Φ is the phase angle. The wave moves towards us from an arbitrary direction (Fig. 2.3). Let's call this direction r (so that the wave is moving in the negative r -direction). In the rest frame S we can then

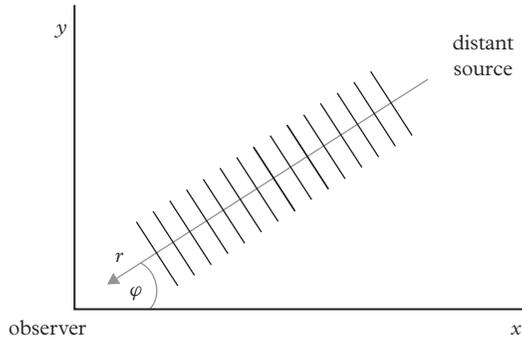


Figure 2.3 *A plane electromagnetic wave from a distant source moves towards the observer from a direction r , which makes an angle φ to the x -direction.*

represent the wave as $A \sin(\omega t - kr)$, where ω is the angular frequency ($2\pi\nu$, where ν is the frequency), and k is the wave vector ($2\pi/\lambda$, where λ is the wavelength). The direction r is quite arbitrary, but we can rewrite the phase angle as

$$\Phi = \omega t - k(a_x x + b_y y + c_z z), \quad (2.5)$$

where a_x , b_y , and c_z are direction cosines which map the coordinate axes x , y , and z on to r . Obviously, ω and k are interrelated, and for an electromagnetic wave travelling at the speed of light, $k = \omega/c$ and

$$\Phi = \omega t - \frac{\omega}{c}(a_x x + b_y y + c_z z). \quad (2.6)$$

Let's now look at this wave from the perspective of a reference frame S' moving with velocity v in the positive x -direction (in other words, the standard configuration). We know that times will dilate and distances will contract, and we're interested to discover how this affects the measurement of the *frequency* of the wave. We anticipate that the phase angle in the moving frame is now

$$\Phi' = \omega' t' - \frac{\omega'}{c}(a'_x x' + b'_y y' + c'_z z'). \quad (2.7)$$

The speed of light is the same in the moving frame, but we also know that the phase angles in (2.6) and (2.7) must be identical, $\Phi = \Phi'$, since we learned from the Michelson–Morley experiment that there is no interference resulting from measurements in different inertial frames of reference, and this can only happen if not only the speeds but the phase angles of the waves are preserved. Time dilation is compensated by distance contraction, such that the phase angle is unaffected.

We now apply the Lorentz transformation to the expression for the phase angle in the moving frame, Φ' :

$$\Phi' = \omega' \gamma \left(t - \frac{v}{c^2} x \right) - \frac{\omega'}{c} [\gamma a'_x (x - vt) + b'_y y + c'_z z]. \quad (2.8)$$

If we gather together the terms involving t and x we have

$$\Phi' = \omega' \gamma \left(1 + \frac{v}{c} a'_x \right) t - \frac{\omega'}{c} \left[\gamma \left(a'_x + \frac{v}{c} \right) x + b'_y y + c'_z z \right]. \quad (2.9)$$

Comparing (2.9) with (2.6) term by term suggests that

$$\omega' \gamma \left(1 + \frac{v}{c} a'_x \right) t = \omega t \quad (2.10)$$

$$\frac{\omega' \gamma}{c} \left(a'_x + \frac{v}{c} \right) x = \frac{\omega}{c} a_x x \quad (2.11)$$

$$\frac{\omega'}{c} b'_y y = \frac{\omega}{c} b_y y \quad (2.12)$$

$$\frac{\omega'}{c} c'_z z = \frac{\omega}{c} c_z z \quad (2.13)$$

We can rearrange (2.11) to give an expression for a'_x ,

$$a'_x = \frac{\omega}{\omega' \gamma} a_x - \frac{v}{c}, \quad (2.14)$$

which we can substitute into (2.10) to give, after some manipulation (remember $\gamma^2 = 1/(1 - v^2/c^2)$),

$$\omega' = \omega \gamma \left(1 - \frac{v}{c} a_x \right). \quad (2.15)$$

This is actually the result we need. But for completeness we can use this to find expressions for all the direction cosines in the moving frame:

$$a'_x = \frac{a_x - v/c}{\left(1 - \frac{v}{c} a_x \right)} \quad (2.16)$$

$$b'_y = \frac{b_y}{\gamma \left(1 - \frac{v}{c} a_x \right)} \quad (2.17)$$

$$c'_z = \frac{c_z}{\gamma \left(1 - \frac{v}{c} a_x \right)}. \quad (2.18)$$

Our next step is to divide both sides of Eq. (2.15) by 2π and note from Fig. 2.3 that $a_x = \cos \varphi$, giving

$$\nu' = \nu\gamma \left(1 - \frac{v}{c} \cos \varphi\right). \quad (2.19)$$

As Einstein noted in his paper, this last expression is actually a summary of the *transverse* Doppler Effect. If we arrange for the light wave to travel directly along the x -axis (such that $\varphi = 0$, $\cos \varphi = 1$), we can see that

$$\nu' = \nu\gamma \left(1 - \frac{v}{c}\right) = \nu \frac{\left(1 - \frac{v}{c}\right)}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}} = \nu \frac{\sqrt{\left(1 - \frac{v}{c}\right)\left(1 - \frac{v}{c}\right)}}{\sqrt{\left(1 + \frac{v}{c}\right)\left(1 - \frac{v}{c}\right)}} = \nu \sqrt{\frac{\left(1 - \frac{v}{c}\right)}{\left(1 + \frac{v}{c}\right)}}. \quad (2.20)$$

This is just the Doppler shift formula. In the moving frame of reference, the frequency of the light is increased, or ‘blue-shifted’, as it moves towards the observer, ‘red-shifted’ if the light is moving away. The equivalent phenomenon with sound waves will be familiar to anyone who has listened to the siren of an ambulance or police car as it speeds past.

Now we’re building up to a derivation of the iconic formula $E = mc^2$, so we’re obviously interested in energy, not frequency. But that’s okay, because we simply reach for the Planck–Einstein equation from Chapter 1 and set $E = h\nu$ giving, from (2.19),

$$E' = E\gamma \left(1 - \frac{v}{c} \cos \varphi\right). \quad (2.21)$$

The energy of a plane light wave moving towards the observer is *increased* in a moving frame of reference, as indeed we would expect from Eq. (2.19).

In fact, in his original paper Einstein chose not to take this step. Instead, he presented a rather more elaborate derivation based on the energy of a spherical surface (think of this as a spherical ‘sample’) within the light wave, which squashes to an ellipsoid in the moving frame of reference as its dimension in the direction of travel contracts. He arrived at precisely the same expression for E' , and observed: ‘It is noteworthy that the energy and the frequency of a light complex vary with the observer’s state of motion, according to the same law.’² Comparison of (2.19) and (2.21) lead directly to the conclusion $E \propto \nu$.

Einstein had published his paper on the light-quantum hypothesis in April that year (see Chapter 1), but at the time this was a very radical proposal. He obviously preferred to play safe here and stick with a purely classical wave description of light, avoiding ‘contaminating’ his development of relativity with quantum considerations. His scientific biographer, Abraham Pais, later observed:³

he rightly regarded his own quantum hypotheses of 1905 more of a new phenomenological description than a new theory, in sharp contrast to his relativity theory, which he rightly regarded as a true theory with clearly defined first principles.

We, of course, have the benefit of hindsight and would not now hesitate to apply the Planck–Einstein result. But, however we got here, we now have all the ingredients we need. In September 1905, Einstein published a short addendum to his June paper on special relativity. This is what he did.

The Ingredients

1. A thought experiment in which an object emits two bursts of light simultaneously in opposite directions.
2. The expression for the relativistic energy of a plane light wave derived in the June special relativity paper, Eq. (2.21).
3. The law of conservation of energy.
4. The Taylor series expansion for $1/\sqrt{1+x}$.

The Recipe

We've seen from the discussion above that assuming the principle of relativity and a fixed speed of light have all kinds of implications for our understanding of space and time and, in consequence, for all the physical parameters we measure in space and time, such as frequency and energy. That energy will be measured to be larger from the perspective of a moving frame of reference causes us to pause and reflect on the law of conservation of energy. If the energy of an object or a wave is measured to be larger, and the laws of physics are the same for everyone, where then does that additional energy come from?

To explore these implications a little further, Einstein needed a thought experiment which would help to simplify the mathematics and give him access to the before-and-after *differences* in energy in two inertial frames of reference. From these differences, and the law of conservation of energy, he could discover the origin of the 'extra' energy in the moving frame.

We begin in Step (1) by analysing the thought experiment and applying the law of conservation of energy and the expression for the relativistic energy of a plane light wave—Eq. (2.21)—to deduce the relationship between the initial and final energies of the object in both a stationary and moving frame of reference. We discover that the energy difference is slightly greater (γE) in the moving frame than in the stationary frame (E). In Step (2) we trace this energy difference to a difference in the kinetic energies of the object in the two frames, and we note that as the energy carried away by the light bursts increases in the moving frame, so the kinetic energy of the object falls. In Step (3) we simplify the expression for kinetic energy by applying a Taylor series expansion for the Lorentz factor, γ . Finally, in Step (4) we show that the decline in kinetic energy derives not from a decline in the object's velocity (which we might have expected), but from a decline in its *mass*, which falls by an amount $m = E/c^2$. We can, of course, rearrange this to give $E = mc^2$.

Step (1): The Thought Experiment

Einstein imagined an object, at rest in one frame (S) and moving with uniform motion v in another (S'). He did not specify the nature of the object, but if it helps to follow the logic in the following we can suppose it is an atom. It emits two bursts of light, each of energy $\frac{1}{2}E$, in opposite (but arbitrary) directions r . Once again, r makes an angle φ with the x -axis; see Fig. 2.4.

Let's suppose that the energy of the object in the rest frame S before the emission of the light waves is E_i (where the subscript i stands for 'initial'). The energy of the object after emitting the light bursts is E_f ('final'). Conservation of energy then demands that

$$E_i = E_f + \frac{1}{2} \vec{E} + \frac{1}{2} \overleftarrow{E} \quad \text{or} \quad E_i = E_f + E. \tag{2.22}$$

In (2.22) I've used an arrow notation merely to indicate that the two lots of energy $\frac{1}{2}E$ are emitted in opposite directions. The arrows have no other significance: $\frac{1}{2} \vec{E}$ is equal to $\frac{1}{2} \overleftarrow{E}$.

Let's now examine this system from the perspective of the moving frame S' . We know from Einstein's earlier result (2.21) that the energies of the emitted light waves are no longer $\frac{1}{2}E$, but rather $\frac{1}{2}E'$, where

$$\frac{1}{2} \vec{E}' = \frac{1}{2} \vec{E} \gamma \left(1 + \frac{v}{c} \cos \varphi \right) \tag{2.23}$$

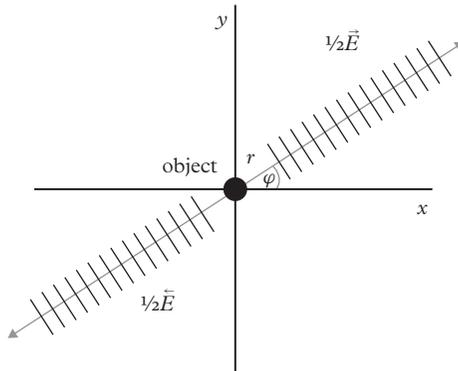


Figure 2.4 In Einstein's thought experiment, an object emits two bursts of light of equal energy $\frac{1}{2}E$ in opposite directions which form an angle φ to the x -axis.

for the light wave emitted away from the observer in the positive r -direction,* and

$$\frac{1}{2} \overleftarrow{E}' = \frac{1}{2} \overleftarrow{E} \gamma \left(1 - \frac{v}{c} \cos \varphi\right) \quad (2.24)$$

for the light wave emitted in the opposite direction, towards the observer. Denoting the initial and final energies of the object in the moving frame as E'_i and E'_f , respectively, we again apply the law of conservation of energy and obtain

$$\begin{aligned} E'_i &= E'_f + \frac{1}{2} \overrightarrow{E}' + \frac{1}{2} \overleftarrow{E}' \quad \text{or} \\ E'_i &= E'_f + \frac{1}{2} \overrightarrow{E} \gamma \left(1 + \frac{v}{c} \cos \varphi\right) + \frac{1}{2} \overleftarrow{E} \gamma \left(1 - \frac{v}{c} \cos \varphi\right). \end{aligned} \quad (2.25)$$

By setting up the thought experiment in this way, we see that the terms in $\cos \varphi$ cancel, leaving us with

$$E'_i = E'_f + \gamma E. \quad (2.26)$$

The Lorentz factor is a positive quantity that is greater than or equal to 1. This means that the energy emitted in the bursts of light in the moving frame of reference is greater or at least equal to the energy emitted in the rest frame.

Step (2): Work out the Difference in Kinetic Energy in the Two Frames

We can rearrange the above expressions for E_i (2.22) and E'_i (2.26) as follows:

$$E_i - E_f = E \quad \text{and} \quad E'_i - E'_f = \gamma E. \quad (2.27)$$

Subtracting the first of these expressions from the second gives

$$(E'_i - E'_f) - (E_i - E_f) = E(\gamma - 1) \quad (2.28)$$

Let's ponder on this result for a while. In situations where the relative velocity of the moving frame of reference is very much smaller than the speed of light, v/c is very much smaller than 1 and the Lorentz factor γ tends towards a low-velocity limiting value of 1 (see Fig. 2.5). In these circumstances, the change in the initial and final energy differences in moving from one frame to another is zero—the energies of the emitted light waves are the same in both frames.

* The sign in the bracket changes because this wave is moving *away* from the observer in the positive r -direction, $A \sin(kr - \omega t)$.

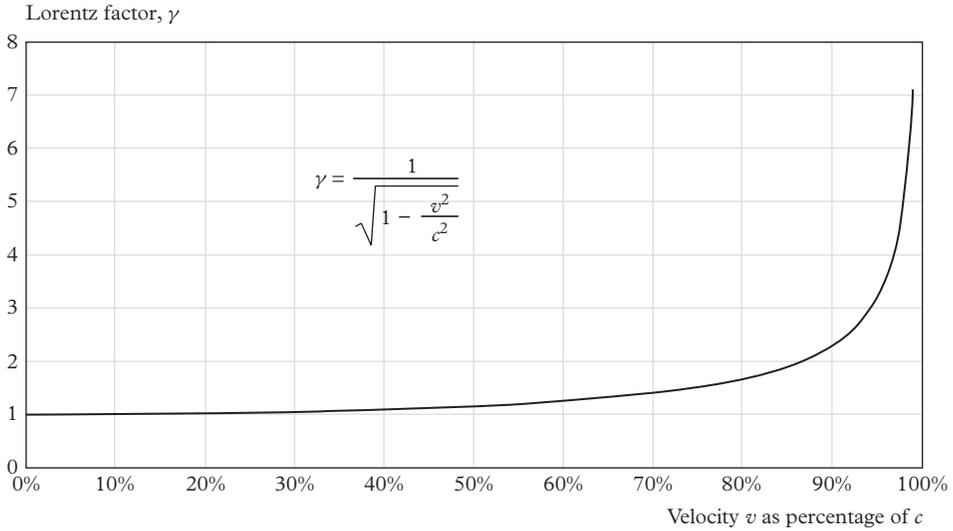


Figure 2.5 Variation of the Lorentz factor as a function of v/c .

But as the relative velocity increases, γ becomes larger than 1 and the change in the initial and final energy differences becomes significant. For example, if v is 86.6% of the speed of light, we know that $\gamma = 2$ and the energy of the light bursts in the moving frame is therefore $2E$, compared with E in the rest frame. The total energy (object plus light bursts) must be conserved, so *where does this additional energy come from?*

Now, the only difference between the two frames is that one is *moving* relative to the other, so it seems perfectly logical to trace the origin of this additional energy to the difference in the *kinetic energy* of the object in the two frames.

We set ΔT_i as the *difference* in the object’s kinetic energy in the two frames *before* it emits the light bursts, such that

$$\Delta T_i = E'_i - E_i. \tag{2.29}$$

This simply represents the increase in the object’s kinetic energy before emission of the light associated with the switch to the moving frame. Likewise,

$$\Delta T_f = E'_f - E_f. \tag{2.30}$$

We now see that

$$\begin{aligned} \Delta T_f - \Delta T_i &= (E'_f - E_f) - (E'_i - E_i) \quad \text{or} \\ \Delta T_f - \Delta T_i &= (E_i - E_f) - (E'_i - E'_f) = -E(\gamma - 1). \end{aligned} \tag{2.31}$$

In other words, in the moving frame of reference the change in kinetic energy of the object *falls* (ΔT_f is *less* than ΔT_i) by an amount $E(\gamma - 1)$. So, here's the answer. The additional energy carried away by the light waves in the moving frame of reference is compensated by a corresponding fall in the object's kinetic energy. Energy is indeed conserved.

If we halted our analysis at this point and relied on intuition, we might be tempted to conclude that the kinetic energy of the object in the moving frame falls because the object *slows down* on emitting the light bursts.

This is why it's never a good idea to stop before we get right to the end.

Step (3): Simplify the Expression for the Lorentz Factor, γ

The difference $\Delta T_f - \Delta T_i$ represents the extent by which the kinetic energy of the object falls in the moving frame. To get a better understanding of what's going on here we need to dig a little deeper into the expression:

$$\Delta T_f - \Delta T_i = -E(\gamma - 1) = -E \left(\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} - 1 \right). \quad (2.32)$$

This is rather cumbersome. But we can now employ a trick to simplify this a little. Complex functions like γ can be recast as the sum of an infinite series of simpler terms, called a Taylor series (for eighteenth-century English mathematician Brook Taylor). The good news is that for many complex functions we can simply look up the relevant Taylor series. Even better, we often find that the first two or three terms in the series provide an approximation to the function that's good enough for most practical purposes.

The Taylor series in question is

$$\frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 + \frac{35}{128}x^4 - \frac{63}{256}x^5 + \dots \quad (2.33)$$

Substituting $x = -v^2/c^2$ means that terms in x^2 are actually terms in v^4/c^4 and so on for higher powers of x . Now the speed of light c in a vacuum is about $2.998 \times 10^8 \text{ ms}^{-1}$, so c^4 is about $8.074 \times 10^{33} \text{ m}^4\text{s}^{-4}$. As this appears in the divisor the higher-order terms in the expansion will be very small and Einstein was happy to leave them out. This gives rise to a small error on the order of a few per cent for speeds v up to about 50% of the speed of light, but of course the error grows dramatically as v increases further.

Applying this approximation in (2.32) gives

$$\Delta T_f - \Delta T_i = -\frac{1}{2} \left(\frac{E}{c^2} \right) v^2. \quad (2.34)$$

Step (4): The Proof: $E = mc^2$

The final step is now really rather obvious. The classical expression for kinetic energy is $\frac{1}{2}mv^2$, as we saw in the Prologue, Eq. (P.3), where m is the mass of the object. What we notice in Eq. (2.34) is that the velocity v is unchanged—the kinetic energy in the moving frame does not fall because the object slows down: v doesn't depend on E . Instead the kinetic energy falls because the *mass* of the object falls by an amount

$$m = \frac{E}{c^2}. \quad (2.35)$$

Now in his short September addendum to his paper on special relativity, Einstein did not go on to rearrange this last expression to give the iconic formula $E = mc^2$. He was content to conclude:⁴

If a body emits the energy [E] in the form of radiation, its mass decreases by an amount [E/c^2]. Here it is obviously inessential that the energy taken from the body turns into radiant energy, so we are led to the more general conclusion: *The mass of a body is a measure of its energy content.*

And indeed, there are arguments to suggest that this conclusion is by far the more profound.⁵

Variable Energy Content

In September 1905, Einstein was quite doubtful that his 'very interesting conclusion' would have any practical applications, although he did note: 'It is not excluded that it will prove possible to test this theory using bodies whose energy content is variable to a high degree (e.g. radium salts).'⁶ Later that year he wrote to his friend Conrad Habicht: 'The line of thought is amusing and fascinating, but I cannot know whether the dear Lord doesn't laugh about this and has played a trick on me'.⁷

Thirty years later, experimental physics would turn up many examples of bodies whose energy content is variable to a high degree, and the discovery of nuclear fission in late 1938 would expose the possibility of a dreadful practical demonstration of $E = mc^2$. The disintegration of just a small amount of mass from a 56-kg bomb core consisting of 90% pure uranium-235, an isotope present in small quantities of naturally occurring uranium, releases an energy equivalent to 12,500 tons of TNT. This was sufficient to destroy utterly the Japanese city of Hiroshima, on 6 August 1945.

The destructive potential of $E = mc^2$ is clear. But, in fact, our very existence on Earth depends on this equation. In the so-called proton–proton (or p–p) chain which operates at the centres of stars, including the Sun, four protons are fused together in a sequence which produces the nucleus of a helium atom, consisting of two protons and two neutrons. If we carefully add up all the masses of the nucleons involved we discover

a small discrepancy, called the *mass defect*. About 0.7% of the mass of the four protons is converted into about 26 MeV (mega electron-volts) of radiation energy, which when it comes from the Sun we call sunlight. This might not sound too impressive, but this is the energy generated from the fusion of just four protons. It is estimated that in the Sun's core about 4×10^{38} protons react every second, releasing energy equivalent to about 4 million billion billion 100-W light bulbs. Every second.

Needless to say, in the years since its publication, many physicists have picked over Einstein's derivation of this, his most famous equation. Some have criticized the derivation as circular. (For example, by its very nature, the Lorentz transformation assumes a fixed relative velocity v between frames, so by definition this cannot be the source of the decline in kinetic energy in the moving frame.) Others have criticized the critics. It seems that Einstein himself was not entirely satisfied with it and during his lifetime developed other derivations, some of which were variations on the same theme and others involving entirely different hypothetical physical situations.

Despite his efforts, all these different approaches seemed to involve situations that, it could be argued, are rather exceptional or contrived. As such, these are perhaps insufficiently general to warrant declaring $E = mc^2$ to be a deep, fundamental relationship—which Einstein called an 'equivalence'—between mass and energy.

The Role of c

This equation has by now become so familiar that we've likely stopped thinking about where it comes from or what it represents. So, let's pause to reflect further on it here. Perhaps the first question we should ask ourselves concerns the *basis* of the equation. If this is supposed to be a fundamental equation describing the nature of material substance, why is the speed of light c involved in it? What has light (or electromagnetic radiation in general) got to do with matter?

The basic form of $E = mc^2$ appears to tie the relationship between mass and energy to electrodynamics, the theory of the motion of electromagnetic bodies. For sure, any kind of physical measurement—hypothetical or real—will likely depend on light in some way. After all, we need to *see* to make our observations. But if this relationship is to represent something deeply fundamental about the nature of matter (and its mass), then we must be able to make it more generally applicable. This means separating it from the motions of bodies that are electrically charged or magnetized and from situations involving the absorption or emission of electromagnetic radiation.

Either we find a way to get rid of c entirely from the equation or we find another interpretation for c that has nothing to do with the speed of light.

This project was begun in 1909, by Gilbert Lewis and Richard Tolman, but was arguably completed only in 1972, by Basil Landau and Sam Sampanthar, mathematicians from Salford University in England.⁸ In their formulation, a quantity equivalent to c appears as a constant of integration, representing an absolute upper limit on the speed that any object can possess.

What this suggests is that nothing—but nothing—in the universe can travel faster than this limiting speed. And for reasons that remain essentially mysterious, light (and, indeed, all particles thought to have zero rest mass) travels at this ultimate speed. We don't need to eliminate c from the equation connecting mass and energy; we just reinterpret it as a universal limit.

Of course, electromagnetic radiation might have zero rest mass, but it can never be brought to rest. This implies that radiation has an inertial mass associated with its energy, and this strengthened Einstein's belief that radiation possesses particle-like properties.

.....

NOTES

1. Albert Einstein, 'Zur Elektrodynamik bewegter Körper', *Annalen der Physik*, 17 (1905), 891–921. English translation quoted in John Stachel (ed.), *Einstein's Miraculous Year: Five Papers That Changed the Face of Physics*, Princeton University Press, Princeton, NJ, 2005, p. 124.
2. Einstein, 'Zur Elektrodynamik bewegter Körper'. English translation quoted in Stachel (ed.), *Einstein's Miraculous Year*, p. 150.
3. Abraham Pais, *Subtle is the Lord: The Science and the Life of Albert Einstein*. Oxford University Press, Oxford, 1982, p. 147.
4. Albert Einstein, 'Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?' *Annalen der Physik*, 18 (1905), 639–41. English translation quoted in Stachel (ed.), *Einstein's Miraculous Year*, p. 164. The italics are mine.
5. In quantum chromodynamics, as much as 95% of the mass of protons and neutrons arises from the *energy* of the interactions between their constituent quarks, and the masses of quarks and electrons are, in turn, derived from their interactions with the Higgs field. See, for example, Jim Baggott, *Mass: The Quest to Understand Matter from Greek Atoms to Quantum Fields*, Oxford University Press, Oxford, 2017.
6. Einstein, 'Ist die Trägheit eines Körpers'. English translation quoted in Stachel (ed.), *Einstein's Miraculous Year*, p. 164.
7. Albert Einstein, letter to Conrad Habicht, autumn 1905 (undated), quoted in Pais, *Subtle is the Lord*, pp. 148–9.
8. See, for example, Max Jammer, *Concepts of Mass in Contemporary Physics and Philosophy*, Princeton University Press, Princeton, NJ, 2000, p. 46.

.....

FURTHER READING

- Crease, Robert P., *A Brief Guide to the Great Equations: The Hunt for Cosmic Beauty in Numbers*, Robinson, London, 2009, Chapter 7.
- Cushing, James T., *Philosophical Concepts in Physics: The Historical Relation between Philosophy and Scientific Theories*, Cambridge University Press, Cambridge, UK, 1998, pp. 242–6.
- Einstein, Albert, *Relativity: The Special and the General Theory*, 100th anniversary edition, Princeton University Press, Princeton, NJ, 2015.

- Galison, Peter, 'The Sextant Equation: $E = mc^2$ ', in Graham Farmelo (ed.), *It Must Be Beautiful: Great Equations of Modern Science*, Granta, London, 2002, pp. 68–86.
- Kennedy, Robert E., *A Student's Guide to Einstein's Major Papers*, Oxford University Press, Oxford, 2012, Chapter 4.
- Pais, Abraham, *Subtle is the Lord: The Science and the Life of Albert Einstein*. Oxford University Press, Oxford, 1982, Chapters 6–7.
- Robinson, Andrew (ed.), *Einstein: A Hundred Years of Relativity*, revised and updated edition, Princeton University Press, Princeton, NJ, 2015, Chapter 3.
- Stachel, John (ed.), *Einstein's Miraculous Year: Five Papers that Changed the Face of Physics*, Princeton University Press, Princeton, NJ, 2005, pp. 99–164.

Bohr's Derivation of the Rydberg Formula

Quantum Numbers and Quantum Jumps

The notion that energy might be quantized was forged in the porcelain furnaces used to study black body radiation. As we saw in Chapter 1, the result of Planck's act of desperation was interpreted by Einstein to suggest that electromagnetic radiation could itself be considered to consist of discrete light-quanta, of energy $E = h\nu$, which today we call photons. But Einstein suspected that quantum principles, and Planck's fundamental constant h , held even more significance for physics. In 1907, he wrote:¹

If the elementary structures that are to be assumed in the theory of energy exchange between radiation and matter cannot be perceived in terms of the current molecular-kinetic theory, are we then not obliged also to modify the theory for the other periodically oscillating structures considered in the molecular theory of heat? In my opinion the answer is not in doubt. If Planck's radiation theory goes to the root of the matter, then contradictions between the current molecular-kinetic theory and experience must be expected in other areas of the theory of heat as well.

Planck had arrived at his conclusion by studying the distribution of energy (in the form of indistinguishable energy elements, E) over the 'oscillators' of the cavity material. Einstein now suggested that similar principles should apply to the distribution of *heat energy* over the oscillating atoms in the lattice structures of crystalline solids. In other words, the heat capacities of solids might also be quantum in nature.

It was understood that crystalline solids such as diamond absorb heat which is stored internally in its lattice vibrations. In the classical theory of Pierre Dulong and Alexis Petit, the heat capacities of all such solids are the same and given approximately by $3N_A k_B$, where N_A is Avogadro's constant and k_B is Boltzmann's constant.* This is roughly true at room temperature (although diamond is a notable exception), but experimental studies at

* This can be written as $3R$, where R is the ideal gas constant.

low temperatures had shown substantial departures from this general rule and differences between different solids that couldn't be explained using classical physics.

Einstein now applied the same kind of quantum logic that Planck had used in his derivation of radiation entropy. He simplified the problem by assuming that all the lattice atoms vibrate at the same frequency (ν), and further assumed that energy is absorbed or emitted by these vibrations only in integral multiples of $h\nu$. The end result is that the Dulong–Petit rule is modified by a function that depends on temperature, such that the heat capacity declines to zero at the absolute zero of temperature. In his 1907 paper, Einstein included a diagram showing general agreement between his theory and the experimental heat capacity of diamond as a function of temperature which, as Pais remarks, 'represents the first published graph in the history of the quantum theory of the solid state'.²

Einstein's theory would eventually be replaced by more sophisticated treatments of the lattice vibrations, but he had more than made his point. Momentum was building, and in 1910 the eminent German chemist Walther Nernst paid Einstein a visit (by this time Einstein was back in Zurich, but was still a relative unknown). Nernst's interest encouraged greater respect for Einstein and his work, with one Zurich colleague repeating the general opinion: 'This Einstein must be a clever fellow, if the great Nernst comes so far from Berlin to Zurich to talk to him.'³ Nernst was instrumental in drawing attention to Einstein's quantum approach and from early 1911 a growing number of scientists began to cite Einstein's papers and embrace quantum ideas.

The Structure of the Atom

In the meantime, atoms had evolved from hypothetical entities, dismissed by physicists such as Mach as the result of wildly speculative metaphysics, into the objects of detailed laboratory study. The discovery of the negatively charged electron ('cathode rays') by English physicist Joseph John Thomson in 1897 implied that atoms, indivisible for more than 2000 years, now had to be recognized as having some kind of internal structure.

Thomson himself developed a simplified model consisting of point-like, negatively charged electrons moving in rings around the centre of a sphere of uniformly positively charged fluid. He initially suggested that even the simplest atoms would contain a very large number of electrons (for example, a hydrogen atom would hold about 1000 electrons), but as evidence showed that the numbers involved were much smaller, he adapted his model accordingly. This became popularly (and unfairly) known as Thomson's 'plum pudding' model of the atom.⁴

Alternatives included variations on a 'planetary' model, following a long tradition in science of applying ideas from macroscopic physics to describe the microscopic world. Further clues were revealed in 1909 when, under the direction of Ernest Rutherford in Manchester, Hans Geiger and Ernest Marsden reported the results of experiments in which energetic alpha-particles were fired through samples of thin gold foil. To their astonishment, they found that about one in 8000 alpha-particles was deflected by the foil, sometimes by more than 90°. This is no less startling than shooting high-velocity

machine-gun bullets through tissue paper, and being forced to duck when one of them bounces back at you.

In 1911 Rutherford interpreted these results to mean that most of the atom's mass is concentrated in a small central nucleus, surrounded by electrons which account for much of the volume of the atom. This was a 'nuclear atom' rather than a planetary atom, as Rutherford was ambiguous about how the electrons might be distributed: 'The electrons may be supposed to be distributed throughout a spherical volume or in concentric rings in one plane.'⁵

Rutherford's hesitancy is not difficult to understand. Unlike the Sun and planets of the Solar System, electrons and atomic nuclei carry electrical charge. It was known from Maxwell's theory that electrical charges moving in an electromagnetic field radiate energy in the form of electromagnetic waves. This radiation carries energy away from the orbiting electrons, presumably causing them to slow down and leaving them exposed to the irresistible pull of the positively charged nucleus. The electrons in a planetary model would spiral down towards the nucleus as they lost their energy and the atoms would collapse in on themselves within about one hundred-millionth of a second.

Though they might be compelling, such atomic models are also disastrous.

A Little Bit of Reality

On completing his PhD on the electron theory of metals, in September 1911 the young Danish physicist Niels Bohr journeyed to Cambridge in England, supported by a stipend from the Carlsberg Foundation.

Bohr joined J. J. Thomson's research team, but unfortunately their relationship got off to a poor start from which it seems never to have recovered. As a young postdoctoral student, Bohr's grasp of the English language was poor. Though always polite and courteous, Bohr's manner could sometimes appear brusque and was open to misinterpretation. His first meeting with Thomson was not auspicious. He entered Thomson's office with a copy of one of Thomson's books on atomic structure, pointed to a particular section and declared: 'This is wrong.'⁶ It is perhaps hardly surprising that Thomson did not immediately warm to him.

Bohr struggled on, growing increasingly frustrated. He met Rutherford for the first time in early November 1911, and transferred to the Manchester laboratory in December 1911. He started work there the following March, and learnt what he needed to know about radioactivity and Rutherford's nuclear atom from George von Hevesy and Charles Darwin (who Bohr would always introduce to others as the grandson of the 'real' Darwin).

It was whilst working on some problems raised by Darwin that he became absorbed by the challenges posed by the structure of the atom. The classical picture said that atoms should not exist. Yet atoms clearly do exist and so it was likely to be impossible to deduce a theoretical description based only on classical mechanics and electromagnetism. Perhaps, he reasoned, some progress could be made by employing quantum ideas. He had become convinced that the inner electronic structure of the Rutherford model was governed in

some way by Planck's quantum of action. 'This seems to be nothing else than what was to be expected as it seems rigorously proved that the [classical] mechanics cannot explain the facts in problems dealing with single atoms.'⁷ To attempt a classical description seemed hopeless.

In June 1912 he wrote a letter to his brother Harald:

Perhaps I have found out a little about the structure of atoms. Don't talk about it to anybody, for otherwise I couldn't write to you about it so soon. If I should be right it wouldn't be a suggestion of the nature of a possibility (i.e., an impossibility, as J. J. Thomson's theory) but perhaps a little bit of reality . . . You understand that I may yet be wrong; for it hasn't been worked out fully yet (but I don't think so); also, I do not believe that Rutherford thinks that it is completely wild . . . Believe me, I am eager to finish it in a hurry, and to do that I have taken off a couple of days from the laboratory (this is also a secret).

Bohr's model was still nevertheless rife with contradictions. He summarized his work in a manuscript (now referred to variously as the Rutherford Memorandum or the Manchester Memorandum) which he submitted to Rutherford on 6 July, but which was never published. He left Manchester and returned to Copenhagen a couple of weeks later, carrying the problems back with him in his briefcase.

Bohr continued to work on the problem of atomic structure through the rest of 1912 and into early 1913, when he was given a clue that would unlock the entire mystery. During a conversation with Hans Marius Hansen, a young professor of physics who had performed some experiments on atomic spectroscopy at the University of Göttingen in Germany, Bohr's attention was drawn to something called the *Balmer formula*.

The Balmer and Rydberg Formulas

Whilst it might have been anticipated from classical physics that atoms such as hydrogen should absorb and emit energy continuously, favouring no particular radiation frequencies, physicists in the nineteenth century had found the opposite to be true. The radiation emitted by excited hydrogen atoms, to take the simplest example, forms a 'line' spectrum, consisting of a series of discrete, narrowly defined frequencies (see Fig. 3.1). In 1885, the Swiss mathematician Johann Jakob Balmer found that the wavelengths of one series of hydrogen emission lines that span the visible region conform to a relatively simple pattern, described by

$$\lambda = A \left(\frac{n^2}{n^2 - 2^2} \right), \quad (3.1)$$

where λ is the wavelength of the emission line, A is a constant (with a value of about 364.7 nm), and n is an integer greater than 2: $n = 3, 4, 5$, etc. This is the Balmer formula.

In 1888 Swedish physicist Johannes Rydberg generalized Balmer's formula to include further series of hydrogen emission lines. Today, we write this as

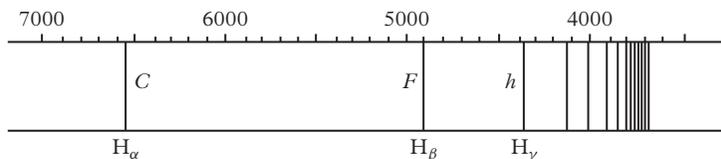


Figure 3.1 This picture of the Balmer series appears in a 1910 textbook *Lærebog i Physik*, by Christian Christiansen, who taught Bohr at the University of Copenhagen. Wavelengths are recorded in ångstroms (tenths of a nanometre) along the top, with the H_α (656.3 nm), H_β (486.1 nm), and H_γ (434.0 nm) lines marked.

Source: Reproduced in Helge Kragh, *Niels Bohr and the Quantum Atom*, Oxford University Press, 2012 (Figure 2.2, p. 57).

$$\frac{1}{\lambda} = R_H \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right), \quad (3.2)$$

in which R_H is now the Rydberg constant for the hydrogen atom (with a measured value of $109,675 \text{ cm}^{-1}$), n_1 and n_2 are integers, and $n_2 > n_1$.

The Balmer series is then a particular instance of the Rydberg formula with $n_1 = 2$. In 1908, Swiss theorist Walther Ritz further generalized this as a *combination principle* for all atoms: the reciprocal wavelength (called the wavenumber) of any spectral line can be calculated as the difference between two terms: $1/\lambda = T_1 - T_2$. For hydrogen we identify the term T_m as R_H/n_m^2 .

‘As soon as I saw Balmer’s formula, the whole thing was immediately clear to me,’ Bohr later claimed.⁸ It would seem most unlikely that Bohr was completely unaware of this formula before February 1913, but after wrestling with the problems of atomic structure for many months its relevance was now suddenly clear to him. Bohr could immediately see where the integer numbers were coming from. They are *quantum numbers*.

We can now get a sense of what’s involved in Bohr’s atomic model by following the logic of his derivation of the Rydberg formula. This is what Bohr did.

The Ingredients

1. The expression for the classical centripetal force associated with uniform circular motion.
2. Coulomb’s law for the force between electrically charged particles.
3. The First Assumption: the quantized electron orbital energy is given as an integral multiple of Planck’s constant times half the mechanical frequency of the orbit (with which we will associate the symbol ω to avoid confusion with radiation frequency, ν).
4. The Second Assumption: radiation energy is emitted or absorbed by the electron in a hydrogen atom only in discontinuous transitions or ‘jumps’ between orbits,

with the frequency of radiation governed by the difference in the energies of the orbits according to the Planck–Einstein relation: $\Delta E = h\nu$.

5. The realization that the First Assumption is equivalent to assuming that the electron *orbital angular momentum* is constrained to integral multiples of \hbar , the ‘reduced’ Planck constant, $\hbar = h/2\pi$.

The Recipe

Perhaps the first thing to note about the list of ingredients for this recipe is the prevalence of a couple of key assumptions. The First Assumption (ingredient 3), in particular, is somewhat curious. It was only after presenting his derivation that Bohr commented (in the same paper) that this is equivalent to assuming that the *angular momentum* of the electron orbit is quantized. Most textbook presentations of Bohr’s derivation therefore make use of this variation of the First Assumption, which with hindsight is arguably much more reasonable (although, as we will see in subsequent chapters, it was later shown to be quite incorrect).

We begin in Step (1) by adopting a simple model involving an electron moving with uniform circular motion around a central nucleus.* This is very much a classical, ‘planetary’ model, and initially Bohr even assumed the orbits to be elliptical, although the results are not affected by assuming circular orbits. We identify the classical centripetal force associated with a stable orbit with the Coulomb force of electrostatic attraction between the electron and the nucleus. This allows us to deduce a relationship between the mechanical frequency (ω) and the *classical* energy of the orbit.

In Step (2) we impose a quantum condition on the energies of the stable orbits, assuming that they must be represented by integral multiples of Planck’s constant times an ‘average’ mechanical frequency, taken to be $\frac{1}{2}\omega$. We rearrange this expression and equate it to the classical result we got for ω in Step (1). Further manipulations allow us to derive an expression for the *quantum* energies of the stable orbits. Assuming that radiation is emitted or absorbed by the electron only when moving between these stable orbits allows us to deduce the Rydberg formula and an expression for the Rydberg constant in Step (3).

In Step (4) we examine an alternative to the quantum condition imposed through the First Assumption, in which the electron orbital angular momentum is constrained to integral multiples of \hbar . Whilst this gives exactly the same results, it implies a potentially much more powerful conclusion: in the hydrogen atom both electron orbital energy *and* angular momentum are quantized.

* I won’t call this a proton just yet, as the proton wasn’t discovered until 1917.

Step (1): Derive the Classical Frequencies of the 'Planetary' Electron Orbits

Bohr published his derivation in the opening sections of the first paper in a series of three entitled 'On the Constitution of Atoms and Molecules', known as the 'trilogy'. Here Bohr essentially asks us to set aside our preconceptions and prejudices and examine the interior structure of the hydrogen atom *as though* it behaves entirely classically. The atom is then conceived much like a planetary system, as an electron (with mass m_e) orbiting a central nucleus. Strictly speaking, we should use the reduced mass $\mu = m_e m_H / (m_e + m_H)$, where m_H is the mass of the hydrogen atom nucleus. But if we assume that the electron is much lighter than the nucleus, $m_e \ll m_H$ then we can assume that $\mu \cong m_e$.⁹ If we further stretch the analogy we would assume an elliptical orbit, as Bohr does initially, but we will see that it makes no difference to our conclusions if we assume the orbit is circular.

The orbit then has a fixed radius r and an orbital or angular velocity v (see Fig. 3.2), measured—for example—in radians per second. We know from the application of Newton's second law to such a system of uniform circular motion that from the perspective of a stationary observer, there will be a centripetal force and an acceleration directed inwards towards the centre, given by $m_e v^2 / r$. From the perspective of the frame of reference of the orbiting electron (in which the electron is stationary), there is no net force—the centripetal force is balanced by an opposing centrifugal force which is directed outwards:

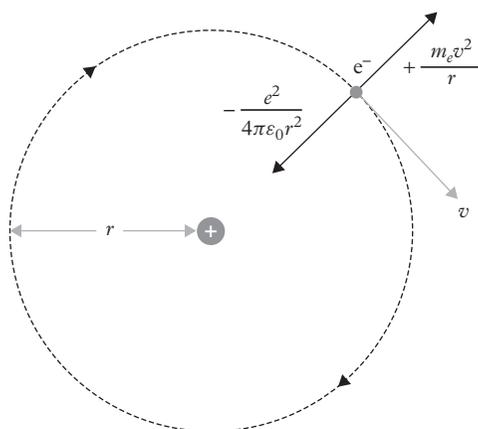


Figure 3.2 The 'planetary' orbit of the electron is assumed to be circular, with orbital or angular velocity v and radius r .

$$F_{out} = +\frac{m_e v^2}{r}. \quad (3.3)$$

In the case of an object tethered to, and rotating around, a central point, we know that the inward centripetal force is provided by the tension in the tether. In the case of a planetary hydrogen atom, the centripetal force is provided by the electrostatic attraction between the electron and the nucleus according to Coulomb's law.

Generally, for a system with electrical charges q_1 and q_2 separated by a distance r , Coulomb's law is $F = q_1 q_2 / 4\pi \epsilon_0 r^2$, where ϵ_0 is the permittivity of free space. If the charges are both positive or both negative, then the Coulomb force is positive and the charges repel. If the charges are of opposite sign then the force is negative and attractive. For the hydrogen atom, q_1 is the charge on the electron, $-e$, and q_2 is the charge on the nucleus, $+e$. The centripetal force acting on the electron is then

$$F_{in} = -\frac{e^2}{4\pi \epsilon_0 r^2}. \quad (3.4)$$

The net force on the electron in a stable orbit then given by

$$F_{net} = F_{out} + F_{in} = \frac{m_e v^2}{r} - \frac{e^2}{4\pi \epsilon_0 r^2} = 0. \quad (3.5)$$

We can rearrange Eq. (3.5) to give an expression for v^2 as follows:

$$v^2 = \frac{e^2}{4\pi \epsilon_0 m_e r}. \quad (3.6)$$

We now turn our attention to the total energy of the electron orbit. We know that this total energy is $E = T + V$, where T is the kinetic energy of the electron in the orbit and V is the potential energy associated with the Coulomb force. Evaluating the kinetic energy is straightforward:

$$T = \frac{1}{2} m_e v^2 = \frac{e^2}{8\pi \epsilon_0 r}. \quad (3.7)$$

Recall from Eq. (P.7) that the potential energy is defined as minus the integral of the force with respect to distance (in this case r), or $V = -\int F dr$. Applying this to the expression for the Coulomb force acting on the electron in a hydrogen atom, Eq. (3.4), we get

$$V = \frac{e^2}{4\pi \epsilon_0} \int \frac{1}{r^2} dr = -\frac{e^2}{4\pi \epsilon_0 r}. \quad (3.8)$$

We note in passing that the relationship between the kinetic and potential energies of this system conforms to the *virial theorem*. For a system with V proportional to r^n , then twice the time average of the kinetic energy $\langle T \rangle$ is equal to n times the time average of the total potential energy. In this case, $n = -1$ and $2\langle T \rangle = \langle V \rangle$.

So, the total energy of a classical stable orbit is

$$E = T + V = \frac{e^2}{8\pi\epsilon_0 r} - \frac{e^2}{4\pi\epsilon_0 r} = -\frac{e^2}{8\pi\epsilon_0 r}. \quad (3.9)$$

The energy of the orbit is negative as this represents a state of lower energy compared with the situation in which the electron is removed to infinity, taken as the (arbitrary) energy zero.

Our final task in this step is to deduce an expression for the mechanical frequency of the orbit, ω . This is simply calculated from the orbital velocity divided by the circumference, $\omega = v/2\pi r$. From (3.9) we have

$$r = \frac{e^2}{8\pi\epsilon_0 (-E)}. \quad (3.10)$$

We can substitute this expression for r in Eq. (3.6) and take the square root:

$$v = \sqrt{\frac{2(-E)}{m_e}}, \quad (3.11)$$

giving

$$\omega = \frac{v}{2\pi r} = \frac{4\sqrt{2}\epsilon_0(-E)^{3/2}}{e^2\sqrt{m_e}}. \quad (3.12)$$

This is the first result that Bohr gives in the first paper of the trilogy.

If this seems a rather clumsy or cumbersome expression, it might be helpful to think of it this way. The orbital period τ is simply the reciprocal of the angular frequency, $\tau = 1/\omega$. Thus,

$$\frac{\tau^2}{r^3} = \frac{16\pi^3\epsilon_0 m_e}{e^2}, \quad (3.13)$$

which is a constant, independent of the orbital energy E . This is, of course, Kepler's third law of planetary motion, in which we have approximated the semi-major axis of the elliptical (planetary) orbit as the radius of a circular (electron) orbit.

This is all fine as far as it goes, but of course this is an inherently unrealistic classical model, as Bohr goes on to acknowledge:¹⁰

Let us now, however, take the effect of the energy radiation into account, calculated in the ordinary way from the acceleration of the electron. In this case the electron will no longer describe stationary orbits. $[E]$ will continuously increase, and the electron will approach the nucleus describing orbits of smaller and smaller dimensions, and with greater and greater frequency; the electron on the average gaining in kinetic energy at the same time as the whole system loses energy. This process will go on until the dimensions of the orbit are the same order of magnitude as the dimensions of the electron or those of the nucleus. A simple calculation shows that the energy radiated out during the process considered will be enormously great compared with that radiated out by ordinary molecular processes.

This is clearly not how real atoms behave. Bohr understood that in the classical planetary model energy is assumed to be continuously variable, in striking contrast to the evidence from atomic spectroscopy, which shows that atoms emit or absorb energy only in discrete quantities. To recover the pattern of lines in the atomic emission and absorption spectra, it is necessary to impose a *quantum* condition.

Step (2): Derive the Frequency of Quantum Electron Orbits, and Compare

But how should this quantum condition be applied? Bohr wrote:¹¹

Now the essential point in Planck's theory of radiation is that the energy radiation from an atomic system does not take place in the continuous way assumed in the ordinary electrodynamics, but that it, on the contrary, takes place in distinctly separated emissions, the amount of energy radiated out from an atomic vibrator of frequency ν in a single emission being equal to $[n] h\nu$, where $[n]$ is an entire number, and h is a universal constant.

Bohr imagined an electron removed to an infinite distance from the nucleus. In this situation, the electron has no orbital velocity or frequency to speak of—from Eq. (3.10), if $r = \infty$, then $E_\infty = 0$, and, from (3.12), $\omega_\infty = 0$. As the electron and nucleus are brought together, the energy falls. Conservation of energy demands that the difference (ΔE) is emitted in the form of electromagnetic radiation. But the Planck–Einstein equation suggests that such radiation is constrained to integral multiples of $h\nu$, or $\Delta E = nh\nu$. This must mean that the electron orbits are likewise constrained to energies determined by n , where n can take values 1, 2, 3, . . . , etc. So, let's label the orbital energy characterized by the integer n as E_n , with orbital frequency ω_n .

In his derivation of the radiation formula (Chapter 1), Planck had assumed that the frequency of the imaginary oscillators and the frequency of the radiation absorbed or emitted by them was identical. So Bohr now equated the frequency of the radiation emitted as an electron is brought from infinity to bind with the nucleus to the *average* of the mechanical frequencies of the electron at infinity and in the final orbit, i.e. $\nu = (\omega_n + \omega_\infty) / 2 = \frac{1}{2}\omega_n$. He wrote: 'If we assume that the radiation emitted is homogeneous, the second assumption concerning the frequency of the radiation suggests itself, since the frequency of revolution of the electron at the beginning of the emission is 0.'¹²

Based on this assumption, the difference in energies at infinity and in the n th orbit is given by

$$E_\infty - E_n = \frac{1}{2}nh\omega_n \quad \text{or} \quad E_n = -\frac{1}{2}nh\omega_n \quad (3.14)$$

(remember, $E_\infty = 0$).

If you have some doubts about this First Assumption, please be reassured—it's clear that Bohr had doubts, too, which is why he chose to revisit it later in his paper. We will look at this assumption again in Step (4).

We can rearrange (3.14) as follows:

$$\omega_n = \frac{2(-E_n)}{nh}. \quad (3.15)$$

And we can now compare this result with the expression (3.12), replacing ω with ω_n and E with E_n in the latter, giving

$$\frac{2(-E_n)}{nh} = \frac{4\sqrt{2}\varepsilon_0(-E_n)^{3/2}}{e^2\sqrt{m_e}} \quad (3.16)$$

or, after some rearrangement,

$$E_n = -\frac{m_e e^4}{8\varepsilon_0^2 h^2} \cdot \frac{1}{n^2}. \quad (3.17)$$

The dependence of the orbital energy on $1/n^2$ is heartening, as this clearly takes us a step closer to the Rydberg and Balmer formulas. But before we take the next step it is worth substituting this result for E_n into the expression for the orbital radius, which we now write as r_n , in Eq. (3.10):

$$r_n = n^2 \cdot \frac{\varepsilon_0 h^2}{\pi m_e e^2}. \quad (3.18)$$

The lowest-energy orbit of the electron in a hydrogen atom corresponds to $n = 1$, so the radius of this orbit is fixed and specified by a collection of physical and geometric constants. Using the values of these constants allowed Bohr to deduce that r_1 has a value of 5.5×10^{-11} m, or 0.055 nm.¹³ This is called the *Bohr radius*.

Step (3): Impose the Planck Condition on Transitions between Orbits

The next step is relatively straightforward. Absorption of radiation promotes the electron from a lower to a higher energy orbit (let's say from an orbit characterized by n_1 to an orbit characterized by n_2 , where $n_2 > n_1$). Likewise, the emission of radiation signals the electron falling from a higher to a lower energy orbit, let's say from n_2 to n_1 . But energy can be absorbed or emitted only in integral multiples of $h\nu$. Bohr now imposed his Second Assumption—the frequency of the absorbed or emitted radiation is determined by the difference in the energies of the orbits according to

$$\Delta E = E_{n_2} - E_{n_1} = h\nu. \quad (3.19)$$

There's a little more to this condition. Transitions between the orbits would have to occur instantaneously in 'jumps', because if the electron were to move gradually from one orbit to another, it would again be expected to radiate energy continuously during the process. Transitions between inherently non-classical stable orbits must themselves involve non-classical discontinuous jumps. Bohr wrote that¹⁴

the dynamical equilibrium of the systems in the [stable orbits] can be discussed by help of the ordinary [classical] mechanics, while the passing of the systems between the different [stable orbits] cannot be treated on that basis.

The integer numbers that characterize the electron orbits would come to be known as *quantum numbers*, and the transitions between different orbits would become *quantum jumps*.

From (3.17) we can deduce that

$$E_{n_2} - E_{n_1} = \frac{m_e e^4}{8\epsilon_0^2 h^2} \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) = h\nu = \frac{hc}{\lambda}. \quad (3.20)$$

And so we get

$$\frac{1}{\lambda} = \frac{m_e e^4}{8\epsilon_0^2 h^3 c} \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right), \quad (3.21)$$

which is just the Rydberg formula, Eq. (3.2), in which the Rydberg constant R_H is now revealed to be a collection of physical constants, $R_H = m_e e^4 / 8\epsilon_0^2 h^3 c$. Using the values of the physical constants available to Bohr at the time, he got an estimate for R_H of $109,740 \text{ cm}^{-1}$. This is just fractionally higher than the experimentally observed value ($109,675 \text{ cm}^{-1}$) and equivalent within the bounds of accuracy of the physical constants.

We know that setting $n_1 = 2$ yields the Balmer series in the visible region. In his paper Bohr noted that setting $n_1 = 3$ gives another series in the infrared that had previously been identified by German physicist Friedrich Paschen in 1908. Setting $n_1 = 1$ and $n_1 = 4, 5$ predicted further emission series in the ultraviolet and infrared that had not at that time been observed 'but the existence of which may be expected'.¹⁵ These would become known as the Lyman series, discovered in the period 1906–14 (and named for Theodore Lyman), the Brackett series (1922, named for Frederick Sumner Brackett), and the Pfund series (1924, named for August Herman Pfund)—see Fig. 3.3.

A further series of emission lines named for American astronomer and physicist Edward Charles Pickering was thought by experimentalists also to belong to the hydrogen atom. However, at the time, the Pickering series was characterized by terms involving half-integer numbers which are not admissible in Bohr's theory. Instead, Bohr proposed that the formula be rewritten in terms of integer numbers, suggesting that the Pickering series belongs not to hydrogen atoms but rather to ionized helium.

If we generalize the Balmer formula for atoms with nuclear charges Z greater than 1, we find that the Rydberg constant becomes $R_Z = Z^2 R_H$. This suggests that the Rydberg constant for helium should be exactly four times that measured for the hydrogen atom. But English astronomer Alfred Fowler objected. The spectroscopic data were instead consistent with a Rydberg constant for helium that was 4.0016 times that for hydrogen. The difference, though very small, could not be attributed to experimental error.

Bohr puzzled over this for a while, before realizing that the discrepancy lay in the use of the electron mass m_e instead of the reduced mass μ . Denoting the Rydberg constant for

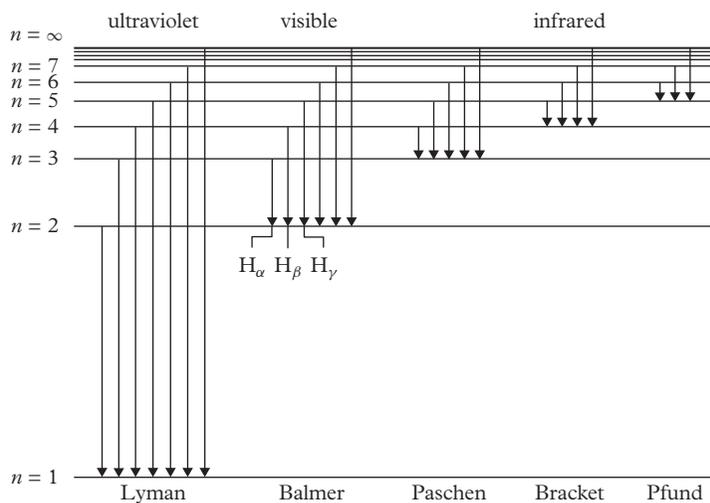


Figure 3.3 Energy levels (not shown to scale) and emission transitions in the spectrum of atomic hydrogen. The Balmer series lies in the visible and the distinctive H_α (656.3 nm) line lends a reddish hue to many true-colour photographs of astronomical objects such as the Orion nebula.

helium as R_{He} , and replacing the electron mass in both R_{He} and R_H with the expression for the reduced mass, we get

$$\frac{R_{He}}{R_H} = 4 \left(\frac{m_e m_{He}}{m_e + m_{He}} \right) \left(\frac{m_e + m_H}{m_e m_H} \right) = 4 \frac{m_{He}}{m_H} \left(\frac{m_e + m_H}{m_e + m_{He}} \right). \quad (3.22)$$

Using the available data on the nuclear masses allowed Bohr to deduce that the ratio R_{He}/R_H should be 4.00163, compared to 4.0016 from observation. Bohr's biographer Abraham Pais wrote: 'Up to that time no one had ever produced anything like [this precision] in the realm of spectroscopy, agreement between theory and experiment to five significant figures.'¹⁶ This kind of agreement was simply unprecedented.

Step (4): Recast the First Assumption as a Quantization of Orbital Angular Momentum

Bohr returned to his rather *ad hoc* First Assumption later in his first paper. He pondered some more on the relationship between the radiation frequency and the mechanical frequency of the orbit, concluding that¹⁷

We are thus led to assume that the interpretation of [Eq. (3.14)] is not that the different [stable orbits] correspond to an emission of different numbers of energy-quanta, but that the frequency of the energy emitted during the passing of the system from a state in which no energy is yet radiated out to one of the different [stable orbits], is equal to different multiples of $\frac{1}{2}\omega$, where ω is the frequency of revolution of the electron in the state considered.

In other words, the integer numbers n do not originate from the quantization of the radiation (as $nh\nu$), but rather from the quantization of the orbits themselves ($\frac{1}{2}nh\omega_n$). Later on in the paper Bohr traces this quantization to the orbital angular momentum:¹⁸

In any molecular system consisting of positive nuclei and electrons in which the nuclei are at rest relative to each other and the electrons move in circular orbits, the angular momentum of every electron round the centre of its orbit will in the permanent state of the system be equal to $h/2\pi$, where h is Planck's constant.

In a system involving circular motion, the orbital angular momentum is the vector or cross product of the position and momentum vectors, $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, where \mathbf{L} , \mathbf{r} , and \mathbf{p} are the angular momentum, radial, and linear momentum vectors, respectively.* The scalar magnitude of the angular momentum is then given by $L = |\mathbf{r}||\mathbf{p}|\sin\theta$, where θ is the angle between the radial and linear momentum vectors. But, of course, the linear momentum vector forms a tangent with the orbit and so $\theta = 90^\circ$, $\sin\theta = 1$, and $L = rp$. If we

* There's a little more background on this available in Appendix 6.

assume the classical expression for linear momentum, then for an electron p is given by $m_e v$, where v is again the orbital or angular velocity. In other words, $L = m_e v r$. The quantum condition is now that the angular momentum in the n th orbit, L_n , must be an integral multiple of the fundamental quantum of angular momentum, $h/2\pi$:

$$L_n = m_e v_n r_n = n \frac{h}{2\pi} = n\hbar, \quad (3.23)$$

where, as before, r_n is the radius of the n th orbit and v_n the orbital or angular velocity. If we make use of the result for v_n derived from Eq. (3.6) we can rearrange (3.23) to give an expression for r_n . Not surprisingly, this is the same as Eq. (3.18). Everything else follows.

There was a precedent. A year earlier in 1912, John Nicholson, an English mathematician at the Cavendish Laboratory in Cambridge, had developed a planetary atomic model and had sought to apply quantum principles to it. He had already concluded that the existence of spectral line series implied that the orbital angular momentum of an electron in an atom might be constrained to integral multiples of $h/2\pi$. He went on to suggest that the series 'may not emanate from the same atom, but from atoms whose internal angular momentum have, by radiation or otherwise, run down by various discrete amounts from some standard value'.¹⁹

By the second paper in Bohr's trilogy, the quantization of orbital angular momentum had evolved from almost an afterthought to a universal condition although, as we will see in Chapter 5, Eq. (3.23) is quite wrong.

Alarm Bells

Bohr wrote to Rutherford on 6 March 1913, enclosing with his letter a manuscript copy of the first paper. In his reply Rutherford reacted favourably but raised some difficult questions. He was particularly puzzled by the fact that, in Bohr's model, an electron in a high-energy orbit would somehow need to 'know' beforehand the energy of the final, destination, orbit in order to emit radiation of just the right frequency.²⁰

how does an electron decide what frequency it is going to vibrate at when it passes from one [stable orbit] to the other? It seems to me that you have to assume that the electron knows beforehand where it is going to stop.

Rutherford was already ringing alarm bells about the implications of the new quantum theory for our understanding of cause and effect, bells that would continue to peal loudly for another century.

He also warned that Bohr's manuscript was rather long: 'I do not know if you appreciate the fact that long papers have a way of frightening readers,' he wrote. Bohr was nonplussed. The day before he received Rutherford's letter he had sent a revised draft of his manuscript. It was even longer.

Bohr resolved to go at once to Manchester and discuss the paper directly with Rutherford. He wrote back on 26 March and declared his intention to visit Rutherford in the first days of the following week. Rutherford was patient. After discussions through several long evenings, during which he declared he had never thought Bohr should prove to be so obstinate, he agreed to leave all the detail in the final paper and communicate it to the journal *Philosophical Magazine* on Bohr's behalf. It appeared in July 1913. The second and third parts of the trilogy appeared in the same journal in September and November that year.

Bohr's model was a triumph. But, much like Planck's achievement in 1900, it was also rather mysterious, the result of shoehorning quantum principles into an otherwise disastrous classical description. There were many unanswered questions. The most pressing of these concerned the quantum numbers. What did they mean? From where, exactly, had they come?

.....

NOTES

1. A. Einstein, 'Die Plancksche Theorie der Strahlung und die Theorie der spezifischen Wärme', *Annalen der Physik*, 22 (1907), 180. An English translation is available online from *The Collected Papers of Albert Einstein*, published by Princeton University Press: <http://einsteinpapers.press.princeton.edu/vol2-trans/228>. This quote appears on pp. 218–19.
2. Abraham Pais, *Subtle is the Lord: The Science and the Life of Albert Einstein*. Oxford University Press, Oxford, 1982, p. 389.
3. George von Hevesy, interview with Thomas Kuhn and Emilio Segre, 25 May 1962, quoted in Thomas S. Kuhn, *Black-Body Theory and the Quantum Discontinuity 1894–1912*. University of Chicago Press, Chicago, 1978, p. 215.
4. See Giora Hon and Bernard R. Goldstein, 'J. J. Thompson's Plum-Pudding Atomic Model: The Making of a Scientific Myth', *Annalen der Physik*, 525 (2013), A129–33. A traditional plum pudding actually consists of a uniform distribution of raisins baked in a suet pudding. As the raisins are stationary, they don't really serve as a helpful metaphor for the electrons in Thomson's model, which are supposed to be continually moving.
5. E. Rutherford, *Radioactive Substances and Their Radiations*, Cambridge University Press, Cambridge, UK, 1913, p. 620.
6. Niels Bohr, quoted in Abraham Pais, *Niels Bohr's Times, in Physics, Philosophy and Polity*, Oxford University Press, Oxford, 1991, p. 120.
7. Niels Bohr, *Collected Works*, Vol. 2: *Work on Atomic Physics (1912–1917)*, edited by Leon Rosenfeld and Ulrich Hoyer, Elsevier, Amsterdam, 1981, p. 136, quoted in Pais, *Subtle is the Lord*, p. 137.
8. Niels Bohr, quoted in Leon Rosenfeld, *Niels Bohr: On the Constitution of Atoms and Molecules*, Munksgaard, Copenhagen, 1963, p. xxxiv, quoted in Helge Kragh, *Niels Bohr and the Quantum Atom: The Bohr Model of Atomic Structure 1913–1925*, Oxford University Press, Oxford, 2012, p. 56.
9. The mass of the electron is 9.109×10^{-31} kg, and the mass of the proton is 1.673×10^{-27} kg. So the reduced mass μ is 9.104×10^{-31} kg, about 0.05% smaller than the electron mass.

10. N. Bohr, 'On the Constitution of Atoms and Molecules', *Philosophical Magazine*, 26 (1913), 1. This quote appears on pp. 3–4.
11. Bohr, 'On the Constitution of Atoms and Molecules', 4.
12. *Ibid.*
13. Using currently accepted values of the physical constants m_e (9.109×10^{-31} kg), e (1.602×10^{-19} C), h (6.626×10^{-34} Js), and ϵ_0 (8.854×10^{-12} Fm⁻¹) suggests that r_1 has a value of 5.292×10^{-11} m, or 0.05292 nm. See Peter J. Mohr, David B. Newell, and Barry N. Taylor, 'CODATA: Recommended values of the fundamental physical constants 2014', *Reviews of Modern Physics*, 88 (2016) 035009, 56–7.
14. Bohr, 'On the Constitution of Atoms and Molecules', 7.
15. *Ibid.* 8.
16. Pais, *Subtle is the Lord*, p. 149.
17. Bohr, 'On the Constitution of Atoms and Molecules', 13.
18. *Ibid.* 23. Italics in the original.
19. J. W. Nicholson, *Monthly Notices of the Royal Astronomical Society*, 72 (1912), 730, quoted in Kragh, *Niels Bohr and the Quantum Atom*, pp. 26–7.
20. Ernest Rutherford, letter to Niels Bohr, 20 March 1913, in Bohr, *Collected Works*, Vol. 2, p. 583, quoted in Pais, *Subtle is the Lord*, pp. 152–3.

.....

FURTHER READING

- Aaserud, Finn, and Heilbron, John L., *Love, Literature, and the Quantum Atom: Niels Bohr's 1913 Trilogy Revisited*, Oxford University Press, Oxford, 2013.
- Cushing, James T., *Philosophical Concepts in Physics: The Historical Relation between Philosophy and Scientific Theories*, Cambridge University Press, Cambridge, UK, 1998, pp. 278–81.
- Darrigol, Oliver, Duplantier, Bertrand, Raimond, Jean-Michel, and Rivasseau, Vincent (eds), *Niels Bohr, 1913–2013: Poincaré Seminar 2013, Progress in Mathematical Physics*, 68 (2016), Springer International Publishing, Switzerland, Chapters 1 and 2.
- Gamow, George, *Thirty Years That Shook Physics*, Dover, New York, 1966, Chapter II.
- Heilbron, J. L., and Kuhn, T. S., 'The Genesis of the Bohr Atom', *Historical Studies in the Physical Sciences*, 1 (1969), University of Pennsylvania Press, Philadelphia.
- Heilbron, John L., 'Bohr's First Theories of the Atom', in A. P. French and P. J. Kennedy, *Niels Bohr: A Centenary Volume*, Harvard University Press, 1985.
- Kragh, Helge, *Niels Bohr and the Quantum Atom: The Bohr Model of Atomic Structure 1913–1925*, Oxford University Press, Oxford, 2012, Chapters 1 and 2.
- Kuhn, Thomas S., *Black-Body Theory and the Quantum Discontinuity 1894–1912*. University of Chicago Press, Chicago, 1978, Chapter 9.
- Pais, Abraham, *Niels Bohr's Times, in Physics, Philosophy and Polity*, Oxford University Press, Oxford, 1991, Chapter 8.

4

De Broglie's Derivation of $\lambda = h/p$

Wave–Particle Duality

In the years following Einstein's outrageous suggestion that light might actually consist of quanta, the supporting evidence was gradually accumulated. As we saw in Chapter 1, Einstein had made some predictions concerning the photoelectric effect, and these were borne out in experiments by American physicist Robert Millikan ten years later.

For many physicists, the status of light-quanta was put virtually beyond question in 1923, when Arthur Compton and Pieter Debye showed that these could be bounced off electrons (the technical term is 'scattered'), with a consequent (and entirely predictable) change in their frequencies. This phenomenon is now known as *Compton scattering*. These experiments strongly suggest that light does indeed consist of particles with directed momenta, behaving like small projectiles. Despite this some physicists, including Planck and Bohr, remained rather stubborn. They preferred to think of quantization as having its origin in atomic structure, retaining Maxwell's classical description based on continuous electromagnetic waves.

Whatever was going to replace classical physics in the description of radiation and atomic phenomena had to confront the difficult task of somehow reconciling the wave-like and particle-like aspects of light in a single theoretical structure.

An important clue would come from Einstein's special theory of relativity. As we saw in Chapter 2, special relativity suggests that energy can be considered equivalent to mass, and that all mass represents energy. We tend to want to associate mass (and the related property of linear momentum) with material particles. But the Planck–Einstein relation connects energy with frequency. So, here are two very simple yet fundamental equations connecting energy to mass and energy to frequency. This suggests an obvious question: can they be combined?

French physicist Prince Louis de Broglie* certainly thought so. Thirty-one-year-old de Broglie was the younger son of Victor, fifth duc de Broglie. He had originally intended to pursue a career in the humanities, and had studied medieval history and law at the Sorbonne, receiving a degree in 1910. But he gained a passion for physics through the

* Pronounced 'de Broy'.

influence of his older brother, Maurice, and his experiences during World War I serving in the French Army, in field radio communications, stationed at the Eiffel Tower.

After the war, de Broglie joined a private physics laboratory headed by his brother which specialized in the study of X-rays. It was whilst he was working at this laboratory in 1923 that he thought to combine the two most iconic equations of special relativity and quantum theory. The result would be another very simple yet iconic equation, $\lambda = h/p$, where the wavelength λ is an unambiguously wave-like property and linear momentum p is equally unambiguously particle-like.

If this were true only for photons, then we likely wouldn't be getting overly excited about it. For photons, the de Broglie relation doesn't really take us much further forward than $E = hv$. What makes the relation much more interesting is de Broglie's further insight:¹

After long reflection in solitude and meditation, I suddenly had the idea, during the year 1923, that the discovery made by Einstein in 1905 should be generalized by extending it to all material particles and notably to electrons.

De Broglie's relation summarizes a profound *wave-particle duality* in nature.

Many modern textbooks either give the de Broglie relation without a derivation or present it as an axiom or postulate of quantum theory. To a certain extent this is understandable. As we will see in later chapters, the classical concept of momentum undergoes a rather radical revision in quantum mechanics, and the meaning of p is much less straightforward than the de Broglie relationship would seem to imply.

Nevertheless, I believe it is still very instructive to appreciate where this relation came from and how it was first derived. But before we can continue, it is first necessary to become familiar with a few more of the consequences of special relativity.

Energy and Linear Momentum in Special Relativity

Einstein's equation $E = mc^2$ represents a starting point for a much broader exploration of energy and momentum in special relativity. If we adopt the classical expression for the momentum, $p = mv$, where v is the velocity, then the mass m is given by $m = p/v$ and

$$E = \frac{pc^2}{v} \quad \text{or} \quad v = \frac{pc^2}{E}. \quad (4.1)$$

Now, if a small incremental change in the energy of a physical system dE is derived exclusively from a change in its kinetic energy T through the application of a force F , then $dE = dT$ and from Eqs. (P.5) and (P.1) we have

$$dE = Fdx \quad \text{or} \quad dE = \frac{dp}{dt} dx = \frac{dx}{dt} dp = vdp. \quad (4.2)$$

If we substitute for v from (4.1) we get

$$dE = \frac{pc^2}{E} dp \quad \text{or} \quad EdE = c^2 p dp. \quad (4.3)$$

On integrating, we get

$$\int EdE = c^2 \int p dp \quad \text{or} \quad E^2 = p^2 c^2 + E_0^2. \quad (4.4)$$

In Eq. (4.4), E_0^2 is a constant of integration, and the reason for writing this explicitly as the square of an energy quantity will become obvious very shortly.

From Eq. (4.1) we have $pc = Ev/c$, which we can insert directly into (4.4):

$$E^2 = \frac{E^2 v^2}{c^2} + E_0^2 \quad \text{or, on rearranging} \quad E^2 = \frac{E_0^2}{\left(1 - \frac{v^2}{c^2}\right)}. \quad (4.5)$$

You can probably now sense where this is heading. Taking the positive root of both sides of (4.5) gives

$$E = \gamma E_0, \quad \text{where} \quad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (4.6)$$

The constant of integration is thus revealed to be the square of the energy of the system in the rest frame (hence, the subscript zero). The energy E is then the relativistic energy of the system in the frame S , moving with velocity v , and γ is the Lorentz factor.

Of course, the equivalence of mass and energy applies equally to the rest frame, too, such that $E_0 = m_0 c^2$, where m_0 is the mass of the system at rest, called the 'rest mass'. We can now use this in (4.4) to obtain a general expression for the relativistic energy:

$$E^2 = p^2 c^2 + m_0^2 c^4. \quad (4.7)$$

If it helps, we can represent this in terms of a right-angled triangle with sides pc and $m_0 c^2$. The hypotenuse is then the total energy E (see Fig. 4.1).

Equation (4.7) is entirely equivalent to, but much more generally applicable than, $E = mc^2$. For one thing, we can apply it to particles with zero rest mass (such as photons), without involving ourselves in difficult debates about the origin of 'relativistic' mass in such particles. To see why this might be a problem we can extend (4.6) to mass and momentum as follows:

$$E = mc^2 = \gamma m_0 c^2, \quad \text{suggesting} \quad m = \gamma m_0 \quad (4.8)$$

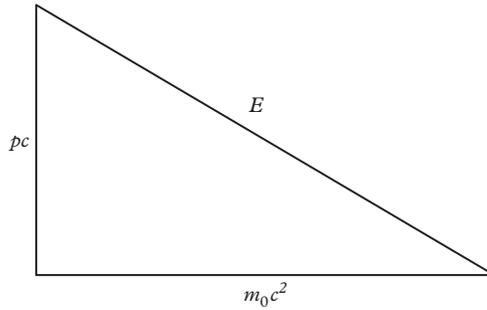


Figure 4.1 Equation (4.7) can be pictured as a right-angled triangle with sides pc and $m_0 c^2$ and hypotenuse E .

and

$$p = mv = \gamma m_0 v. \quad (4.9)$$

The problem with (4.8) and (4.9) is that they can't be applied to photons, because photons have zero rest mass and move at the speed of light $v = c$, giving $\gamma m_0 = 0/0$, which is generally said to be 'indeterminate'. But we know that photons not only possess energy, they also have momentum, which we experience directly through phenomena such as radiation pressure and Compton scattering. This is where Eq. (4.7) comes to the rescue, since for photons with $m_0 = 0$, we have (again taking the positive root)

$$E = pc. \quad (4.10)$$

This might be an unfamiliar (and therefore somewhat surprising) result. If photons move with velocity c and we can imagine them to possess a relativistic mass m and momentum mc (without worrying overmuch what m might mean in this context or where it comes from), then why isn't the kinetic energy given by $\frac{1}{2}mc^2 = \frac{1}{2}pc$? Then we realize that this is muddled thinking. From (4.6) we have

$$E = \gamma E_0 = \frac{E_0}{\sqrt{1 - \frac{v^2}{c^2}}} \approx E_0 \left(1 + \frac{1}{2} \frac{v^2}{c^2} - \dots \right), \quad (4.11)$$

where, once again, we have substituted a Taylor series expansion for the Lorentz factor, γ —look back at Eq. (2.33), replacing x with $-v^2/c^2$. Setting $E_0 = m_0 c^2$ and neglecting higher-order terms in the expansion gives the following *approximate* expression, valid for speeds v much less than c :

$$E \approx \frac{1}{2} m_0 v^2 + m_0 c^2. \quad (4.12)$$

This is just the Newtonian kinetic energy (based on the rest mass) plus the rest energy. Note that we cannot apply this approximation in the case of photons because photons only travel at the speed of light (in whatever medium they're moving in) and because they have zero rest mass. Thus, we see that the Newtonian kinetic energy $\frac{1}{2}mv^2$ and, by the same token, $p = mv$, are limiting expressions *applicable only for systems moving with non-relativistic speeds*.

In fact, it's interesting to note that the correct result $E = pc$ was already well known to physicists in the early twentieth century based on the relationship between electromagnetic energy flux and momentum density in Maxwell's theory of electrodynamics.² It can be derived without recourse to special relativity or the presumption of photons with zero rest mass, and has been verified by experiments on radiation pressure.

Equations (4.6) and (4.9) allow us to deduce the relativistic energy and momentum in a reference frame S moving with velocity v relative to the rest frame, which we will call S_0 . This is all fine, but in what follows we will need to refer to a more general Lorentz transformation for energy and momentum between the arbitrary frames S and S' , where v is now the relative velocity of the two frames. This can be quite easily done. We begin by using the general Lorentz transformation for position and time (the right-hand side of Table 2.1) to establish relationships between all the velocities involved and, most importantly, their corresponding Lorentz factors. We then use Eqs. (4.6) and (4.9) to transform energy and momentum from S_0 to S and from S_0 to S' and use the relationships between the Lorentz factors to compare the two sets of results for S and S' . This gives us the Lorentz transformation equations for energy and momentum.

Although quite straightforward, the algebraic manipulations involved are relatively uninformative and, frankly, rather tedious. I've therefore chosen to confine this derivation to Appendix 2, and will here proceed directly to the end results. For a system in S with energy E and momentum p , which is resolved into Cartesian components p_x, p_y , and p_z (such that $p^2 = p_x^2 + p_y^2 + p_z^2$), the Lorentz transformations to another frame S' (energy E' and momentum p') are

$$E' = \gamma (E - vp_x) \quad (4.13)$$

$$p'_x = \gamma \left(p_x - \frac{v}{c^2} E \right) \quad (4.14)$$

$$p'_y = p_y \quad (4.15)$$

$$p'_z = p_z, \quad (4.16)$$

where v is the velocity of S' relative to S and $p'^2 = p_x'^2 + p_y'^2 + p_z'^2$. We can see immediately that when S is the rest frame S_0 , $p_x = 0$ and (4.13) reduces to (4.6). From (4.14) we have (since $E_0 = m_0c^2$) $p'_x = -\gamma m_0v$. This is equivalent to (4.9) as in this case $p'^2 = p_x'^2$ (as $p_y'^2 = p_z'^2 = 0$) and we can choose to take the positive root for p' .

Lorentz Transformation for a System of Plane Waves

Before we can go on to consider the main derivation of this chapter we need one last ingredient—the Lorentz transformation for a system of plane waves. In fact, we have already considered this problem in Chapter 2—see Eqs. (2.5)–(2.18) and Fig. 2.3. However, this referred to a plane *electromagnetic* wave moving at the speed of light. What we need here is a more general transformation for plane waves moving at any speed.

But the logic is nevertheless much the same. In the frame S we represent the wave as $A \sin(\omega t - kr)$, where r is the direction of a sine wave travelling towards the observer, ω is the angular frequency, and k is the wave vector. We resolve the wave vector along the three Cartesian coordinates with components k_x , k_y , and k_z (where $k^2 = k_x^2 + k_y^2 + k_z^2$), and rewrite the phase angle as

$$\Phi = \omega t - (k_x x + k_y y + k_z z). \quad (4.17)$$

Likewise, in the frame S'

$$\Phi' = \omega' t' - (k'_x x' + k'_y y' + k'_z z'). \quad (4.18)$$

We proceed as before, applying the Lorentz transformation for t' , x' , y' , and z' to Eq. (4.18) and gathering together terms in t and x to give

$$\Phi' = \gamma (\omega' + vk'_x) t - \left[\gamma \left(k'_x + \frac{v}{c^2} \omega' \right) x + k'_y y + k'_z z \right]. \quad (4.19)$$

We now set $\Phi' = \Phi$ and compare (4.17) and (4.19) term by term:

$$\gamma (\omega' + vk'_x) t = \omega t \quad (4.20)$$

$$\gamma \left(k'_x + \frac{v}{c^2} \omega' \right) x = k_x x \quad (4.21)$$

$$k'_y y = k_y y \quad (4.22)$$

$$k'_z z = k_z z. \quad (4.23)$$

As before, we use (4.21) to derive an expression for k'_x which we substitute into (4.20) giving, after some rearrangement,

$$\omega' = \gamma (\omega - vk_x). \quad (4.24)$$

Substituting this expression for ω' into (4.21) gives, again after some rearrangement,

$$k'_x = \gamma \left(k_x - \frac{v}{c^2} \omega \right). \quad (4.25)$$

And, of course, from (4.22) and (4.23),

$$k'_y = k_y \quad (4.26)$$

$$k'_z = k_z. \quad (4.27)$$

We're now all set. But before going on to consider the main derivation I want to pause to consider de Broglie's original.

De Broglie's Original Derivation

Chapter 1 of de Broglie's PhD thesis opens with a discussion of Einstein's $E = mc^2$ and the Planck–Einstein relation, $E = h\nu$. He writes:³

The notion of a quantum makes little sense, seemingly, if energy is to be continuously distributed through space; but, we shall see that this is not so. One may imagine that, by cause of a meta law of Nature, to each portion of energy with a proper mass m_0 , one may associate a periodic phenomenon of frequency ν_0 , such that one finds. . .

$$E_0 = m_0c^2 = h\nu_0 \quad (4.28)$$

valid in the rest frame, S_0 .

De Broglie assumed that the periodic phenomenon in question is a standing wave somehow confined *within* the particle, and which is carried along with the particle moving at a velocity v . However, when he then tried to generalize this for moving frames, he hit a problem. He figured that in a frame S moving with velocity v relative to the rest frame, the observed 'periodic phenomenon of frequency ν_0 ' would be affected by time dilation according to

$$\nu_i = \nu_0 \frac{t_0}{t} = \frac{\nu_0}{\gamma}, \quad (4.29)$$

where t_0 is the time measured in the rest frame, $t = \gamma t_0$, and ν_i is the observed frequency of the 'inner' standing wave. Because of the effects of time dilation, the frequency of the inner wave is observed to be *less* than ν_0 . But the equation for the relativistic energy, (4.6), instead implies that

$$E = h\nu = \gamma h\nu_0 \quad \text{or} \quad \nu = \gamma \nu_0, \quad (4.30)$$

which implies a frequency *greater* than ν_0 . These two frequencies can't be the same. The only way de Broglie could reconcile them was further to assume the existence of

an associated travelling *external* wave (hence, v instead of v_i) which remains in phase with the inner wave but possesses a *phase velocity* V which is different from v . De Broglie called this a ‘theorem of phase harmony’:⁴

A periodic phenomenon is seen by a stationary observer to exhibit the frequency [$\nu_i = \nu_0/\gamma$] that appears constantly in phase with a wave having frequency [$\nu = \gamma\nu_0$] propagating in the same direction with velocity [$V = c^2/v$].

De Broglie arrived at this result for the phase velocity as follows. We assume that the inner and external waves start out in phase, with the particle and the external wave moving in tandem along the x -coordinate. Now, assuming a simple sinusoidal form for the inner wave, the phase is given by $\omega_i t$, where the angular frequency $\omega_i = 2\pi\nu_i$ (this internal wave is a standing wave, so the phase depends only on the angular frequency). The phase of the travelling external wave is $\omega t - kx$, where k is the wave vector. The theorem of phase harmony then requires

$$\omega_i t = \omega t - kx \quad \text{or} \quad \nu_i t = \nu \left(t - \frac{x}{V} \right), \quad (4.31)$$

where we have substituted $k = 2\pi/\lambda = 2\pi\nu/V$ and cancelled all the factors of 2π . Within the time t the particle moves a distance $x = vt$ and so we can substitute $t = x/v$ in (4.31) to give

$$\nu_i \frac{x}{v} = \nu \left(\frac{x}{v} - \frac{x}{V} \right) = \nu \frac{x}{v} \left(1 - \frac{v}{V} \right) \quad \text{or} \quad \frac{\nu_i}{\nu} = \left(1 - \frac{v}{V} \right) \quad (4.32)$$

We can now use the results from Eqs. (4.29) and (4.30) to deduce an expression for v :

$$\frac{\nu_0}{\gamma^2 \nu_0} = \left(1 - \frac{v}{V} \right) \quad \text{or} \quad \left(1 - \frac{v^2}{c^2} \right) = \left(1 - \frac{v}{V} \right). \quad (4.33)$$

From which it follows that $V = c^2/v$.

This means that the phase velocity of the travelling external wave is greater than the speed of light, so this can't be a ‘real’ wave capable of transporting energy. De Broglie writes: ‘Our theorem teaches us, moreover, that this wave represents a spatial distribution of phase, that is to say, it is a “*phase wave*”.’⁵ Observers of a crowd at a sports stadium will appreciate that a ‘Mexican wave’, created by spectators standing and raising their arms before sitting down synchronously with their neighbours, travels around the stadium at a speed much greater than that of the spectators themselves.

We can now complete de Broglie's original derivation. Recall from Eq. (4.30) that the relativistic energy is given by $E = h\nu = \gamma m_0 c^2$. From Eq. (4.9) we know that the relativistic momentum is $p = \gamma m_0 v$. We can combine these with the expression for the

phase velocity V to give

$$p = \frac{h\nu}{c^2}v = \frac{h\nu}{V}. \quad (4.34)$$

But ν/V is simply the reciprocal of the wavelength (λ) of the external wave, so

$$p = \frac{h}{\lambda} \text{ or } \lambda = \frac{h}{p}. \quad (4.35)$$

De Broglie didn't immediately recognize this as the iconic equation of wave-particle duality (it makes an appearance in his PhD thesis only in the last chapter, in a discussion of quantum statistical mechanics). However, in the Nobel Prize lecture which he delivered in 1929, he declared it 'a fundamental relation of the theory'.⁶

The assumption of faster-than-light phase waves might appear somewhat arbitrary, or even unphysical, but de Broglie nevertheless attached great significance to them. He went on to show how the frequencies ν_0 and ν combine to define a *refractive index* for free space and he equated the velocity of the particles ν with the 'group velocity' of a 'packet' of waves with similar frequencies.

He refined his understanding some years later. For de Broglie, wave-particle duality was a duality of waves *and* particles, the external wave serving to guide or pilot the motion of the particle with which it is associated.⁷ It is this external 'pilot wave' that is responsible for phenomena such as electron diffraction and interference, with particles following trajectories governed by the resulting pattern of wave intensities. But, as we will see in what follows, this duality would become much more commonly ascribed to one of waves *or* particles. In this interpretation, material particles such as electrons show *either* wave behaviour (diffraction and interference) *or* particle behaviour (mass, linear momentum, defined trajectory), depending on the nature of the experiment used to observe them. But they cannot show both types of behaviour simultaneously.

If you think that this latter interpretation suggests that it should be possible to derive de Broglie's relation without recourse to faster-than-light waves, then you would be right. There is a much more direct derivation which is actually simpler and which demonstrates quite clearly the roles of both the Planck-Einstein relation and special relativity. Consequently, I have chosen this as the main derivation for this chapter.

The Ingredients

1. The Planck-Einstein relation $E = h\nu$.
2. The Lorentz transformation for energy and linear momentum.
3. The Lorentz transformation for a system of plane waves.

The Recipe

Step (1) involves an almost trivial derivation of $\lambda = h/p$ for photons, by comparing the energy derived from special relativity with the Planck–Einstein relation. This emphasizes the dual wave–particle nature of electromagnetic radiation—a wave-like property (wavelength) is inversely related to a particle-like property (linear momentum). But it is actually Step (2) that is quite breath-taking. This involves a generalization of the derivation for matter particles. It follows much the same logic as Step (1), but requires the full Lorentz transformations for both energy and linear momentum *and* a system of plane waves. Applying the Planck–Einstein relation to material particles *assumes* that such particles possess associated wave-like properties, such as frequency. Comparison of these two sets of transformations then leads directly to a much more general derivation of $\lambda = h/p$ for matter particles.

Step (1): Wave–Particle Duality in Photons

This is likely to be one of the simplest derivations in the entirety of physics. We take the energy of a photon derived from special relativity (or, if you prefer, from Maxwell’s classical electrodynamics), Eq. (4.10), and relate this directly with the energy of a photon based on the Planck–Einstein relation:

$$E = pc = hv = \frac{hc}{\lambda} \quad \text{hence} \quad \lambda = \frac{h}{p}. \quad (4.36)$$

That’s all there is to it.

But, of course, it was de Broglie’s next step that was astounding. If electromagnetic waves, characterized by a frequency ν , possess associated particle-like properties such as linear momentum, then, de Broglie reasoned, perhaps particles such as electrons, characterized by a mass m and linear momentum p , might possess associated wave-like properties. He continued:⁸

An electron is for us the archetype of [an] isolated parcel of energy, which we believe, perhaps incorrectly, to know well; but, by received wisdom, the energy of an electron is spread over all space with a strong concentration in a very small region, but otherwise whose properties are very poorly known. That which makes an electron an atom of energy is not its small volume that it occupies in space, I repeat: it occupies all space, but the fact that it is undividable, that it constitutes a unit.

Step (2): Lorentz Transformations for Particles and Waves

It will be helpful at this stage to draw together the earlier results for the Lorentz transformation for energy and linear momentum and a system of plane waves; see Table 4.1.

Table 4.1 Lorentz Transformations in the Standard Configuration

Particles Energy and momentum		Waves Angular frequency and wave vector	
From S to S'	From S' to S	From S to S'	From S' to S
$E' = \gamma(E - vp_x)$	$E = \gamma(E' + vp'_x)$	$\omega' = \gamma(\omega - vk_x)$	$\omega = \gamma(\omega' + vk'_x)$
$p'_x = \gamma(p_x - \frac{v}{c^2}E)$	$p_x = \gamma(p'_x + \frac{v}{c^2}E')$	$k'_x = \gamma(k_x - \frac{v}{c^2}\omega)$	$k_x = \gamma(k'_x + \frac{v}{c^2}\omega')$
$p'_y = p_y$	$p_y = p'_y$	$k'_y = k_y$	$k_y = k'_y$
$p'_z = p_z$	$p_z = p'_z$	$k'_z = k_z$	$k_z = k'_z$

The similarity of these two sets of transformations is rather obvious, and we could complete the derivation simply by inspection. Nevertheless, we can still proceed quite quickly.

If we apply the Planck–Einstein relation for both E and E' ,

$$E = hv = \hbar\omega \quad \text{and} \quad E' = hv' = \hbar\omega', \quad (4.37)$$

we can rewrite Eq. (4.13) as follows:

$$\hbar\omega' = \gamma(\hbar\omega - vp_x). \quad (4.38)$$

Likewise, if we multiply (4.24) by the reduced Planck constant \hbar we get

$$\hbar\omega' = \gamma(\hbar\omega - v\hbar k_x). \quad (4.39)$$

From a comparison of (4.38) and (4.39) we deduce that

$$p_x = \hbar k_x. \quad (4.40)$$

In the standard configuration we assume that the motion of S' relative to S is confined to the x -direction. But we can obviously follow precisely the same logic by confining the motion separately to the y -direction and the z -direction to conclude that

$$p_y = \hbar k_y \quad \text{and} \quad p_z = \hbar k_z. \quad (4.41)$$

From which we deduce that

$$p^2 = p_x^2 + p_y^2 + p_z^2 = \hbar^2 (k_x^2 + k_y^2 + k_z^2) = \hbar^2 k^2. \quad (4.42)$$

Taking the positive root gives

$$p = \hbar k = \frac{h}{\lambda}, \text{ hence } \lambda = \frac{h}{p}. \quad (4.43)$$

The significance of this last expression is that, unlike (4.36), it applies to *any* system of plane waves, including ‘matter waves’ that may not travel at the speed of light.

Matter Waves

Whatever they were, de Broglie’s matter waves could not be considered to be in any way equivalent to more familiar wave phenomena, such as sound waves or ripples on the surface of a pond. But the simple fact that they can nevertheless be described in terms of the mathematics of wave motion suggested a further insight. Musical notes produced by string or wind instruments are the result of standing waves—a variety of standing wave patterns is possible provided they fit between the string’s secured ends or the ends of the pipe. This means that they must have zero amplitude at each end. This is possible only for vibrational patterns which contain an integral number of half-wavelengths (see Fig. A1.1 in Appendix 1).

De Broglie saw that the quantum numbers in Bohr’s theory of atomic structure might emerge naturally from a model in which an electron wave is confined to a circular orbit around the nucleus. Perhaps, he reasoned, the stable electron orbits of Bohr’s theory represented standing electron waves, just like the standing waves in strings and pipes responsible for musical notes. The same kinds of arguments would apply. For a standing wave to be produced in a circular orbit, the electron wavelengths must fit exactly within the orbit circumference.

Put simply, by the time the electron wave has performed one complete orbit and returned to its starting point, the value of the associated wave amplitude and the phase must be the same as at the starting point. If this is not the case, the wave will not ‘join up’ with itself: it will interfere destructively and no standing wave will be produced—Fig. 4.2.

To satisfy the requirement, the wavelength of the electron wave must be such that an *integral* number of wavelengths will fit into the circumference of the orbit. This is called the resonance or phase condition. Bohr’s quantum numbers could therefore be thought of as the number of wavelengths of the electron wave present in each orbit.

‘Thus,’ wrote de Broglie, ‘the resonance condition can be identified with the stability condition from quantum theory.’⁹

De Broglie published his ideas in a series of three short papers in the *Comptes Rendus* (proceedings) of the Paris Academy in September and October 1923. He collected these papers together, extended his ideas, and submitted them as a PhD thesis to the Faculty of Science at Paris University. He produced three typed copies and passed one of these to the prominent French physicist Paul Langevin, who was to act as an examiner. De Broglie’s other examiners were Jean Perrin, Élie Cartan, and Charles Mauguin.

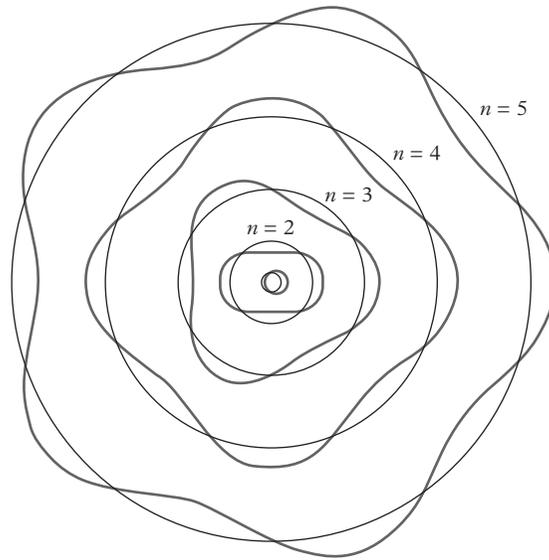


Figure 4.2 De Broglie imagined that the quantum numbers of Bohr's atomic energy levels might arise from the condition for standing electron waves confined to each orbit.

Source: Adapted from <https://skullsinthestars.com/2015/05/20/1975-the-year-that-quantum-mechanics-met-gravity/>.

Langevin didn't know quite what to make of de Broglie's bold ideas, which he thought were rather far-fetched, and sought guidance from Einstein, by now professor at the University of Berlin. Einstein asked for a further typed copy to be sent to him. Mauguin didn't believe a word of it. Perrin could only comment that: 'All I can tell you is that your brother [Maurice] is very intelligent.'¹⁰

But Einstein may have recognized a little of himself in de Broglie's revolutionary ideas. He wrote back to Langevin offering encouragement: 'He [de Broglie] has lifted a corner of the great veil.'¹¹ This was enough for Langevin. He accepted the thesis and de Broglie was duly awarded his doctorate in November 1924. De Broglie's thesis was published in its entirety in the French scientific journal *Annales de Physique* in 1925.

Not everyone was convinced. Most physicists were rather sceptical: 'some wit, in fact, dubbed de Broglie's theory "la Comédie Française".'¹²

But experimental evidence for the dual wave-particle nature of electrons was quite quickly forthcoming. In a series of experiments over the period 1926-7, Clinton Davisson and Lester Germer showed that a beam of electrons passed through a single crystal of nickel exhibit diffraction phenomena, with a wavelength very close to that predicted using the de Broglie relation. Diffraction of electrons by thin metal films was demonstrated at around the same time by English physicist George P. Thomson (son of J. J.). Davisson and Thomson shared the 1937 Nobel Prize in Physics. Thus it was that J. J. Thomson won the Nobel Prize for showing that the electron is a particle, but his son G. P. Thomson won it for showing that the electron is a wave.

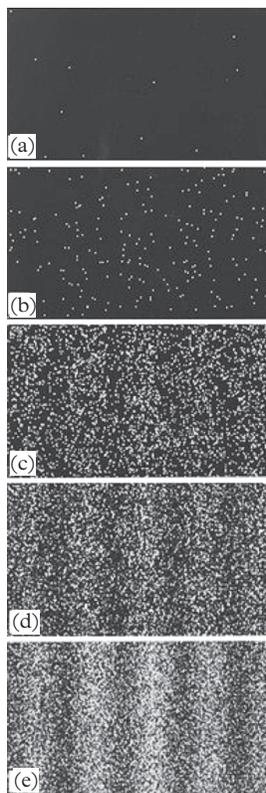


Figure 4.3 We can observe electrons as they pass, one at a time, through a two-slit apparatus by recording where they strike the screen. Photographs (a) to (e) show the resulting images when, respectively, 10, 100, 3,000, 20,000, and 70,000 electrons have been detected.

Source: Reproduced from Tonomura, A. et al., Demonstration of single-electron buildup of an interference pattern, *American Journal of Physics* 57, 117 (1989), with the permission of the American Association of Physics Teachers.

In experiments that continue to haunt anyone who has paused to reflect on what quantum theory tells us about the nature of physical reality, in 1989 a group of Japanese physicists reported a series of interference patterns derived from passing a beam of electrons through two parallel slits (Fig. 4.3). Be assured—we will be returning to ponder on what this sequence of pictures might be telling us.

More Alarm Bells

De Broglie's ideas were illuminating but they were far from representing a solution. Assuming that material particles like electrons could possess associated wave-like properties had led de Broglie to identify the quantum numbers of Bohr's theory with the

resonance condition for standing waves. But this was nothing more than an idea. De Broglie had not *derived* the quantum numbers from some kind of wave theory of the electron in an atom. And he had as yet no explanation for the mysterious quantum jumps:¹³

From this we see why certain orbits are stable; but, we have ignored passage from one to another stable orbit. A theory for such a transition can't be studied without a modified version of electrodynamics, which so far we do not have.

Einstein was already concerned about the implications of quantum jumps for the principles of causality, the notion that every effect in the physical universe is traceable to a cause. It was clear that an electron in a high-energy atomic orbit will fall to a more stable, lower-energy orbit. It is caused to do so by the mechanics of the atom.

But, just as Rutherford had noted, it seemed that the electron needed to 'know' in advance which orbit it was jumping into. And the precise moment of the jump and the direction of the consequently emitted photon also appeared to be left entirely to chance. These were things that could not be predicted.

Einstein was not at all comfortable with this. In 1920 he had written to German physicist Max Born on the subject, noting that he: 'would be very unhappy to renounce complete causality'.¹⁴ From the very beginning, Einstein had viewed the quantum hypothesis as provisional, to be eventually replaced by a new, more complete theory that would properly explain quantum phenomena.

After pioneering quantum theory through one of its most testing early periods, Einstein was beginning to have grave doubts about what it meant.

.....

NOTES

1. Louis de Broglie, from the 1963 re-edited version of his PhD thesis. Quoted in Abraham Pais, *Subtle is the Lord: The Science and the Life of Albert Einstein*. Oxford University Press, Oxford, 1982, p. 436.
2. See, for example, Fulvio Melia, *Electrodynamics*, University of Chicago Press, Chicago, 2001, pp. 73 and 131.
3. Louis de Broglie, 'Recherches sur la Théorie des Quanta', PhD thesis, Faculty of Science, Paris University, 1924. English translation by A. F. Kracklauer, p. 8.
4. *Ibid.* 9. Italics in the original.
5. *Ibid.* 10.
6. Louis de Broglie, 'The Wave Nature of the Electron', Nobel Prize lecture, 12 December 1929, *Nobel Lectures, Physics 1922-1941*, Elsevier, Amsterdam, 1965, p. 249. Chapter IV, Section I of Louis de Broglie, *Matter and Light: The New Physics*, translated by W. H. Johnston, Norton, New York, 1939, pp. 165-79, is very closely based on his Nobel lecture.
7. This is de Broglie's 'double solution' interpretation. See Max Jammer, *The Philosophy of Quantum Mechanics*, Wiley, New York, 1974, pp. 44-9.
8. De Broglie, 'Recherches sur la Théorie des Quanta', p. 8.

9. Ibid. 29.
10. Jean Perrin, quoted in Anatole Abragam, 'Louis Victor Pierre Raymond de Broglie, 15 August 1892–19 March 1987', *Biographical Memoirs of Fellows of the Royal Society*, 34 (1988), 30.
11. Albert Einstein, letter to Paul Langevin, 16 December 1924. Quoted in Walter Moore, *Schrödinger: Life and Thought*, Cambridge University Press, Cambridge, UK, 1989, p. 187.
12. George Gamow, *Thirty Years that Shook Physics*, Dover, New York, 1966, p. 81.
13. De Broglie, 'Recherches sur la Théorie des Quanta', p. 29.
14. Albert Einstein, letter to Max Born, 27 January 1920. Quoted in Pais, *Subtle is the Lord*, p. 412.

.....

FURTHER READING

- Abragam, Anatole, 'Louis Victor Pierre Raymond de Broglie, 15 August 1892–19 March 1987', *Biographical Memoirs of Fellows of the Royal Society*, 34 (1988), 21–41.
- Broglie, Louis de, *Matter and Light: The New Physics*, translated by W. H. Johnston, Norton, New York, 1939, particularly Chapter IV.
- Gamow, George, *Thirty Years that Shook Physics*, Dover, New York, 1966, Chapter IV.
- Moore, Walter, *Schrödinger: Life and Thought*, Cambridge University Press, Cambridge, UK, 1989, pp. 183–7.
- Pais, Abraham, *Subtle is the Lord: The Science and the Life of Albert Einstein*. Oxford University Press, 1982, Chapter 24.
- Pais, Abraham, *Niels Bohr's Times, in Physics, Philosophy and Polity*, Oxford University Press, Oxford, 1991, pp. 239–41.

5

Schrödinger's Derivation of the Wave Equation

Quantization as an Eigenvalue Problem

Bohr's new theory of the atom was in trouble just as soon as it was formulated. Although there could be no doubting that it contained part of some elusive truth, it was far from the whole story. Already in 1887 Albert Michelson and Edward Morley (of speed of light fame) had reported that, when studied under high resolution, the hydrogen atom H_α emission line has a *fine structure*. It is, in fact, not one line, but two. It may be that Bohr was aware of this result, but he did not refer to it in his trilogy. The H_β line had also been shown to exhibit a splitting of similar magnitude.

In Bohr's theory, the H_α and H_β lines arise, respectively, from transitions from the higher energy levels $n_2 = 3$ and $n_2 = 4$ to the lower level $n_1 = 2$. Bohr wondered if interactions with electric fields or the nucleus might be responsible, or if the relativistic mass of the electron moving in an elliptical orbit might account for the splitting.

German physicist Arnold Sommerfeld showed that if the orbits are assumed to be elliptical, then another quantum number is needed to characterize the orbit eccentricity. He called the new quantum number k , referred to as the auxiliary or azimuthal quantum number, constrained to the integer values $0, 1, 2, \dots, n$, where n is now referred to as the principal quantum number. When combined with special relativity this seemed to work, and the fact that Bohr had assumed circular orbits (for which $k = n$) did not invalidate any of his earlier conclusions, although he was later (in 1920) forced to abandon the assumption that the orbital angular momentum is constrained to the values $n\hbar$, Eq. (3.23).

But Sommerfeld wasn't quite done. The lines in the hydrogen atom emission spectrum were found to split further in the presence of an electric field (this is the Stark effect, named for Johannes Stark). To explain this, Sommerfeld introduced a third quantum number, which he designated as m . These new quantum numbers were thought to be related in some way to the quantization of the *geometry* of the electron orbits.

These were still classical models of the atom, to which a system of 'quantum rules' had been imposed. Confusion reigned. In his biography of Bohr, Pais explained:¹

As in any crisis, reactions varied quite considerably. Many preferred to sit this one out until it would go away. Those who had the courage (I cannot find a better word) to persist tried, in a variety of ways depending on personal temperament, to move ahead in this muddled territory It was not ideal but there was nothing better—yet.

And indeed the Bohr–Sommerfeld model was successful for a time. In addition to the fine structure and the Stark effect, it explained the splitting of spectral lines in a weak magnetic field (the ‘normal’ Zeeman effect, named for Pieter Zeeman). But it was all rather ad hoc. Although it would turn out that Sommerfeld had arrived at the correct formula for the fine structure splitting, his reasoning was not correct. His derivation would later be considered as ‘perhaps the most remarkable numerical coincidence in the history of physics’.²

It couldn’t last, and by 1925 the theory was in real crisis. Spectroscopists could find no evidence for transitions involving energy levels with $k = 0$, although there was no theoretical reason why they should be excluded, or ‘forbidden’. The theory certainly could not explain the further splitting seen in a strong magnetic field, which became known as the ‘anomalous’ Zeeman effect, although it is, in fact, more common than the ‘normal’ effect. And attempts to apply the model to heavier atoms such as helium were wholly unsuccessful.

Quantum Numerology

We’ll reserve discussion of both the normal and anomalous Zeeman effects for a later chapter. For now, let’s focus on the quantum numbers n , k , and m . Over time, these became the *principal* quantum number n , k was replaced by the *orbital angular momentum* quantum number l ($= k - 1$), with values $0, 1, 2, \dots, n - 1$, and m was replaced by the *magnetic* quantum number m_l , with $2l + 1$ values ranging in integral steps from $-l, -(l - 1), \dots, 0, \dots, (l - 1), l$.^{*} Spectroscopists had designated the spectral lines as ‘sharp’, ‘principal’, ‘diffuse’, and ‘fine’, based on how they appeared, and this naming convention was carried over to the energy levels involved, as ‘s’ ($l = 0$), ‘p’ ($l = 1$), ‘d’ ($l = 2$), and ‘f’ ($l = 3$).

The combinations of the quantum numbers suggested that each principal quantum number n is associated not with a single energy level, as in Bohr’s older model, but rather with n^2 levels. For $n = 1$, l can take only the value 0 ($n - 1$) and so $m_l = 0$, giving 1 ($= 1^2$) level. This lowest energy level is designated 1s. For $n = 2$, l can take the values 0 (2s) and 1 , and for $l = 1$ (2p), m_l can take three values, $-1, 0$, and 1 , giving a total of 4 ($= 2^2$) levels. For $n = 3$ there is one 3s, three 3p, and five 3d levels, making 9 ($= 3^2$) in total. Some of these levels are illustrated in Fig. 5.1.

This figure also shows how the levels involved in the H_α transition result in the appearance of two lines (called a ‘doublet’). Two of the lower 2p levels are degenerate—

* Note that switching to l in preference to k immediately rules out energy levels characterized by values of l less than 0, such as $k = 0$.

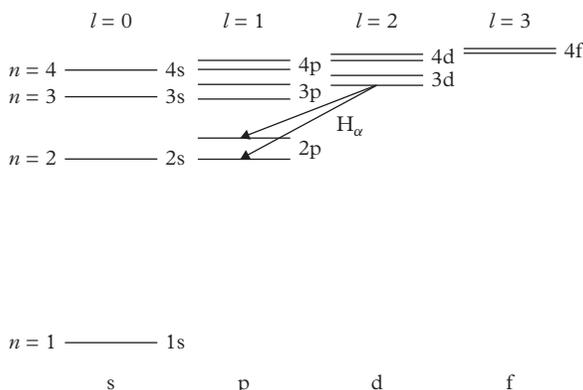


Figure 5.1 A selection of the energy levels of the hydrogen atom, showing the origin of fine structure (not to scale). Note how the magnitude of the splitting between levels declines as n increases.

they have the same energy, and these are split from the third for reasons we will come to understand in a later chapter.

Aside from this kind of quantum ‘numerology’, it was obvious from the hydrogen atom spectrum that there must be some rules governing which transitions would appear and which were, for some reason, forbidden. There was clearly no constraint on changes to the principal quantum number—transitions between any pair of values of n are possible, as Fig. 3.3 illustrates. But this is not true for l : only transitions involving an increase or decrease of 1 appear in the spectrum, suggesting the ‘selection rule’ $\Delta l = \pm 1$. Thus, transitions $3d \rightarrow 2p$ are allowed (as shown in Fig. 5.1), but $3p \rightarrow 2p$ are not. For m_l the selection rule is $\Delta m_l = 0, \pm 1$, depending of the polarization of the light absorbed or emitted. There was no real explanation for any of these rules.

But the immediate concern was the quantum numbers themselves. Where did they come from? Could de Broglie’s outrageous hypothesis shed any light?

Enter Schrödinger

By this time the Austrian physicist Erwin Schrödinger had established a reputation as a competent, if somewhat unspectacular, physicist. He had drawn praise for his versatility and the breadth of his knowledge and accomplishments, but he had yet to make a noteworthy contribution to any branch of physical science with which he was familiar. As he grew older, he had little choice but to watch as a younger generation of physicists overtook him. Earlier in 1925 he had celebrated his 38th birthday, and it was common knowledge that many (though certainly not all) physicists whose efforts had merited Nobel Prize-level recognition had made their breakthroughs when young (Einstein had been 26 in his ‘miracle year’ of 1905).

It seemed that Schrödinger would be sidelined, worthy of little more than a footnote in the history of physics. When, in 1924, he was invited to attend the fourth in a series of

prestigious conferences on physics established by the wealthy Belgian industrialist Ernest Solvay, he was not asked to present a paper.

But in October 1925 his attention was drawn to a footnote in one of Einstein's recent papers. It mentioned 'a very notable contribution' by de Broglie. Intrigued, Schrödinger acquired a copy of de Broglie's PhD thesis.

Schrödinger held a professorship at the University of Zurich, and physicists in his department had established biweekly colloquia jointly with colleagues at the neighbouring ETH,* at which they would discuss topics of mutual interest. Pieter Debye asked Schrödinger if he would be prepared to present a colloquium on de Broglie's thesis, which had recently been published in the French journal *Annales de Physique*. Schrödinger agreed.

The seminar was held on 23 November. In the audience was a young Swiss student called Felix Bloch, who recalled the event more than 50 years later:³

Schrödinger gave a beautifully clear account of how de Broglie associated a wave with a particle and how he could obtain the quantization rules of Niels Bohr and Sommerfeld by demanding that an integer number of waves should be fitted along a stationary orbit. When he had finished, Debye casually remarked that this way of talking was rather childish. As a student of Sommerfeld he had learned that, to deal properly with waves, one had to have a wave equation.

A few days before Christmas 1925, Schrödinger left Zurich for a short vacation in the Swiss Alps, at a resort near Davos. His relationship with his wife Anny was in deep trouble, so he chose to invite an old girlfriend from Vienna to join him. He also took with him his notes on de Broglie's thesis. We do not know who the girlfriend was or what influence she might have had on him, but when he returned on 8 January 1926, he had discovered *wave mechanics*.

In order to follow Schrödinger's logic, we first need to become a little more familiar with the equations of classical wave motion.

The Classical Wave Equation

We have already encountered the one-dimensional wave equation in the Prologue, as Eq. (P.19):

$$\frac{\partial^2}{\partial x^2} \Psi(x, t) = \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \Psi(x, t), \quad (5.1)$$

where v is the wave velocity in the x -direction. Now, we could take this equation on trust and solve it to find an expression for the wavefunction, $\Psi(x, t)$. Or, we could choose just to see how it works by reaching for a handy, already-known solution.

* ETH stands for the Eidgenössische Technische Hochschule, formerly the Zurich Polytechnic which Einstein had attended to study for his degree.

I've opted for the latter, so let's use the general plane wave solution, $\Psi(x, t) = Ae^{i(k_x x - \omega t)}$, where, as before, A is the peak amplitude, $k_x (= 2\pi/\lambda)$ is the wave vector in the x -direction, $\omega (= 2\pi\nu)$ is the angular frequency, and $i = \sqrt{-1}$. This wavefunction has the general form $Ae^{i\theta}$, where θ is the phase angle, equal to $(k_x x - \omega t)$. We can see the connection with sinusoidal wave motion more clearly if we plot this function in the complex plane, with one 'imaginary' axis and one 'real' x -axis—see Fig. 5.2(a). As θ increases from 0° to 360° (2π), it describes a complete circle in the complex plane and traces a single sine wave with wavelength λ , Fig. 5.2(b).

We could obviously have chosen a simple (non-complex) solution such as $A \sin(k_x x - \omega t)$, so the introduction of a complex solution containing i might seem

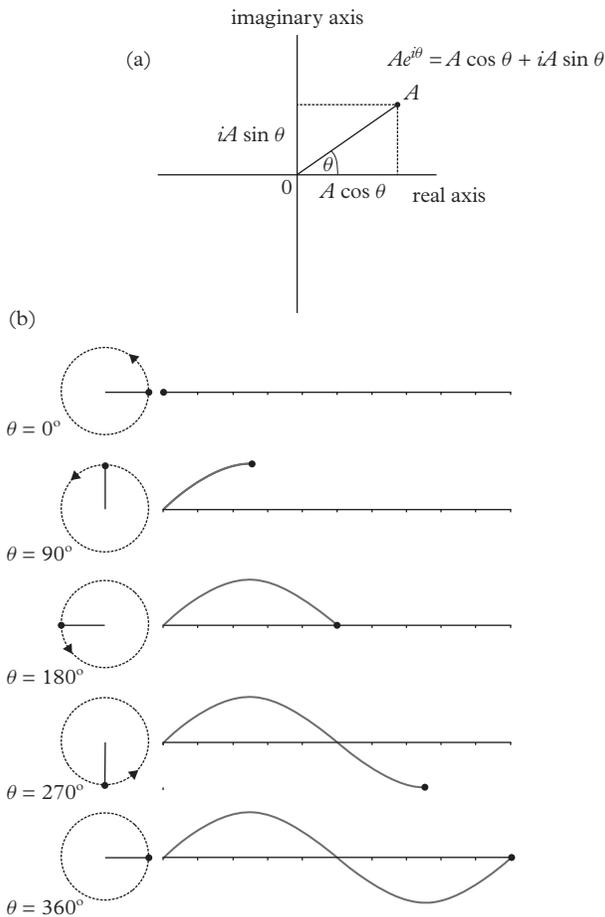


Figure 5.2 The relationship between circular movement of a point in the complex plane and wave motion.

an unnecessary complication. But please bear with me—we will soon see why this is both necessary and important.

We can see that this particular choice of solution means that $\Psi(x, t)$ can be written as the product of separate and distinct spatial and time-dependent functions:

$$\Psi(x, t) = \psi(x)\varphi(t), \text{ where } \psi(x) = Ae^{ik_x x} \text{ and } \varphi(t) = e^{-i\omega t}, \quad (5.2)$$

and this allows us to decompose the partial differential wave Eq. (5.1) into two *differential* equations:

$$\frac{\partial}{\partial x}\Psi(x, t) = \varphi(t)\frac{d}{dx}\psi(x), \text{ where } \frac{d}{dx}\psi(x) = ik_x\psi(x) \quad (5.3)$$

from which it follows that

$$\frac{\partial^2}{\partial x^2}\Psi(x, t) = \varphi(t)\frac{d^2}{dx^2}\psi(x), \text{ where } \frac{d^2}{dx^2}\psi(x) = -k_x^2\psi(x) \quad (5.4)$$

(remember, $i^2 = -1$). Similarly,

$$\frac{\partial}{\partial t}\Psi(x, t) = \psi(x)\frac{d}{dt}\varphi(t), \text{ where } \frac{d}{dt}\varphi(t) = -i\omega\varphi(t), \quad (5.5)$$

and

$$\frac{\partial^2}{\partial t^2}\Psi(x, t) = \psi(x)\frac{d^2}{dt^2}\varphi(t), \text{ where } \frac{d^2}{dt^2}\varphi(t) = -\omega^2\varphi(t). \quad (5.6)$$

We note in passing that from (5.4) and (5.6) we have

$$\frac{\partial^2}{\partial x^2}\Psi(x, t) = -k_x^2\psi(x)\varphi(t) = -k_x^2\Psi(x, t), \quad (5.7)$$

$$\frac{\partial^2}{\partial t^2}\Psi(x, t) = -\omega^2\psi(x)\varphi(t) = -\omega^2\Psi(x, t), \text{ and so} \quad (5.8)$$

$$\frac{\partial^2}{\partial x^2}\Psi(x, t) = \frac{k_x^2}{\omega^2}\frac{\partial^2}{\partial t^2}\Psi(x, t) = \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\Psi(x, t), \quad (5.9)$$

where we have made use of the fact that $k_x^2/\omega^2 = 1/\lambda^2 v^2 = 1/v^2$ (remember, $v = \lambda\nu$). This simply proves that the wavefunction $\Psi(x, t)$ as written in Eq. (5.2) is indeed a solution of the classical wave equation. Take it from me that *any* valid solution for $\Psi(x)$ will be consistent with Eqs. (5.7), (5.8), and (5.9).

In what follows much of our attention will be absorbed by the time-independent, spatial function $\psi(x)$. From (5.4) and $k_x = 2\pi/\lambda$, we have

$$\frac{d^2}{dx^2}\psi(x) = -\frac{4\pi^2}{\lambda^2}\psi(x). \quad (5.10)$$

The spatial function $\psi(x)$ is often referred to as ‘the wavefunction’ but, strictly speaking, this is a term that should be reserved for the full time-dependent function $\Psi(x, t)$. Old habits die hard, and I’m sorry to say that I will indeed frequently refer to $\psi(x)$ as the wavefunction. However, as long as we agree not to overlook the important time-dependent component $\varphi(t)$, I believe we should be okay.

All of this discussion has so far centred on the mechanics of classical wave motion, and we now have everything we need to consider Schrödinger’s derivation. But before we begin, I should tell you that it is, in fact, *impossible* to provide a rigorous derivation of Schrödinger’s wave equation starting from classical physics. Just like Planck’s derivation of $E = h\nu$ in Chapter 1, the quantum Schrödinger equation had to be tortured from the classical theories of waves and particle mechanics. Consequently, many textbooks on quantum mechanics present it almost as a postulate or a given, as though handed to science inscribed on a tablet of stone.

But Schrödinger must have got his equation from *somewhere*, and it is indeed possible (and instructive) to follow his reasoning from notebooks he kept at the time. This is more or less what he did.

The Ingredients

1. The classical equation for $\psi(x)$, Eq. (5.10).
2. The de Broglie relation, $\lambda = h/p$.
3. The equation for the classical Hamiltonian, $H = T + V$.
4. The Laplacian operator, $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$, needed when we generalize the wave equation to three dimensions.
5. The Planck–Einstein relation, $E = h\nu = \hbar\omega$.

The Recipe

We apply the de Broglie relation to the classical equation for the spatial function $\psi(x)$ in Step (1), thereby deducing an equation which features the kinetic energy, T . In Step (2), we make use of the classical Hamiltonian to substitute for T and so derive an expression for the ‘wave-mechanical’ Hamiltonian which we then generalize to three dimensions. The form of this equation requires that we also fundamentally reinterpret the nature of the ‘observable’ quantities in the physical mechanics of quantum wave–particles, such as momentum and energy. In Step (3) we then extend the three-dimensional wave equation

to include time, using the Planck–Einstein relation to translate angular frequency into energy. This gives us a time-dependent, wave-mechanical equivalent of the classical wave equation, (5.1).

Step (1): Wave–Particle Momentum and Kinetic Energy

This first step is, once again, extraordinarily simple and straightforward. To create a ‘wave mechanics’, we adopt de Broglie’s wave–particle duality, relating the classical wavelength of the wave to the linear momentum of its associated particle using $\lambda = h/p$. Now, we’re still considering motion only in one dimension, so for consistency we should note that the momentum enters as p_x , the component of p in the x -direction. Then, from Eq. (5.10), we have

$$\frac{d^2}{dx^2} \psi(x) = -\frac{4\pi^2 p_x^2}{h^2} \psi(x) = -\frac{p_x^2}{\hbar^2} \psi(x). \quad (5.11)$$

But we also know that the classical kinetic energy T in the x -direction is given by $p_x^2/2m$, where m is the mass of the associated particle—see Eq. (P.3). And so:

$$\frac{d^2}{dx^2} \psi(x) = -\frac{2m}{\hbar^2} T \psi(x) \quad \text{or} \quad -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = T \psi(x). \quad (5.12)$$

As simple as this is, the result is extraordinary and, as we will soon see, it represents a marked break with classical mechanics.

Step (2): The Hamiltonian

The classical Hamiltonian summarizes the total energy E according to Eq. (P.10),

$$H(= E) = T + V, \quad (5.13)$$

where V is the potential energy, which we assume is independent of time. It follows that $T = E - V$ and we can rewrite Eq. (5.12) as

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = (E - V) \psi(x), \quad (5.14)$$

or

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + V \psi(x) = E \psi(x). \quad (5.15)$$

This is the one-dimensional, time-independent Schrödinger wave equation.

We can now generalize this to three dimensions by introducing a spatial wavefunction $\psi(r)$ where, as usual, the position r can be specified by $r^2 = x^2 + y^2 + z^2$. Equation (5.15) then becomes

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \psi(r) + V\psi(r) = E\psi(r) \quad (5.16)$$

We have reverted to partial differentials since differentiating along the x -direction (for example) now means assuming constant y and z . We can write (5.16) a little more succinctly as

$$-\frac{\hbar^2}{2m} \nabla^2 \psi(r) + V\psi(r) = E\psi(r), \quad \text{where} \quad (5.17)$$

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad \text{is the Laplacian operator.} \quad (5.18)$$

What Eq. (5.17) says is that if we differentiate the spatial wavefunction $\psi(r)$ twice, multiply the result by $-\hbar^2/2m$, and add the potential energy multiplied by the wavefunction, this set of mathematical operations will return the total energy of the system multiplied by the wavefunction.

In other words, provided the wavefunction $\psi(r)$ is a valid solution for whatever physical system is being considered (which obviously depends on the shape of its potential energy function, V), it must somehow ‘contain’ the value for the total energy. In classical mechanics, we get at the total energy simply by multiplying half the mass by the velocity-squared, and adding the potential energy for a particular set of spatial coordinates. But to get at the total energy in wave mechanics we need to perform a *mathematical operation* on the wavefunction that is much more involved than simply multiplying or dividing it. *To extract the value of an observable quantity (in this case E) from the wavefunction, we need to apply its corresponding operator.*

Of course, the notion of mathematical operators operating on a wavefunction which returns some value multiplied by the wavefunction is inherent in the equations of classical wave motion, such as Eq. (5.10). What is entirely novel is that this logic is now brought to bear on quantities more familiarly associated with particles, a direct consequence of the assumption of wave–particle duality.

To avoid confusion with their classical counterparts, we distinguish the wave-mechanical operators by means of a caret, $\hat{}$, sometimes referred to colloquially as a ‘hat’. If the classical Hamiltonian is H , then its wave-mechanical counterpart is \hat{H} , such that

$$\hat{H}\psi(r) = E\psi(r), \quad \text{where} \quad \hat{H} = -\frac{\hbar^2}{2m} \nabla^2 + V. \quad (5.19)$$

\hat{H} is known as the *Hamiltonian operator*.

The analogy with classical mechanics can be made explicit by writing \hat{H} in terms of a linear momentum operator, \hat{p} :

$$\text{Classical: } H = T + V = \frac{p^2}{2m} + V \quad (5.20)$$

$$\text{Quantum: } \hat{H} = \hat{T} + V = \frac{\hat{p}^2}{2m} + V. \quad (5.21)$$

Note that in (5.21) the significance of \hat{p}^2 is that the momentum operator is applied twice, just as we generate a second-order differential by differentiating a function twice.

If you're unfamiliar with operators in quantum mechanics, you might be somewhat disconcerted by equations such as (5.19), which appear to leave the operators 'hanging' without some function to operate on. Please be reassured, the assumption of a suitable function on which to operate is implicit in equations like this. But, as we will see later, we can actually deduce quite a lot of physics just from considerations of the properties of the operators, without ever actually knowing what the wavefunctions are.

Incidentally, equations such as (5.19) have a general form known as an *eigenvalue equation*, such as $\hat{O}f(n, x) = \lambda_n f(n, x)$, where \hat{O} is an operator and the λ_n are the *eigenvalues* which are returned by the operation of \hat{O} on the *eigenfunctions* $f(n, x)$. This is why Schrödinger chose to title the series of papers he published on wave mechanics as 'Quantization as an Eigenvalue Problem'.

Clearly, the solutions to the operator equation (5.19) depend on the nature of the potential energy function V and its variation in space. Students are usually offered a not-so-gentle introduction to wave mechanics by assuming some much-simplified examples of V in one, two, and three dimensions, known as 'particle in a box', 'particle on a ring', and 'particle on a sphere'. The simple one-dimensional harmonic oscillator is also a favourite, and particularly useful as it is applicable to low-energy chemical bond vibrations in diatomic molecules.

Before we move on to Step (3) it's worth noting that in order to reconcile (5.21) with (5.19), we must *define* the linear momentum operator as*

$$\hat{p} = -i\hbar\nabla = -i\hbar(\partial/\partial x + \partial/\partial y + \partial/\partial z) \quad \text{such that} \quad \hat{p}^2 = -\hbar^2\nabla^2 \quad (5.22)$$

(once again, remember that $-i \cdot -i = i^2 = -1$). In (5.22), $\hat{p} = \hat{p}_x + \hat{p}_y + \hat{p}_z$, and $\hat{p}_x = -i\hbar\partial/\partial x$, etc. We'll return to take a closer look at the consequences of this form for the momentum operator in Chapter 7.

Now, astute readers will have noticed that not only has this procedure been quite simple (so far, at least), it has also involved some considerable sleight-of-hand. To obtain Eq. (5.12) we assumed the classical expression for the kinetic energy, $p_x^2/2m$, only to reinterpret the linear momentum as an operator in (5.21) and (5.22), with a form that

* Readers might quibble that I could have chosen to define \hat{p} as $i\hbar\nabla$, as this also returns the right form for \hat{p}^2 . The reason for my choice will become apparent in Chapter 7.

bears absolutely no resemblance to the classical $p = mv$. This is not really acceptable, from either physical or mathematical perspectives, which is almost certainly one of the reasons why Schrödinger chose to present two very different (and much more difficult) derivations in the first of his published papers. His biographer, Walter Moore, called them 'almost deliberately cryptic'.⁴

Step (3): The Time-dependent Wave Equation

Let's now put time back into the picture. The three-dimensional form for our generalized wavefunction is $\Psi(r, t) = \psi(r)\varphi(t)$, or, from Eq. (5.2),*

$$\Psi(r, t) = \psi(r)e^{-i\omega t}. \quad (5.23)$$

We can now use the Planck–Einstein relation $E = h\nu = \hbar\omega$ to rewrite (5.23) as

$$\Psi(r, t) = \psi(r)e^{-iEt/\hbar}. \quad (5.24)$$

If we differentiate (5.24) with respect to time we have

$$\frac{\partial}{\partial t}\Psi(r, t) = -\frac{iE}{\hbar}\psi(r)e^{-iEt/\hbar} = \frac{E}{i\hbar}\Psi(r, t) \quad (5.25)$$

(remember, $-i \cdot \frac{i}{i} = \frac{1}{i}$). Multiplying through by $i\hbar$ gives

$$i\hbar\frac{\partial}{\partial t}\Psi(r, t) = E\Psi(r, t). \quad (5.26)$$

But we can generalize Eq. (5.19) by applying the Hamiltonian operator to the time-dependent wavefunction $\Psi(r, t)$, giving (note that \hat{H} operates only on the spatial part of $\Psi(r, t)$ and the energy E is a simple quantity which can be distributed in the normal way)

$$\hat{H}\Psi(r, t) = \varphi(t)\hat{H}\psi(r) = \varphi(t)E\psi(r) = E\psi(r)\varphi(t) = E\Psi(r, t). \quad (5.27)$$

Combining Eqs. (5.26) and (5.27) allows us to deduce the full time-dependent Schrödinger wave equation as follows:

$$\hat{H}\Psi(r, t) = i\hbar\frac{\partial}{\partial t}\Psi(r, t). \quad (5.28)$$

The first thing to note about this final result is that it is not consistent with the classical wave equation, (5.1), which is a second-order differential in both space *and* time. In

* See? I didn't forget.

arriving at Eq. (5.28), we were obliged to run with a first-order differential in time as this returns an expression which is first-order in the energy E , such that it can be equated with (5.27), which involves a second-order differential in the spatial coordinates but is also first order in E . This means that space and time are not treated equally, and so (5.28) is not consistent with the requirements of special relativity. As written, the time-dependent Schrödinger wave equation is *non-relativistic*.

It's interesting to note that Schrödinger actually started out by deriving a fully relativistic version of the wave equation, but found that the result did not agree with the experimental hydrogen atom spectrum. He went on to discover that the non-relativistic version worked perfectly satisfactorily. He therefore withdrew a paper he had already submitted on the relativistic version and submitted another on the non-relativistic form, which became the first published paper on wave mechanics. Schrödinger's relativistic wave equation was rediscovered by Oskar Klein and Walter Gordon, and is now known as the *Klein–Gordon equation*. We will revisit this, and the reasons for Schrödinger's failure, in Chapter 9.

Application to the Hydrogen Atom

If Schrödinger had stopped at the non-relativistic wave equation, then the impact of his first paper on wave mechanics would have been much less significant. De Broglie had speculated that standing or stationary electron waves confined to an orbit around the nucleus might explain how the hydrogen atom spectrum had come to depend on the quantum numbers n , l , and m_l . Having derived the wave equation, all Schrödinger had to do now was apply it.

The first step involves switching to a different coordinate system. In three dimensions the Coulomb potential around the nucleus is spherically symmetric, and if we assume that the hydrogen atom nucleus is very much heavier than the orbiting electron, then it serves as a suitable origin for a system of spherical coordinates. These are *spherical polar coordinates*, characterized by the radial distance r of the electron from the nucleus, the co-latitude or polar angle θ between the line $0-r$ and the z -axis, and the longitude or azimuthal angle ϕ , measured from the x -axis to the projection of $0-r$ onto the xy plane (see Fig. 5.3). From this we can deduce that $x = r \sin \theta \cos \phi$, $y = r \sin \theta \sin \phi$, and $z = r \cos \theta$. As before, $r^2 = x^2 + y^2 + z^2$.

Changing the variables from x, y, z to r, θ, ϕ transforms the Laplacian operator to the form

$$\nabla^2 = \frac{1}{r} \frac{\partial^2}{\partial r^2} r + \frac{1}{r^2} \Lambda^2, \quad (5.29)$$

where

$$\Lambda^2 = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \quad (5.30)$$

is called the *legendrian*.

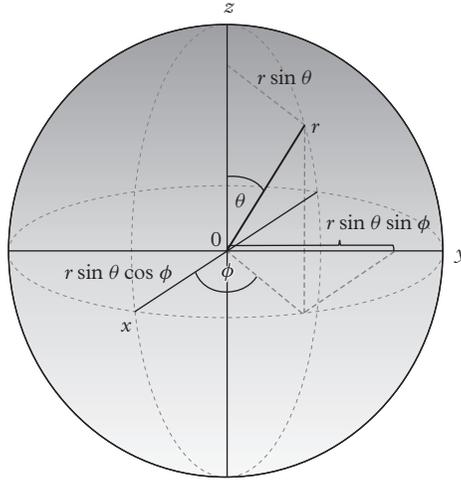


Figure 5.3 The system of spherical polar coordinates.

I don't propose to give the details of this transformation here, but they can be found in most introductory textbooks on quantum mechanics.⁵

We assume that the wavefunctions of the electron in a hydrogen atom can be written as $\Psi(r, \theta, \phi, t) = \psi(r, \theta, \phi) \varphi(t)$, and for now we will confine ourselves to the consideration of the spatial functions only. An expression for the Coulomb potential is available from Eq. (3.8), and we replace m in the wave equation with m_e , the mass of the electron. So, we have everything we need to write the wave equation for the hydrogen atom:

$$-\frac{\hbar^2}{2m_e} \left[\frac{1}{r} \frac{\partial^2}{\partial r^2} r + \frac{1}{r^2} \Lambda^2 \right] \psi(r, \theta, \phi) - \frac{e^2}{4\pi \epsilon_0 r} \psi(r, \theta, \phi) = E \psi(r, \theta, \phi). \quad (5.31)$$

We can hope to simplify this by assuming that the radial variable r can be separated from the angular variables θ and ϕ . If this is the case, then we suppose that $\psi(r, \theta, \phi)$ can be written as the product of a radial function $R(r)$, which depends only on r , and a function $Y(\theta, \phi)$, which depends on the angles θ and ϕ , and is independent of r . In other words, $\psi(r, \theta, \phi) = R(r)Y(\theta, \phi)$, and we note that only the legendrian in (5.31) operates on the angular function $Y(\theta, \phi)$. This allow us to rewrite (5.31) as

$$-\frac{\hbar^2}{2m_e} \frac{Y(\theta, \phi)}{r} \frac{\partial^2}{\partial r^2} r R(r) - \frac{\hbar^2}{2m_e} \frac{R(r)}{r^2} \Lambda^2 Y(\theta, \phi) - \frac{e^2}{4\pi \epsilon_0 r} R(r) Y(\theta, \phi) = ER(r) Y(\theta, \phi). \quad (5.32)$$

I propose that we cheat at this stage (much as Schrödinger would have done in 1925), and recognize that terms of the type $\Lambda^2 Y(\theta, \phi)$ are well known. They form an eigenvalue equation for which the functions $Y(\theta, \phi)$ are eigenfunctions. The eigenvalues of this

equation are also well known and are given by $-l(l+1)$, where l is an integer ranging from 0, 1, 2, ..., such that

$$\Lambda^2 Y(\theta, \phi) = -l(l+1) Y(\theta, \phi). \quad (5.33)$$

Because they depend on two angles, the functions $Y(\theta, \phi)$ are actually characterized by two integer numbers which (given that we know where this is taking us) we will call l and m_l , where m_l ranges in integral steps from $-l$ to $+l$. These are the *spherical harmonics*, a set of functions defined on the surface of a sphere, first investigated in detail by Laplace in 1782. These functions are written as $Y_l^{m_l}(\theta, \phi)$,

$$Y_l^{m_l}(\theta, \phi) = \sqrt{\frac{(2l+1)(l-|m_l|)!}{2(l+|m_l|)!}} P_l^{|m_l|}(\cos \theta) \frac{1}{\sqrt{2\pi}} e^{im_l\phi}, \quad (5.34)$$

where the functions $P_l^{|m_l|}(\cos \theta)$ are the *associated Legendre polynomials*, first studied in the context of Newtonian gravity by French mathematician Adrien-Marie Legendre, and given by

$$P_l^{|m_l|}(\cos \theta) = (1 - \cos^2 \theta)^{|m_l|/2} \frac{d^{|m_l|}}{d(\cos \theta)^{|m_l|}} P_l(\cos \theta), \quad \text{and} \quad (5.35)$$

$$P_l(\cos \theta) = \frac{1}{2^l l!} \frac{d^l}{d(\cos \theta)^l} (\cos^2 \theta - 1)^l. \quad (5.36)$$

The functions $P_l(\cos \theta)$ are the *Legendre polynomials*. The first few examples of the associated Legendre polynomials are given alongside the corresponding spherical harmonics $Y_l^{m_l}(\theta, \phi)$ in Table 5.1.

For now, we just need to note that if we substitute for $\Lambda^2 Y(\theta, \phi)$ from (5.33) in (5.32), then in the resulting expression there are no longer any operators to operate on the angular functions. This means that we can simply cancel them out, leaving a purely radial equation of the form:

$$-\frac{\hbar^2}{2m_e} \frac{1}{r} \frac{\partial^2}{\partial r^2} rR(r) + \left(\frac{-e^2}{4\pi\epsilon_0 r} + \frac{l(l+1)\hbar^2}{2m_e r^2} \right) R(r) = ER(r) \quad (5.37)$$

In this last expression, we can see that the attractive (and negative) Coulomb potential term has now been modified in a way that has a fairly straightforward physical interpretation. For solutions with $l=0$, which we take to be solutions with zero orbital angular momentum, the potential is attractive for all values of r and tends towards $-\infty$ as the electron gets closer and closer to the nucleus. But for solutions with $l > 0$, the orbital angular momentum contributes a kind of centrifugal force which tends to throw the electron outwards, resisting the attraction of the positively charged nucleus. The

Table 5.1 Associated Legendre Polynomials and Spherical Harmonics for Low-Order Values of l and m_l for the Hydrogen Atom

l	m_l	$P_l^{ m_l }(\cos \theta)$	$Y_l^{m_l}(\theta, \phi)$	
0	0	$P_0^0 = 1$	$Y_0^0 = \frac{1}{2\sqrt{\pi}}$	
1	0	$P_1^0 = \cos \theta$	$Y_1^0 = \frac{1}{2}\sqrt{\frac{3}{\pi}} \cos \theta$	
1	± 1	$P_1^1 = \sin \theta$	$Y_1^{\pm 1} = \mp \frac{1}{2}\sqrt{\frac{3}{2\pi}} \sin \theta e^{\pm i\phi}$	
2	0	$P_2^0 = \frac{1}{2}(3\cos^2\theta - 1)$	$Y_2^0 = \frac{1}{4}\sqrt{\frac{5}{\pi}}(3\cos^2\theta - 1)$	
2	± 1	$P_2^1 = 3 \sin \theta \cos \theta$	$Y_2^{\pm 1} = \mp \frac{1}{2}\sqrt{\frac{15}{2\pi}} \cos \theta \sin \theta e^{\pm i\phi}$	
2	± 2	$P_2^2 = 3\sin^2\theta$	$Y_2^{\pm 2} = \frac{1}{4}\sqrt{\frac{15}{2\pi}} \sin^2\theta e^{\pm 2i\phi}$	

corresponding centrifugal potential increases more quickly towards $+\infty$ as r declines (since it depends on $1/r^2$), and so wins out at small distances.

It was here that Schrödinger reached the limit of his mathematical competence. He could solve equation for the spherical harmonics but he struggled to solve the radial equation. He had nevertheless seen enough. On 27 December he wrote a letter to Wien:⁶

At the moment I am struggling with a new atomic theory. If only I knew more mathematics! I am very optimistic about this thing and expect that if I can only ... solve it, it will be *very* beautiful... I hope I can report soon in a little more detailed and understandable way about the matter. At present, I must learn a little more mathematics.

He returned to Zurich on 8 January and immediately sought help from his friend and colleague, the mathematician Herman Weyl. Within about a week he had solved the radial equation, switched to the non-relativistic form of the wave equation (which is what we have used here), and withdrawn his paper on the relativistic version. The resulting expression for the radial function is given in Eq. (5.38). There was little doubt that Schrödinger now possessed the solution to all the mysteries:

$$R_{nl}(r) = \sqrt{\left(\frac{2}{na_0}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}} (\rho_n)^l e^{-\rho_n/2} L_{n-l-1}^{2l+1}(\rho_n). \tag{5.38}$$

The radial functions depend on two integer numbers, n and l —so we now write these as $R_{nl}(r)$ —with the maximum value of l now constrained to $n - 1$. The term $\rho_n = 2r/na_0$, where a_0 is the Bohr radius, $\epsilon_0 h^2 / \pi m_e e^2$, see Eq. (3.18), and the functions $L_{n-l-1}^{2l+1}(\rho_n)$ are the *associated Laguerre polynomials*, named for French mathematician Edmond Laguerre, given by

$$L_{n-l-1}^{2l+1}(\rho_n) = (-1)^{2l+1} \frac{d^{2l+1}}{d\rho_n^{2l+1}} L_{n+l}(\rho_n), \quad \text{and} \tag{5.39}$$

$$L_{n+l}(\rho_n) = e^{\rho_n} \frac{d^{n+l}}{d\rho_n^{n+l}} \left(e^{-\rho_n} \rho_n^{n+l} \right), \tag{5.40}$$

where the functions $L_{n+l}(\rho_n)$ are the *Laguerre polynomials*. The first few associated Laguerre polynomials are given alongside the radial solutions in Table 5.2, and the functional form for these is plotted in Fig. 5.4. Note that in getting the explicit forms for the radial functions from Eq. (5.38) it's helpful to remember that $0! = 1$.

Getting to these solutions was not by any means straightforward, and I should point out that on arriving at Eqs. (5.34) and (5.38) some ‘unphysical’ solutions have been discarded, particularly those which tend to diverge (blow up to infinity) at very small or very large values of r . We also require that the energies associated with the radial functions

Table 5.2 Associated Laguerre Polynomials and Radial Functions for Low-Order Values of n and l for the Hydrogen Atom

n	l	Orbital	$L_{n-l-1}^{2l+1}(\rho_n)$	$R_{nl}(r)$
1	0	1s	$L_0^1 = 1$	$R_{10} = \frac{2}{\sqrt{a_0^3}} e^{-\rho_n/2}$
2	0	2s	$L_1^1 = 4 - 2\rho_n$	$R_{20} = \frac{1}{2\sqrt{2a_0^3}} (2 - \rho_n) e^{-\rho_n/2}$
2	1	2p	$L_0^3 = 6$	$R_{21} = \frac{1}{2\sqrt{6a_0^3}} \rho_n e^{-\rho_n/2}$
3	0	3s	$L_2^1 = 3\rho_n^2 - 18\rho_n + 18$	$R_{30} = \frac{1}{9\sqrt{3a_0^3}} (\rho_n^2 - 6\rho_n + 6) e^{-\rho_n/2}$
3	1	3p	$L_1^3 = 96 - 24\rho_n$	$R_{31} = \frac{1}{9\sqrt{6a_0^3}} (4 - \rho_n) \rho_n e^{-\rho_n/2}$
3	2	3d	$L_0^5 = 120$	$R_{32} = \frac{1}{9\sqrt{30a_0^3}} \rho_n^2 e^{-\rho_n/2}$

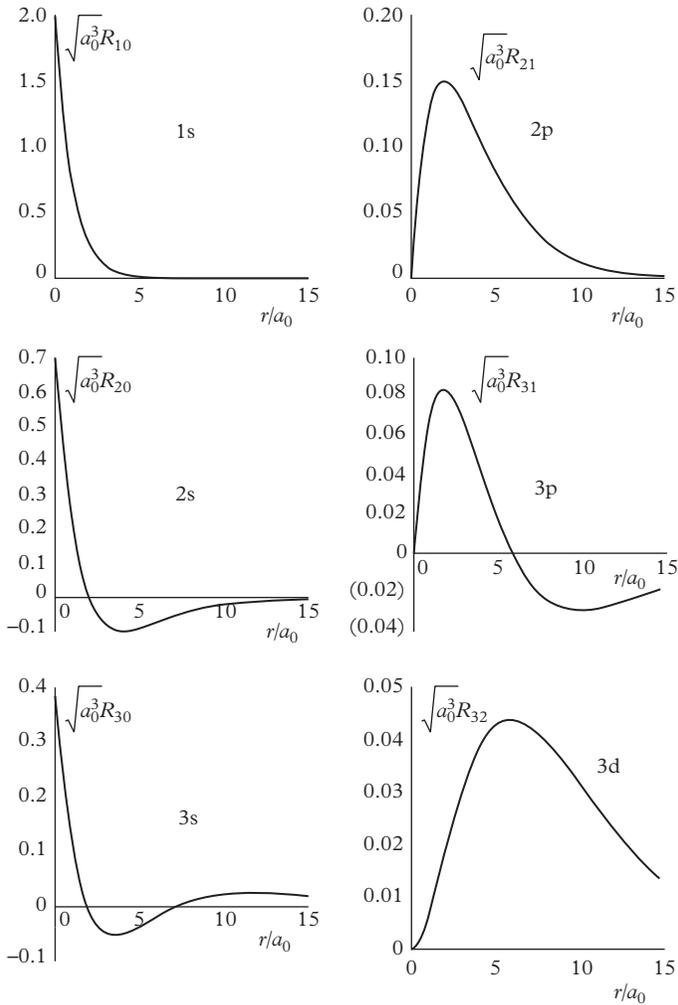


Figure 5.4 The low-energy hydrogen atom radial wavefunctions, plotted as a function of r/a_0 . Note that although the distance scale is the same for each plot, the peak values of the radial functions decline with increasing energy.

have negative values, corresponding to states in which the electron is ‘bound’ to the nucleus, such that E tends to 0 when r tends to ∞ . The wavefunctions themselves are also required to be single-valued, finite, continuous, and *normalized*.⁷ The last is particularly important when we come to interpret what they represent, and we will turn our attention to this in Chapter 6. Imposing such ‘boundary conditions’ on the problem is equivalent to confining the electron in a spherical container whose properties are determined by the Coulomb potential, and gives rise to the quantization manifested through *integer* values

of n , l , and m_l , and their pattern of interrelationships, in just the same way that the pattern of nodes appears in the standing waves of a vibrating string.

We now have everything we need in Tables 5.1 and 5.2 to write down the spatial wavefunctions of the hydrogen atom. Here are just a few:

Orbital		$\psi_{nlm_l}(r, \theta, \phi)$	
1s	$R_{10} Y_0^0$	$\psi_{100} = \frac{1}{\sqrt{\pi a_0^3}} e^{-\rho_n/2}$	(5.41)
2s	$R_{20} Y_0^0$	$\psi_{200} = \frac{1}{4\sqrt{2\pi a_0^3}} (2 - \rho_n) e^{-\rho_n/2}$	(5.42)
2p ₀	$R_{21} Y_1^0$	$\psi_{210} = \frac{1}{4\sqrt{2\pi a_0^3}} \rho_n e^{-\rho_n/2} \cos \theta$	(5.43)
2p _{±1}	$R_{21} Y_1^{\pm 1}$	$\psi_{21\pm 1} = \mp \frac{1}{8\sqrt{\pi a_0^3}} \rho_n e^{-\rho_n/2} \sin \theta e^{\pm i\phi}$	(5.44)

What about the energies? Obviously, we can plug the equation for the radial functions $R_{nl}(r)$ from Eq. (5.38) into the Schrödinger Eq. (5.37) and solve for the energy E (the eigenvalues). This is quite a complex differential equation, but all we need to solve it is some knowledge of algebra, the product rule for differentiation, something called the associated Laguerre equation, and a little patience.

We can learn everything we need to know about the result by looking at the action of the kinetic energy operator (the details are available in Appendix 3):

$$-\frac{\hbar^2}{2m_e} \frac{1}{r} \frac{\partial^2}{\partial r^2} r R_{nl}(r) = -\frac{\hbar^2}{2m_e} \left[\frac{1}{n^2 a_0^2} - \frac{2}{ra_0} + \frac{l(l+1)}{r^2} \right] R_{nl}(r). \quad (5.45)$$

If we insert this into the radial Eq. (5.37) we can cancel the function $R_{nl}(r)$. When we then multiply through by $-\hbar^2/2m_e$ and substitute for the Bohr radius, $a_0 = \epsilon_0 \hbar^2 / \pi m_e e^2 = 4\pi \epsilon_0 \hbar^2 / m_e e^2$, we find that the second term in brackets in (5.45) neatly cancels the Coulomb potential term $-e^2 / 4\pi \epsilon_0 r$, and the third term cancels the centrifugal contribution $l(l+1) \hbar^2 / 2m_e r^2$.

Rather incredibly, what we're left with is

$$E_n = -\frac{\hbar^2}{2m_e a_0^2} \cdot \frac{1}{n^2} = -\frac{m_e e^4}{8\epsilon_0^2 \hbar^2} \cdot \frac{1}{n^2}, \quad (5.46)$$

which is precisely the result that Bohr had obtained in 1913 with his ad hoc planetary model, Eq. (3.17). This really is nothing less than a revelation. The energies of each orbital are indeed 'coded' in the mathematical forms of the radial wavefunctions that

describe them. All we need to do to tease them out is work on them with the right Hamiltonian operator.

Given that the radial function depends on both n and l , it is perhaps surprising that the energy depends only on n . And I think it's particularly interesting that the kinetic energy term involving l in Eq. (5.45) is derived from the radial function, yet the potential energy term that it cancels is derived from the spherical harmonics. This turns out to be a peculiarity of the Coulomb potential—other spherical potentials may give quantized energies that do depend on both n and l .

It is also now clear why Bohr's Second Assumption, that the quantized orbital angular momentum depends exclusively on n , is wrong—see Eq. (3.23). We will see in later chapters that the quantized orbital angular momentum is given by $\sqrt{l(l+1)}\hbar$, not $n\hbar$. Not for the first (or the last) time in science, Bohr got the right answer for the wrong reasons.

The Problem of Interpretation

According to Debye, de Broglie's hypothesis had been, if not a fanciful French comedy, still all rather childish. But there was nothing wrong with de Broglie's intuition. In an extraordinary burst of creativity spanning just a few short weeks, Schrödinger had shown that the quantum numerology inspired by the Bohr–Sommerfeld model did indeed have a comprehensible physical basis. By assuming that the electron in a hydrogen atom behaves as a wave, he was able to show that the quantum numbers n , l , and m_l emerge 'in the same natural way as the integers specifying the number of nodes in a vibrating string'.⁸

There was more. Schrödinger had successfully imported an entirely new mathematical structure and approach to mechanics, complete with its own rather strange, esoteric language. The textbooks of later generations of science students would be filled with operators, eigenfunctions, and eigenvalues.

It was indeed *very* beautiful. But physicists were now faced with another problem. The wavefunctions for the electron in a hydrogen atom could now be written down and their energies computed. But just how were these wavefunctions meant to be interpreted?

Think of it this way. In the 1s orbital described by the radial function R_{10} , Fig. 5.4, and the spherical harmonic Y_0^0 , Table 5.1, *just where, exactly, is the electron?*

NOTES

1. Abraham Pais, *Niels Bohr's Times, in Physics, Philosophy and Polity*, Oxford University Press, Oxford, 1991, p. 177.
2. R. Kronig, in M. Fierz and V. F. Weisskopf (eds), *Theoretical Physics in the Twentieth Century: A Memorial Volume to Wolfgang Pauli*, Interscience, New York, 1960, p. 50. Quoted in Pais, *Niels Bohr's Times*, p. 188.

3. Felix Bloch, ‘Heisenberg and the Early Days of Quantum Mechanics’, *Physics Today*, **29** (1976), 23.
4. Walter Moore, *Schrödinger: Life and Thought*, Cambridge University Press, 1989, Cambridge, UK, p. 200.
5. My personal favourites are Peter Atkins and Ronald Friedmann’s *Molecular Quantum Mechanics*, David J. Griffiths’ *Introduction to Quantum Mechanics*, and A. P. French and E. F. Taylor’s *An Introduction to Quantum Physics* (see this chapter’s Further Reading for more details). Different textbooks are usually pretty consistent in the way that the legendrian and the spherical harmonics are written, but sometimes differ in the form of the leading kinetic energy term in Eq. (5.41). For the purposes of comparison, it might be helpful to note that

$$\frac{1}{r} \frac{\partial^2}{\partial r^2} r R(r) = \left(\frac{2}{r} \frac{\partial}{\partial r} + \frac{\partial^2}{\partial r^2} \right) R(r) = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} R(r) \right).$$

To move back and forth between these expressions, simply apply the product rule: $\partial(uv)/\partial r = v\partial u/\partial r + u\partial v/\partial r$.

6. Erwin Schrödinger, letter to Wilhelm Wein, 27 December 1925, quoted in Moore, *Schrödinger*, p. 196.
7. English translations of Schrödinger’s papers ‘Quantization as an Eigenvalue Problem’, Parts I–IV, are reproduced in Chapter 3 of Stephen Hawking (ed.), *The Dreams that Stuff Is Made Of: The Most Astounding Papers on Quantum Physics and How They Shook the World*, Running Press, Philadelphia, 2011, pp. 251–387. These are worth consulting, but be warned: in his first paper Schrödinger denotes the principal quantum number as l and the azimuthal quantum number as n , which can lead to some confusion!
8. Erwin Schrödinger, ‘Quantisierung als Eigenwertproblem’, *Annalen der Physik*, **79** (1926), 735. English translation quoted in Moore, *Schrödinger*, p. 202.

.....

FURTHER READING

- Atkins, Peter, and Friedman, Ronald, *Molecular Quantum Mechanics*, 5th edn, Oxford University Press, Oxford, 2011, Chapter 3.
- French, A. P., and Taylor, E. F., *An Introduction to Quantum Physics*, Van Nostrand Reinhold, Wokingham, 1978, Chapter 12.
- Gamow, George, *Thirty Years That Shook Physics*, Dover, New York, 1966, Chapter IV.
- Griffiths, David J., *Introduction to Quantum Mechanics*, 2nd edn, Cambridge University Press, Cambridge, UK, 2017, Chapter 4.
- Hawking, Stephen (ed.), *The Dreams That Stuff Is Made Of: The Most Astounding Papers on Quantum Physics and How They Shook the World*, Running Press, Philadelphia, 2011, pp. 251–66.
- Kilmister, C. W. (ed.), *Schrödinger: Centenary Celebration of a Polymath*, Cambridge University Press, Cambridge, UK, 1987, particularly Chapters 3–6.

- Kragh, Helge, *Niels Bohr and the Quantum Atom: The Bohr Model of Atomic Structure 1913–1925*, Oxford University Press, Oxford, 2012, Chapter 4.
- Miller, Arthur I., 'Erotica, Aesthetics and Schrödinger's Wave Equation', in Graham Farmelo (ed.), *It Must Be Beautiful: Great Equations of Modern Science*, Granta, London, 2002, pp. 110–31.
- Moore, Walter, *Schrödinger: Life and Thought*, Cambridge University Press, Cambridge, UK, 1989, Chapter 6.
- Pais, Abraham, *Niels Bohr's Times, in Physics, Philosophy and Polity*, Oxford University Press, Oxford, 1991, pp. 179–99.

6

Born's Interpretation of the Wavefunction

Quantum Probability

The sleight of hand involved in Schrödinger's 'trivial' derivation of the wave equation based on de Broglie's hypothesis had once more allowed quantum conditions to spring from entirely classical concepts. And yet, for reasons that might seem curious to us today, Schrödinger perceived the wave equation not as a route forward into a new discontinuous quantum mechanics, but as a route *back* to a broadly classical description of atomic phenomena based on the notion of continuous matter waves.

According to Schrödinger's interpretation of wave mechanics, elementary particles such as electrons are principally *undulatory* in nature, and their particle-like properties are the illusory result of the collective motion of a superposition of electron waves, known as a 'wave packet'. He argued that particle trajectories arise in a wave theory of matter in much the same way that light rays arise in the wave theory of light:¹

According to the wave theory of light, the light rays, strictly speaking, have only fictitious significance. They are not the physical paths of some particles of light, but are a mathematical device, the so-called orthogonal trajectories of wave surfaces, imaginary guide lines as it were, which point in the direction normal to the wave surface in which the latter advances.

For Schrödinger, it was no coincidence that the wavelength of an electron is comparable with the dimensions of the atom: '*we assert that the atom in reality is merely the diffraction phenomenon of an electron wave captured as it were by the nucleus of the atom*'.²

It follows that there is no room in Schrödinger's continuous wave mechanics for Bohr's incomprehensible quantum jumps. Schrödinger imagined that the differences in the energies associated with the different 'orbits' of the Bohr-Sommerfeld model arise as 'beats' between high-frequency vibrational modes of the electron, and that transitions take place gradually and continuously, not discontinuously in sudden jumps. In his first paper in the series 'Quantisation as an Eigenvalue Problem' he wrote: 'It is hardly necessary to emphasise how much more congenial it would be to imagine that at a

quantum transition the energy changes over from one form of vibration to another, than to think of a jumping electron'.³

Schrödinger had hoped that his wave mechanics would help to re-establish some sense of what he referred to as *anschaulichkeit* (meaning 'visualizability') of the physics going on inside the atom. But he also recognized that this wouldn't be straightforward. The wavefunctions can't possibly represent easily visualizable, real 'waves'. For one thing, although all the 'observables' (such as energy and linear momentum) *must* be exclusively real quantities, kind of by definition, in general the wavefunctions themselves may possess complex phases (involving terms such as $e^{i\theta}$). Schrödinger had tended to gloss over this problem, believing that it would be possible always to take only the real part of the solution, or that the problem would be resolved in a fully relativistic wave equation.

There were also other conceptual difficulties. We can write wavefunctions for the single electron in a hydrogen atom in three spatial coordinates, such as r , θ , and ϕ . But that's it. For atoms containing two or more electrons the solutions are necessarily wavefunctions in a complex multidimensional *configuration space* (see the discussion in 'Classical Mechanics and the Concept of Force', in the Prologue), which is much more difficult to visualize.

So, if the wavefunction does not provide a simple description of an 'electron wave', as we might have naively imagined, what, then, does it represent?

$|\psi|^2$ as a Charge Density

Schrödinger devoted the final section of his fourth paper in the series to a discussion of the physical significance of the wavefunction (which he referred to in this paper as the 'field scalar'). Particles such as electrons possess mass but they also possess electrical charge and, whilst it might be difficult to conceive how the electron mass might be 'distributed' through space as a wave, charge is arguably a much more nebulous (and therefore 'flexible') property. As such, it is easier to imagine a *distribution of charge*.

In his treatment of the wave-mechanical theory of dispersion, Schrödinger found it convenient to focus on the *density* of electrical charge. He associated this density not with the wavefunction directly, but with the wavefunction ψ multiplied by its complex conjugate, which is written ψ^* , and in which every instance of the imaginary number $i = \sqrt{-1}$ is replaced by $-i$. For example, if $\psi = Ae^{i\theta}$, then its complex conjugate is $\psi^* = A^*e^{-i\theta}$, and $\psi^*\psi = |\psi|^2 = |A|^2$. There are obvious logical parallels with the relation between the intensity of a light wave or the energy of a physical wave and the square of the wave amplitude. Schrödinger wrote:⁴

[$\psi^*\psi$] is a kind of *weight function* in the system's configuration space...each point-mechanical configuration contributes to the true wave-mechanical configuration with a certain *weight*, which is given precisely by [$\psi^*\psi$]. If we like paradoxes, we may say that the system exists, as it were, simultaneously in all the positions kinematically imaginable, but not 'equally strongly' in all.

Although he would later come to reject it, hidden in Schrödinger's words is the interpretation that would eventually come to dominate our understanding of the wavefunction.

The condition of normalization, which had featured in Schrödinger's reasoning from the beginning, is then a requirement that 'we would wish its integral over the whole configuration space to remain constantly normalized to the same unchanging value, preferably to unity'.⁵ For the simple reason that the charge on an electron is a fixed quantity—a 'unit' of negative charge—and such a constraint is therefore necessary if this charge is to be conserved.

It will help in what follows to take a quick look at this condition here.

Normalization and Orthogonality

The normalization condition can be written as

$$\int \psi_n^* \psi_n d\tau = 1, \quad (6.1)$$

where ψ_n is a general wavefunction characterized by one or more quantum numbers collectively denoted by n , and $d\tau$ is an infinitesimal volume element in whatever coordinates in real space or configuration space are relevant to the problem under consideration. Although in Eq. (6.1) the integral is written as indefinite, the presumption is that it is definite and runs over *all* possible values of the coordinates.

To see how this works it's helpful to return to the wavefunction of the electron in a hydrogen atom given in Chapter 5, and to keep things simple we'll consider the spherical 1s function ψ_{100} , Eq. (5.41). In the context of the spherical polar coordinates we used to analyse the hydrogen atom, the volume element $d\tau$, located a distance r from the origin, can be calculated as the product of the increments in the radial distance, dr , the polar angle multiplied by r , $r d\theta$, and the azimuthal angle multiplied by $r \sin \theta$, or $r \sin \theta d\phi$ (see Fig. 6.1). Thus, $d\tau = r^2 \sin \theta dr d\theta d\phi$ ⁶ and as ψ_{100} is independent of the angles, the integral to be evaluated simplifies to

$$\int \psi_{100}^2 d\tau = \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi \int_0^\infty \frac{1}{\pi a_0^3} r^2 e^{-2r/a_0} dr. \quad (6.2)$$

Note that in (6.2) we have acknowledged that ψ_{100} is not a complex function, so we can happily replace $|\psi_{100}|^2$ with ψ_{100}^2 . We have also substituted for $\rho_n = 2r/na_0$, using $n = 1$. The polar angle ranges over $0-\pi$ (from north to south) and the azimuthal angle ranges over $0-2\pi$ (all the way around the equator). As the 1s wavefunction is spherically symmetric it doesn't depend on either of these angles and so we can immediately evaluate the first two integrals to give $2 \times 2\pi = 4\pi$. Equation (6.2) then becomes

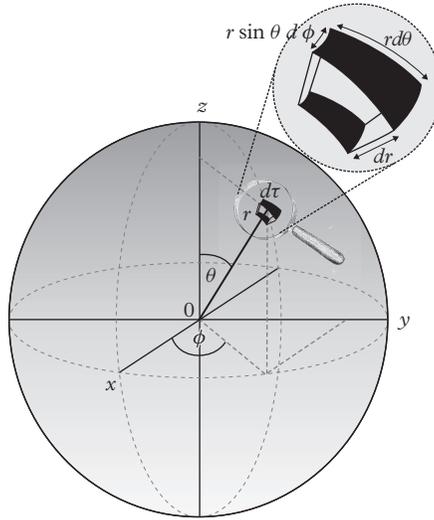


Figure 6.1 The volume element $d\tau$ in spherical polar coordinates is given by the product of its dimensions: $dr \times rd\theta \times r \sin \theta d\phi = r^2 \sin \theta dr d\theta d\phi$.

$$\int \psi_{100}^2 d\tau = \frac{4}{a_0^3} \int_0^\infty r^2 e^{-2r/a_0} dr. \tag{6.3}$$

Now, we can evaluate this integral by parts or we can do what I always prefer to do in these circumstances, which is look it up in a table of standard integrals. The integral we need is

$$\int x^2 e^{bx} dx = \left(\frac{x^2}{b} - \frac{2x}{b^2} + \frac{2}{b^3} \right) e^{bx}. \tag{6.4}$$

We substitute $x = r$ and $b = -2/a_0$. To make this definite, we need to evaluate (6.4) at the limits $r = \infty$ and $r = 0$. When $r = \infty$ the exponential term in (6.4) declines to 0, taking everything else down to 0 with it. When $r = 0$ the exponential term is equal to 1 and the only term to survive is $2/b^3 = -a_0^3/4$. Thus, the integral in (6.3) comes out to $+a_0^3/4$ and

$$\int \psi_{100}^2 d\tau = \frac{4}{a_0^3} \cdot \frac{a_0^3}{4} = 1, \tag{6.5}$$

which simply proves that the wavefunction ψ_{100} as given in (5.41) is already normalized.

To complete the picture, it's worth noting that solutions of the wave equation that do not correspond to the same energy eigenvalues have a further property—they are *orthogonal*, such that

$$\int \psi_m^* \psi_n d\tau = 0. \quad (6.6)$$

Here ψ_m and ψ_n represent two different solutions of the wave equation with different energies (say, E_m and E_n). Once again, we can make use of the simpler hydrogen atom wavefunctions—such as ψ_{100} and ψ_{200} —to show that this is indeed the case (see Appendix 4).

We can combine Eqs. (6.1) and (6.6) in a single expression

$$\int \psi_m^* \psi_n d\tau = \delta_{mn}, \quad (6.7)$$

where δ_{mn} is the *Kronecker delta*, which has the value 1 when $m = n$ and 0 when $m \neq n$. Solutions that are at once both normalized and orthogonal are said to be *orthonormal*.

Matrix Mechanics

Schrödinger's wave mechanics was undoubtedly a triumph, and the fact that it still features prominently in undergraduate textbooks on quantum mechanics demonstrates both its historical importance and utility. But, in truth, by the time Schrödinger submitted his first paper on the subject in January 1926, a successful rival theory was already available, published some months earlier by a young German theorist called Werner Heisenberg. This was *matrix mechanics*, and physicists were now confronted with a choice.

The two theories could not appear more different, at least in terms of the philosophical perspectives that had motivated them. Where Schrödinger had been attracted by the prospect of visualizability in wave mechanics, Heisenberg had become convinced that progress in atomic theory could only be made by abandoning any attempt to visualize or understand the interior workings of the atom. The planetary model, with its appealing images of material particles orbiting a central nucleus, was visually rich but analytically empty. As a classical mechanical model, its usefulness had been prolonged by the addition of rather arbitrary quantum rules, but it was surely doomed to failure. Heisenberg reasoned that it was time for a new language.

Towards the end of May 1925 Heisenberg had succumbed to a severe bout of hay fever. He asked his research supervisor Max Born at the University of Göttingen for 14 days' leave to recuperate, and on 7 June he arrived on the small island of Helgoland, just off the north coast of Germany, hoping that the clear North Sea air would facilitate a speedy recovery.

He had been working on an approach to atomic theory in which he dispensed completely with any notion of an unobservable interior mechanics of the atom, and focused instead on the things that were observable—the jumps or transitions between orbits manifested as lines in the resulting spectra. Free from distractions, he now made swift progress. What he got was an infinite table of mathematical terms, organized into

columns and rows, each term representing a quantum transition from some initial to some final state.

He then sought to calculate the intensities of the spectral lines as the squares of the amplitudes of the terms that appeared in the table. For example, in the case of a jump from a state characterized by a quantum number n to $n - 2$, he found it necessary to multiply the amplitude of the term corresponding to the jump n to $n - 1$ with the amplitude of the term corresponding to the jump from $n - 1$ to $n - 2$. More generally, he found that he could calculate the intensity of a spectral line resulting from any quantum jump as the sum of the products of the amplitudes for all possible intermediate jumps.

Whilst this multiplication rule seemed straightforward and perfectly satisfactory, Heisenberg was aware of a potential paradox. If this same rule is used to calculate the product of two different physical quantities (x and y , say), then it suggests that there may arise situations in which the product x multiplied by y is not equal to the product y multiplied by x . Heisenberg was quite unfamiliar with this kind of result and greatly unsettled by it.

Whilst recuperating on Helgoland, he worked long into the night.⁷

As a result, it was almost three o'clock in the morning before the final result of my computations lay before me. . . . At first, I was deeply alarmed. I had the feeling that, through the surface of atomic phenomena, I was looking at a strangely beautiful interior, and felt almost giddy at the thought that I now had to probe this wealth of mathematical structures nature had so generously spread out before me.

Heisenberg hastily wrote up his calculations and passed a copy of the manuscript to Born. Born was enthusiastic but puzzled by the multiplication rule that Heisenberg had used. It seemed familiar. On 10 July 1925 he finally remembered that he had been taught this multiplication rule as a student. It was the rule for multiplying *matrices*.^{*} Born now worked with another student Pascual Jordan to recast Heisenberg's calculations into the language of matrix multiplication. They discovered that the matrix for the energy of the system is diagonal, meaning that all the elements in the matrix are zero but for those along the diagonal of the array, which are time-independent and represent the stable quantum states (the 'orbits') of the system.

Heisenberg was at once both pleased and relieved: 'Only much later did I learn from Born that it was simply a matter here of multiplying matrices, a branch of mathematics that had hitherto remained unknown to me.'⁸ He acquired some textbooks on the subject and quickly caught up. He was soon collaborating with Born and Jordan and together they published a further paper on the new matrix mechanics in November 1925.

In a paper received by the German Journal *Zeitschrift für Physik* on 17 January 1926, Austrian theorist Wolfgang Pauli used the new matrix mechanics to derive the Rydberg formula for the hydrogen atom, and English theorist Paul Dirac did much the same in a paper received by the *Proceedings of the Royal Society* just a few days later.⁹ There could be no doubting that matrix mechanics *worked*.

^{*} If you're unfamiliar with the mathematics of matrices don't worry—we'll return to this subject in a later chapter.

To underscore the nature of the choice, Schrödinger himself demonstrated that, mathematically, wave mechanics and matrix mechanics are entirely equivalent. So, this was not a choice driven by the theory's 'correctness'. Judgements were instead shaped by the appeal to the physicists' metaphysical preconceptions, their sense of how nature might be or ought to appear. Heisenberg was forthright. In a letter to Pauli he wrote: "The more I think about the physical portion of the Schrödinger theory, the more repulsive I find it. . . . What Schrödinger writes about the visualizability of his theory "is [paraphrasing Bohr] probably not quite right", in other words, it's crap."¹⁰

But matrix mechanics lacked any such visualizability—by design—and, to an older generation of physicists schooled in classical ways of thinking, it was all rather abstract.

It also appeared to be quite limited. Although it had been shown to be equivalent to wave mechanics, this was a theory about transitions taking place between energy levels *within* the atom, and therefore good only for predicting the positions and intensities of spectral lines. But any new quantum mechanics really ought to be applicable to other problems in physics, such as for example the description of the *collision* between an external electron and an atom, a phenomenon generally called 'scattering'. When Born sat down to try to formulate a quantum scattering theory he found that: 'Of the different forms of the theory only Schrödinger's has proved suitable'.¹¹

Scattering phenomena would appear to be firmly 'particle' in nature, so Born's challenge was to find a way to conjure particles from Schrödinger's determinedly wave description. In order fully to appreciate Born's logic, it's helpful first to become a little more familiar with some of the results of classical scattering theory.

Particle Scattering

To keep things reasonably simple, we're going to examine a hypothetical situation in which a small, hard, impenetrable spherical particle is bounced off a much larger, hard, impenetrable spherical target, of radius R . It will help in what follows to make use of the spherical polar coordinate system we introduced in Chapter 5 (see Fig. 5.3), but now turned on its side, such that the positive z -direction (south to north) is drawn horizontally from left to right on the page. Drawn this way, the plane of the paper is still the yz -plane, with the x -axis pointing out of this plane towards and away from the reader.

We imagine that the moving particle approaches the target with velocity v along the z -direction, but with a trajectory that is displaced from the target equator by a distance b (called the impact parameter), such that $b > 0$ implies a 'glancing' collision and $b = 0$ a 'head-on' collision—see Fig. 6.2(a). The particle strikes the surface of the target at an incident angle α to the surface normal, and we assume that the target is so large that it doesn't recoil or in any other way exchange energy or momentum with the incoming particle. Conservation of kinetic energy means that the moving particle bounces off the target with the same outgoing velocity, and conservation of angular momentum means that it bounces off at the same angle as the incident angle. The collision takes place all in the same yz -plane (it doesn't bounce off into or out of the plane of the paper). This

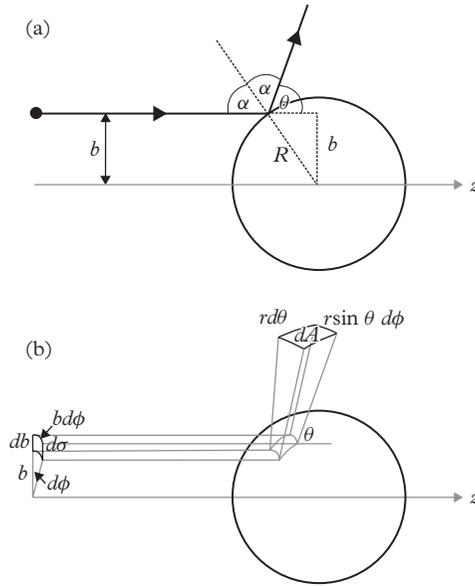


Figure 6.2 (a) Elastic scattering of an incident particle from a large, hard, target sphere. (b) Relationships between the incident beam cross-section $d\sigma$ and detector area dA .

is *elastic scattering*, with a scattering angle θ which is equivalent to the polar angle of our spherical coordinate system.

Simple geometry allows us to deduce that $b = R \sin \alpha$, and $\theta = \pi - 2\alpha$, so that

$$b = R \sin \left(\frac{\pi - \theta}{2} \right) = R \cos \left(\frac{1}{2} \theta \right). \tag{6.8}$$

This is obviously valid only for $b \leq R$, as $b > R$ means that the particle misses the target and there is no collision.

In a realistic physical system, we are likely to be observing the results of a series of collisions involving a *beam* consisting of a large number of identical incident particles. The impact parameters of successive collisions will not be controlled and so will vary, and we're interested in the number or rate of flow of elastically scattered particles that enter a detector placed at different scattering angles.

If we extend the impact parameter by a small increment, to $b + db$, and the azimuthal angle in the xy -plane by a small increment $d\phi$, then we can define a *cross-section* of the beam $d\sigma$ as the area bounded by $|db|$ and $b d\phi$, i.e. $d\sigma = b |db| d\phi$; see the left-hand side of Fig. 6.2(b).

Incident particles pass through the 'window' defined by $d\sigma$ with a certain *flux*, or rate of flow per unit area. Particle physicists refer to the number of incident particles per unit

time per unit area as the beam *luminosity*, \mathcal{L} . The number dN of incident particles passing through $d\sigma$ per unit time is then simply $dN = \mathcal{L}d\sigma$.

The particles are scattered into a detector area dA , which is placed a long distance r from the target object and centred on it. The area dA is then an element on the surface of a sphere of radius r , such that $dA = rd\theta \times r \sin\theta d\phi = r^2 \sin\theta d\theta d\phi$ —see Fig. 6.2(b). It's convenient to define the *solid angle* $d\Omega$, measured in steradians and subtended at the origin of the sphere of radius r by the surface area dA , such that $dA = r^2 d\Omega$, and $d\Omega = \sin\theta d\theta d\phi$. This enables us to define the *differential scattering cross-section*:

$$\frac{d\sigma}{d\Omega} = \frac{b |db| d\phi}{\sin\theta d\theta d\phi} = \frac{b}{\sin\theta} \left| \frac{db}{d\theta} \right|. \quad (6.9)$$

The number of particles dN scattered into the detection solid angle $d\Omega$ per unit time must obviously be the same as the number of particles crossing $d\sigma$ per unit time, so

$$\frac{d\sigma}{d\Omega} = \frac{dN}{\mathcal{L}d\Omega}. \quad (6.10)$$

As quantities such as beam luminosity, detection solid angles, and particle detection rates are relatively easily measured, Eq. (6.10) is often used to *define* the differential scattering cross-section.

From (6.8) and (6.9), we have

$$\frac{d\sigma}{d\Omega} = \frac{R \cos(\frac{1}{2}\theta)}{\sin\theta} \frac{d}{d\theta} \left[R \cos\left(\frac{1}{2}\theta\right) \right] = \frac{R^2 \cos(\frac{1}{2}\theta) \sin(\frac{1}{2}\theta)}{\sin\theta \cdot 2}. \quad (6.11)$$

Or

$$\frac{d\sigma}{d\Omega} = \frac{1}{4}R^2, \quad (6.12)$$

where we have used the trigonometric identity $\sin\theta = 2 \sin(\frac{1}{2}\theta) \cos(\frac{1}{2}\theta)$.

It's rather interesting that the differential scattering cross-section in Eq. (6.11) doesn't depend on the scattering angle, θ . For the case of elastic scattering from a large spherical target, the scattering is *isotropic*—the same number of particles will be scattered into the detector wherever it is placed around the target.

The *total cross-section* σ is given by the integral

$$\sigma = \int \frac{d\sigma}{d\Omega} d\Omega = \frac{1}{4}R^2 \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi = \pi R^2, \quad (6.13)$$

which is simply the cross-sectional area of the target sphere. Particles passing through this area will collide with the target and be scattered. Those outside this area will miss it.

This is probably enough background for us to turn now to the logic and reasoning that Born set out in a paper published a little later in 1926. This did not involve a mathematical derivation, as such, but rather an argument leading to a logical conclusion. Nevertheless, it's useful to bring out this logic in the derivation of some quantum equivalents to the classical scattering parameters we have considered above, such as the differential scattering cross-section, and this is what I propose to do here.

The Ingredients

1. Wave descriptions for systems of linear and spherical plane waves.
2. The de Broglie relation, $\lambda = h/p$, or $p = \hbar k$, where k is the wave vector.
3. The radial Schrödinger wave equation for a localized spherical potential.
4. Legendre polynomials and spherical Hankel functions of the first kind.
5. The differential scattering cross-section, Eq. (6.10).

The Recipe

We begin in Step (1) by setting up the scattering problem in terms of classical waves instead of particles, and applying the de Broglie relation to derive a radial wave equation for the system. We analyse the problem in terms of a system of so-called partial waves, expressing the scattered spherical wave as an infinite sum of partial waves comprised of Legendre polynomials and spherical Hankel functions. This allows us to define a *scattering amplitude* for the outgoing wave. In Step (3) we show that the differential scattering cross section is equal to the modulus-square of the scattering amplitude. We now have to figure out how we get from an *amplitude* in a wave description to a *number* in a particle description. We conclude that the modulus-square of the scattering amplitude (and, by inference, the modulus-square of the wavefunction itself) must represent a *probability density*.

Step (1): The Radial Wave Equation

Classical scattering theory is based on the notion that we scatter small, spherical particles from a target (or target potential). In Schrödinger's wave mechanics, particles give way to waves, so instead of describing the electron as a localized particle we replace this with a simple linear 'electron wave'.

The electron wave impinges on a target represented by a *localized* spherical potential, which we denote as V_{scat} . By 'localized', what we mean is that the potential energy of the target is confined within a small boundary, such that V_{scat} is large within this boundary but can safely be assumed to be negligible or 0 outside it. This means that a Coulomb potential is excluded from this kind of treatment, as its influence declines to 0 only at

infinity. To maintain consistency with our discussion of classical scattering, we assume that the electron wave is travelling in the positive z -direction—see Fig. 6.3. As before, a plane wave such as $\Psi(z, t) = Ae^{i(kz - \omega t)}$ will suffice. We separate this into spatial- and time-dependent functions $\psi(z) = Ae^{ikz}$ and $\varphi(t) = e^{-i\omega t}$.

When the incoming plane wave encounters the scattering potential, it produces a system of outgoing spherical waves (Fig. 6.3). This phenomenon should be familiar to anyone who has watched as ocean waves run into the pillar of a seaside pier. The incoming wave ‘wraps around’ the pillar and it becomes a source of secondary waves which expand outwards in a circle.

We assume that this is still elastic scattering, so the energy of the incoming wave is preserved in the outgoing spherical wave, as is the amplitude. If we think of the incoming wave as a ‘beam’ of electron waves, then if we keep this beam switched on for times longer than it takes to scatter, we anticipate that the system will settle down into a steady-state. This means that we can neglect the time-dependent components of both incoming and outgoing waves and consider only the spatial functions.

When the electron wave is far from the scattering potential, $V_{scat} = 0$ and its total energy is therefore equal to its kinetic energy. If we assume the classical expression for linear momentum, $p = m_e v$, where v is the wave velocity, then $E = p^2/2m_e$, or $p = \sqrt{2m_e E}$. We now make this description decidedly non-classical by invoking the de Broglie relation, $p = \hbar k$, which allows us to write the wave vector k as

$$k = \frac{p}{\hbar} = \frac{\sqrt{2m_e E}}{\hbar}. \quad (6.14)$$

To describe the outgoing waves we adopt spherical polar coordinates, as before, with the z -axis as drawn in Fig. 6.3 and the angles θ and ϕ as already defined in our earlier analysis of classical scattering. We measure the distance r from the centre of the scattering potential and write the outgoing waves as $\psi(r, \theta, \phi)$. We note that these waves will possess the same energy E (and hence wave vector k) and amplitude, A , as the incoming wave.

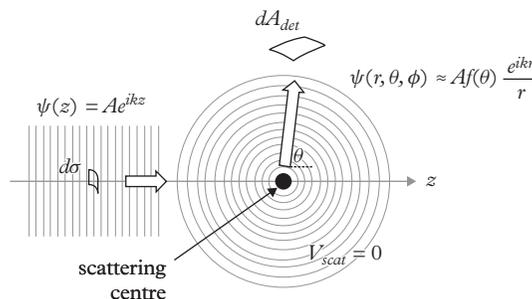


Figure 6.3 Scattering of an electron wave from a localized spherical potential. The potential energy V_{scat} is assumed to be concentrated in a small central region, such that $V_{scat} = 0$ outside this.

As with our consideration of the wavefunctions of the hydrogen atom in Chapter 5, we further assume that the outgoing waves can be factored into radial and angular components, $\psi(r, \theta, \phi) = R(r)Y(\theta, \phi)$.

From Eq. (5.37) we can now set up the radial wave equation as follows:

$$-\frac{\hbar^2}{2m_e} \frac{1}{r} \frac{\partial^2}{\partial r^2} rR(r) + \left(V_{scat} + \frac{l(l+1)\hbar^2}{2m_e r^2} \right) R(r) = ER(r), \quad (6.15)$$

where we have simply replaced the Coulomb potential in (5.37) with the localized scattering potential, V_{scat} .

Although this looks just like the hydrogen atom problem, we're dealing here with a very different potential function and very different boundary conditions. For one thing, we require the incoming plane wave to be one solution of Eq. (6.15) at large distances far from the scattering centre. Unlike the hydrogen atom problem, the energy of the system under consideration is fixed—we're not looking to determine energy eigenvalues from (6.15). There is also no constraint on the value of l , which is not a quantum number as such in this analysis but which is best thought of as an index which characterizes the angular parts of the outgoing spherical waves.

We now simplify (6.15) by considering regions well outside the scattering potential, where $V_{scat} = 0$. If we multiply through by $-2m_e r/\hbar^2$ we get

$$\frac{\partial^2}{\partial r^2} rR(r) - \frac{l(l+1)}{r^2} rR(r) = -\frac{2m_e E}{\hbar^2} rR(r). \quad (6.16)$$

But, from (6.14), $2m_e E/\hbar^2 = k^2$. Substituting $\chi(r) = rR(r)$ allows us to write (6.16) rather succinctly as

$$\frac{\partial^2}{\partial r^2} \chi(r) - \frac{l(l+1)}{r^2} \chi(r) = -k^2 \chi(r). \quad (6.17)$$

As before, I don't propose to solve this equation here and will instead reach for handy solutions that have already been worked out.

Step (2): Partial Waves

The 'standard' solutions for the wavefunctions $\chi(r)$ in Eq. (6.17) are linear combinations of spherical Bessel functions, but these do not give us quite what we're looking for. We will instead reach for spherical *Hankel* functions, which can be derived from linear combinations of the Bessel functions and which, as we will see, more closely fulfil the physical requirements of the problem. There is nothing wrong in principle with this approach: one of the sources of great flexibility in any kind of problem expressed in terms of waves is that waves can be combined any way we like, so long as we respect the physical constraints of the system (such as boundary conditions and normalization).

In fact, because we need solutions describing *outgoing* spherical waves, we choose spherical Hankel functions of the first kind, denoted $h_l^{(1)}(x)$, the first few of which are given in Table 6.1 for the case $x = kr$. We also note that there is an infinite number of such solutions, each corresponding to different values of l , which ranges in integer steps from 0 to infinity. What we get then is a linear combination of these:

$$\chi(r) = c_0 r h_0^{(1)}(kr) + c_1 r h_1^{(1)}(kr) + \dots = \sum_{l=0}^{\infty} c_l r h_l^{(1)}(kr). \quad (6.18)$$

In (6.18), the ‘mixing’ or ‘expansion’ coefficients c_l represent the contribution of each individual solution to the total wavefunction. Since $R(r) = \chi(r)/r$, to get the radial solutions we simply divide (6.18) through by r .

Now, to get the total outgoing wavefunction $\psi(r, \theta, \phi)$ we need to multiply each of these radial solutions by the spherical harmonics. This is all starting to look a little hairy, but there are some things we can do to simplify the problem. Firstly, we note that the way we have set up the scattering problem implies no dependence on the azimuthal angle, ϕ . The incoming wave defines the z -direction, making the scattering angle θ important, but the localized spherical potential is symmetric, so it can't introduce any ϕ -dependence into the scattered wave. This means we can eliminate from consideration all spherical harmonics with $m_l > 0$:

$$\psi(r, \theta, \phi) = R(r) Y(\theta, \phi) = A \sum_{l=0}^{\infty} c_l h_l^{(1)}(kr) Y_l^0(\theta, \phi). \quad (6.19)$$

Note that in (6.19) we have taken the opportunity to normalize the solutions to the amplitude of the incoming wave, A .

From Eq. (5.34) we can deduce that

$$Y_l^0(\theta, \phi) = \sqrt{\frac{(2l+1)}{4\pi}} P_l(\cos \theta), \quad (6.20)$$

Table 6.1 Spherical Hankel Functions of the First Kind and Legendre Polynomials for Low-Order Values of l

l	Spherical Hankel functions, $h_l^{(1)}(kr)$	Legendre polynomials, $P_l(\cos \theta)$
0	$h_0^{(1)}(kr) = -\frac{i}{kr} e^{ikr}$	$P_0(\cos \theta) = 1$
1	$h_1^{(1)}(kr) = -\frac{(i+kr)}{(kr)^2} e^{ikr}$	$P_1(\cos \theta) = \cos \theta$
2	$h_2^{(1)}(kr) = -\frac{[3i+3kr-(kr)^2 i]}{(kr)^3} e^{ikr}$	$P_2(\cos \theta) = \frac{1}{2} (3\cos^2 \theta - 1)$
3	$h_3^{(1)}(kr) = -\frac{[15i+15kr-6(kr)^2 i-(kr)^3]}{(kr)^4} e^{ikr}$	$P_3(\cos \theta) = \frac{1}{2} (5\cos^3 \theta - 3\cos \theta)$

where the functions $P_l(\cos \theta)$ are the Legendre polynomials, the first few of which are included in Table 6.1 alongside the spherical Hankel functions (and see also Eq. (5.36)).

The second thing we can do is assume that our detector is placed far from the scattering potential, such that r is very large or, more specifically, $kr \gg 1$. In these circumstances, the spherical Hankel functions tend towards a limiting value:

$$h_l^{(1)}(kr) \rightarrow (-i)^{l+1} \frac{e^{ikr}}{kr} \quad \text{valid for } kr \gg 1. \quad (6.21)$$

With these simplifications, the expression for $\psi(r, \theta, \phi)$ in Eq. (6.19) takes the approximate form:

$$\psi(r, \theta, \phi) \approx A \sum_{l=0}^{\infty} (-i)^{l+1} c_l \sqrt{\frac{(2l+1)}{4\pi}} P_l(\cos \theta) \frac{e^{ikr}}{kr}. \quad (6.22)$$

We now choose to *replace* the expansion coefficients, c_l , with *partial wave amplitudes*, $f_l(k)$, defined by

$$f_l(k) = \frac{c_l}{i^{l+1} k \sqrt{4\pi} (2l+1)}. \quad (6.23)$$

Such that Eq. (6.22) becomes

$$\psi(r, \theta, \phi) \approx A \sum_{l=0}^{\infty} (2l+1) f_l(k) P_l(\cos \theta) \frac{e^{ikr}}{r}. \quad (6.24)$$

Or

$$\psi(r, \theta, \phi) \approx Af(\theta) \frac{e^{ikr}}{r},$$

where

$$f(\theta) = \sum_{l=0}^{\infty} (2l+1) f_l(k) P_l(\cos \theta). \quad (6.25)$$

The function $f(\theta)$ is called the *scattering amplitude*. It determines the amplitude of the outgoing spherical wave (given by Ae^{ikr}/r) that is scattered in the direction of the scattering angle, θ .

Step (3): The Differential Scattering Cross-Section

How do we now relate the scattering amplitude defined in Step (2) to the differential scattering cross-section? In other words, how do we turn an amplitude into a number?

It seems we have no choice but to think of the modulus-square of the wavefunction as an intensity or *density* of some kind, but a density of what? Charge density won't really work for us here because we're interested in the *fluxes* of incident and scattered particles, not electrical charge.

I propose to leave this question for the next step, and focus efforts here on deriving a version of the differential scattering cross-section from what we know.

If we take the modulus-square of the wavefunction to be some kind of density, then the proportion of the density of the incoming wave passing through the cross-section $d\sigma$ is given by $|\psi(z)|^2 d\sigma$:

$$|\psi(z)|^2 d\sigma = \left(A^* e^{-ikz}\right) \left(A e^{ikz}\right) d\sigma = A^* A d\sigma = |A|^2 d\sigma. \quad (6.26)$$

Likewise, the proportion of the density of the outgoing spherical wave passing through the detector area dA_{det} is given by (I've added the subscript *det* so that we don't get this area confused with the wave amplitude)

$$\begin{aligned} |\psi(r, \theta, \phi)|^2 dA_{det} &= \left(A^* f^*(\theta) \frac{e^{-ikr}}{r}\right) \left(A f(\theta) \frac{e^{ikr}}{r}\right) dA_{det} \\ &= \frac{|A|^2 |f(\theta)|^2}{r^2} dA_{det}. \end{aligned} \quad (6.27)$$

But we know from Step (2) that $dA_{det} = r^2 d\Omega$, so

$$|\psi(r, \theta, \phi)|^2 dA_{det} = |A|^2 |f(\theta)|^2 d\Omega. \quad (6.28)$$

Logic suggests that $|\psi(z)|^2 d\sigma$ must be equal to $|\psi(r, \theta, \phi)|^2 dA_{det}$ (whatever it is, the proportion of the density passing through $d\sigma$ must be conserved in the scattering process) so we can combine Eqs. (6.26) and (6.28) to give

$$\frac{d\sigma}{d\Omega} = \frac{|A|^2 |f(\theta)|^2}{|A|^2} = |f(\theta)|^2. \quad (6.29)$$

In other words, the differential scattering cross-section is the modulus-square of the scattering amplitude.

To complete the story, we note from Eq. (6.13) that the total cross-section is the integral of the differential scattering cross section:

$$\sigma = \int \frac{d\sigma}{d\Omega} d\Omega = \int |f(\theta)|^2 d\Omega. \quad (6.30)$$

Now, $f(\theta)$ is an infinite series, so its modulus-square is a double-sum:

$$|f(\theta)|^2 = \left[\sum_{l=0}^{\infty} (2l+1) f_l^*(k) P_l(\cos \theta) \right] \left[\sum_{l'=0}^{\infty} (2l'+1) f_{l'}(k) P_{l'}(\cos \theta) \right] \quad (6.31)$$

$$= \sum_l \sum_{l'} (2l+1) (2l'+1) f_l^*(k) f_{l'}(k) P_l(\cos \theta) P_{l'}(\cos \theta).$$

Fortunately, we can make use of the fact that the Legendre functions are orthogonal, such that

$$\int P_l(\cos \theta) P_{l'}(\cos \theta) d\Omega = \frac{4\pi}{(2l+1)} \delta_{ll'}, \quad (6.32)$$

where $\delta_{ll'} = 1$ when $l = l'$ and $\delta_{ll'} = 0$ when $l \neq l'$. This means that the only terms to survive in the integral are those for which $l = l'$:

$$\sigma = \int |f(\theta)|^2 d\Omega = 4\pi \sum_{l=0}^{\infty} (2l+1) |f_l(k)|^2. \quad (6.33)$$

We can't go any further with this without deriving an expression for the partial wave amplitudes, $f_l(k)$, which requires some analysis of the effect of the *interior* of the localized scattering potential on the outgoing waves. There are a number of ways of approaching this which I don't propose to get drawn into here, but the details can be found in textbooks on introductory quantum mechanics.¹²

However, there's one final point that's worth making. In situations in which the scattering potential can be assumed to be *infinite* inside a small spherical volume with radius R (and 0 outside), it is possible to deduce that $\sigma = 4\pi R^2$. This is four times the size of the classical cross-section given in Eq. (6.13) and is the *total* surface area of the sphere. This reflects the fact that the long-wavelength incoming wave wraps itself around the sphere and is diffracted.

Our challenge now is to square the result for the differential scattering cross-section in (6.29) with Eq. (6.10), which defines this in terms of the *number* of particles dN scattered into the detection solid angle $d\Omega$ per unit time and the incident beam luminosity, \mathcal{L} , which is the *number* of incident particles per unit time per unit area.

Now we must address the question: *How do we get from wave amplitudes to numbers of particles?*

Step (4): $|\psi|^2$ as a Probability Density

Born submitted a paper titled 'Quantum Mechanics of Collision Phenomena' to the journal *Zeitschrift für Physik* in June 1926. He had no hesitation in concluding that the

only way to reconcile Schrödinger's wave mechanics with the particle description is to interpret the modulus-square of the wavefunction as a *probability* density. He wrote: 'only one interpretation is possible, $[|\psi|^2]^*$ gives the probability for the electron, arriving from [a specific initial] direction, to be thrown out into [a final] direction'.¹³

Born subsequently claimed that he had been influenced by a remark that Einstein had made in one of his unpublished papers. Einstein had suggested that the wavefunction represents a *Gespenssterfeld*—a kind of 'ghost field'—which determines the probability for a quantum particle to follow a specific path or trajectory. But the reason that Einstein had not published his speculations is that this probabilistic interpretation has profound implications for the notions of physical causality and determinism, notions that Einstein held very dear. Born understood these implications only too well. In his June 1926 paper he wrote:¹⁴

Schrödinger's quantum mechanics therefore gives quite a definite answer to the question of the effect of the collision; but there is no question of any causal description. One gets no answer to the question: 'what is the state after the collision,' but only to the question, 'how probable is a specified outcome of the collision'.

Such sentiments would frame a debate that would continue to rage for nearly a century. If the only information available from quantum mechanics concerns the probabilities of certain outcomes, then strict causality and determinism are abandoned. In quantum transitions, we can no longer say: 'if we do this, then that will happen'. We can only say: 'if we do this, then that will happen with a certain probability'.

We tend to associate probabilities with the operation of some kind of statistics. We deal with probabilities in classical physics because we are *ignorant* of the real states of large, complex systems—we just can't hope to keep track of the behaviour of all the system's many components. A good example is Boltzmann's use of statistical methods to describe the properties of atomic and molecular gases. In such situations, we can perhaps be confident that the principles of causality and determinism hold at the microscopic level for each individual particle undergoing a sequence of collisions, but we are simply in no position to follow these motions individually. Instead, we deal with statistical averages, and probabilities.

But this is not what quantum probability implies, as Born explained in a lecture delivered to a meeting of the British Association held in Oxford, England, a few months later in August 1926:¹⁵

The classical theory introduced the microscopic co-ordinates which determine the individual process, only to eliminate them because of ignorance by averaging over their values; whereas the new [quantum] theory gets the same results without introducing them at all. Of course, it is not forbidden to believe in the existence of these co-ordinates; but they will only be of physical significance when methods have been devised for their experimental observation.

* In his original paper Born wrote that the wavefunction represents a probability but corrected this in a footnote added at the proof stage.

Born concluded his lecture with the remark: ‘the fundamental idea of waves of probability will probably persist in one form or another’.¹⁶

‘Finding’ the Electron: The Radial Distribution Function

We’re now in a position to return to the question I posed at the end of Chapter 5. In the hydrogen 1s orbital described by the wavefunction ψ_{100} , *just where, exactly, is the electron?*

In fact, it was Wolfgang Pauli who proposed to interpret the modulus-square of the wavefunction not only as a transition probability or as the probability for the system to be in a specific state, as Born had done, but as the probability of ‘finding’ the electron at a specific position in its orbit inside an atom. This gives rise to an image of a particulate electron which is somehow ‘smeared’ through space with, at any one time, different probabilities for being found in different positions around the nucleus.

To answer the question it is necessary to reframe it slightly. We ask what is the *probability* of ‘finding’ the electron in a thin shell around the nucleus with volume dV ? Based on the logic we have examined in the past section, Step (4), we assume that this probability is given by $|\psi_{100}|^2 dV$. But we know that the 1s orbital is spherically symmetric, meaning that it doesn’t depend on the angles θ and ϕ , so we can write

$$dV = r^2 dr \int_0^\pi \sin \theta \, d\theta \int_0^{2\pi} d\phi = 4\pi r^2 dr. \quad (6.34)$$

Using the result for ψ_{100} in (5.41), the probability function then becomes

$$4\pi r^2 |\psi_{100}|^2 dr = \frac{4r^2}{a_0^3} e^{-2r/a_0} dr. \quad (6.35)$$

The function $4r^2 e^{-2r/a_0} / a_0^3$ is called the *radial distribution function*, and this is plotted (multiplied by a_0) against the ratio r/a_0 in Fig. 6.4. Quite remarkably, this suggests that the electron may be found anywhere, but it has the highest probability of being found precisely at the Bohr radius, a_0 . Another incredible coincidence, courtesy of the Coulomb potential.

But ...

This might seem all perfectly reasonable, but of course it doesn’t really answer our question. Look back at Fig. 4.3 in Chapter 4. Now imagine that we reduce the intensity of the electron beam in this experiment such that, on average, only *one* electron at a time passes through the two slits. Look back at Fig. 6.3, and again imagine a situation in which a single electron is scattered, producing a spherical outgoing electron wave. We know that in both these cases the wavefunction for the electron that passes through

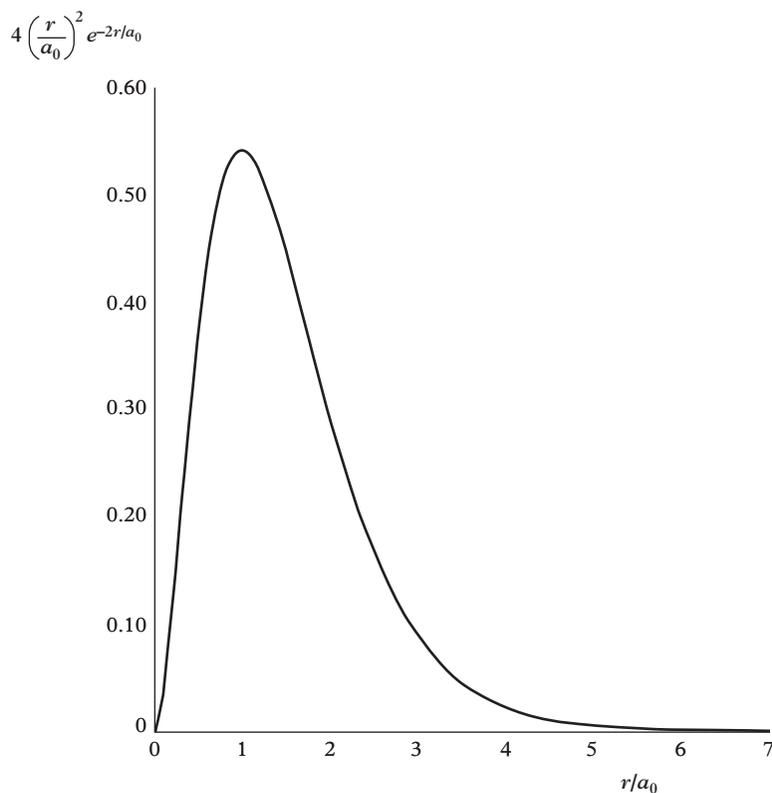


Figure 6.4 The radial distribution function for the hydrogen atom $1s$ orbital, $4\pi r^2 |\psi_{100}|^2$, multiplied by a_0 . The probability of 'finding' the electron is greatest at the Bohr radius, $r/a_0 = 1$.

the two slits or is scattered from the potential is *delocalized*. It is an extended wave with amplitude in many different places. We reach for Born's interpretation and we deduce that the probability of finding the electron is similarly extended and delocalized—it has a certain probability of being found 'here', 'there', and, well, anywhere it has non-zero amplitude. Yet we anticipate that it will register on a photographic plate or in a detector as a single, self-contained particle that is unambiguously 'here', and nowhere else.

It seems we have a recipe for 'finding' the electron, but nothing to tell us precisely *how* it gets there.

Some physicists were deeply troubled by this. Despite the fact that Schrödinger was dropping broad hints in this direction with his reference to an electron being in all places 'kinematically imaginable' but 'not equally strongly' (see the section ' $|\psi|^2$ as a Charge Density' earlier in this chapter), he was not persuaded by Born's arguments. Born had acknowledged his debt of inspiration to Einstein's *Gespenssterfeld* in a letter to Einstein dated 30 November 1926. But Einstein rejected the loss of causality and the element of

chance implied by Born’s interpretation, and in his reply he summarized the nature and extent of his doubts:¹⁷

Quantum mechanics is very impressive. But an inner voice tells me that it is not yet the real thing. The theory produces a good deal but hardly brings us closer to the secret of the Old One. I am at all events convinced that *He* does not play dice.

Einstein, whose speculative light-quantum hypothesis had helped to establish the new quantum theory in 1905, was rapidly becoming one of the theory’s most determined critics.

.....

NOTES

1. Erwin Schrödinger, ‘The Fundamental Idea of Wave Mechanics’, Nobel Lecture, 12 December 1933, *Nobel Lectures, Physics 1922–1941*, Elsevier, Amsterdam, 1965, p. 306.
2. Ibid. 314.
3. Erwin Schrödinger, ‘Quantisation as a Problem of Proper Values (Part I)’, in Stephen Hawking (ed.), *The Dreams That Stuff Is Made Of: The Most Astounding Papers on Quantum Physics and How They Shook the World*, Running Press, Philadelphia, 2011, p. 264.
4. Erwin Schrödinger, ‘Quantisation as a Problem of Proper Values (Part IV)’, in Hawking (ed.), *The Dreams That Stuff Is Made Of*, p. 382.
5. Ibid. 383.
6. At first glance this expression for the volume element $d\tau$ might look a bit complicated and unfamiliar, but we can make it much more familiar simply by integrating it to give the total volume:

$$\int d\tau = \int_0^R r^2 dr \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi = \frac{1}{3}R^3 \cdot 2 \cdot 2\pi = \frac{4}{3}\pi R^3,$$

where R is the radius of the sphere.

7. Werner Heisenberg, *Physics and Beyond: Memories of a Life in Science*, Allen & Unwin, London, 1971, p. 61.
8. Werner Heisenberg, manuscript of a lecture intended for delivery in Göttingen, May 1975. Subsequently published in Werner Heisenberg, *Encounters with Einstein*, Princeton University Press, Princeton, NJ, 1983. This quote appears on p. 45.
9. Many of the most important papers on the development of matrix mechanics have been collected together and published in B. L. van der Waerden, *Sources of Quantum Mechanics*, Dover, New York, 1968. See especially Part II, pp. 261–427.
10. Werner Heisenberg, letter to Wolfgang Pauli, 8 June 1926. Quoted in David C. Cassidy, *Uncertainty: The Life and Science of Werner Heisenberg*, W. H. Freeman, New York, 1992, p. 215.
11. Max Born, ‘Zur Quantenmechanik der Stossvorgänge’, *Zeitschrift für Physik*, 37 (1926), 863–7. An English translation is available in John Archibald Wheeler and Wojciech Hubert Zurek (eds), *Quantum Theory and Measurement*, Princeton University Press, Princeton, NJ, 1983. This quote appears on p. 52.

12. A really good place to look is David J. Griffiths, *Introduction to Quantum Mechanics*, 2nd edn, Cambridge University Press, Cambridge, UK, 2017, Chapter 11.
13. Born, 'Zur Quantenmechanik der Stossvorgänge', reproduced in Wheeler and Zurek (eds), *Quantum Theory and Measurement*. This quote appears on p. 54.
14. Ibid.
15. Max Born, 'Physical Aspects of Quantum Mechanics', paper delivered to the British Association for the Advancement of Science in Oxford on 10 August 1926, translated by J. Robert Oppenheimer and first published in *Nature*, **119** (1927), 354–57. This is reproduced in Max Born, *Physics in My Generation*, 2nd edn, Springer-Verlag, New York, 1970. This quote appears on p. 10.
16. Ibid. 12.
17. Albert Einstein, letter to Max Born, 4 December 1926. Quoted in Abraham Pais, *Subtle is the Lord: The Science and the Life of Albert Einstein*. Oxford University Press, 1982, p. 443.

.....

FURTHER READING

- Born, Max, *Physics in My Generation*, 2nd edn, Springer-Verlag, New York, 1970, pp. 6–30.
- Griffiths, David J., *Introduction to Quantum Mechanics*, 2nd edn, Cambridge University Press, Cambridge, UK, 2017, Chapter 11.
- Hawking, Stephen (ed.), *The Dreams that Stuff Is Made Of: The Most Astounding Papers on Quantum Physics and How They Shook the World*, Running Press, Philadelphia, 2011, pp. 358–87.
- Kibble, Tom W. B., and Berkshire, Frank H., *Classical Mechanics*, 5th edn, Imperial College Press, London, 2004, Chapter 4: especially pp. 90–4.
- Wheeler, John Archibald, and Zurek, Wojciech Hubert (eds), *Quantum Theory and Measurement*, Princeton University Press, Princeton, NJ, 1983, pp. 50–5.

Heisenberg, Bohr, Robertson, and the Uncertainty Principle

The Interpretation of Quantum Uncertainty

Heisenberg was pretty angry. The foundations of his new matrix mechanics were at great risk of being completely undermined by Schrödinger's intuitively appealing and mathematically more familiar and accessible wave mechanics. As more members of the physics community began to defect—with great relief—to the wave approach, Heisenberg grew increasingly dismayed.

Bohr shared Heisenberg's misgivings, and in September 1926 he invited Schrödinger to Copenhagen to debate the issues further. This debate began even as Bohr and Heisenberg collected Schrödinger from the railway station. Bohr, normally considerate and friendly, confronted Schrödinger like some kind of fanatic. But Schrödinger had by now developed an entrenched view on how he believed his wave mechanics should be interpreted. Moreover, he had been greatly encouraged by the warmth with which the physics community had received his approach, and he was in no mood to back down.

Schrödinger cherished the continuity and visualizability afforded by wave concepts, and simply could not see how discontinuous quantum jumps would fit with such a scheme. 'Surely you realize', Schrödinger pleaded, 'that the whole idea of quantum jumps is bound to end in nonsense.' He continued, with growing exasperation: 'If all this damned quantum jumping were really here to stay, I should be sorry I ever got involved with quantum theory.'¹

The debate raged day and night, and Schrödinger succumbed to a feverish cold. Bohr pursued him even to his sickbed. As Bohr's wife Margrethe attempted to nurse him back to health with tea and cake, Bohr sat at the edge of the bed and continued his harangue: 'But you must surely admit that. . .'.² Schrödinger would admit nothing.

The debate had a powerful effect on Bohr. Schrödinger's grasping for classical wave concepts and his relentless pursuit of a theory which provided some kind of space-time visualizability brought home to Bohr the precise nature of the problems they were all grappling with.

Heisenberg, too, paused to reflect. He had long ago abandoned any attempt to use classical space-time concepts to describe what he believed to be an essentially

unvisualizable physics. Inevitably, whilst he was content to embrace Born's probability interpretation, he was rather unhappy that Born had reached for wave mechanics as the 'deepest formulation of the quantum laws'.³ He felt that Born's conclusion still left too much room for alternative interpretations. Born, in the meantime, was beginning to regret the 'victory' of the wave approach over matrix mechanics, as it tended to drag Schrödinger's interpretation along with it.

Heisenberg was now living in a small attic flat in Bohr's Institute, with slanting walls and windows which overlooked the entrance to nearby Fælled Park. Bohr would come to the flat late at night, and the two would discuss their scientific problems into the small hours.

Although they disagreed on many points, Heisenberg had good reason to believe they were heading in more or less the same direction. He was astonished at how simple phenomena—such as the observed trajectory of an electron revealed, for example, in a cloud chamber—were proving so intractable.* Indeed, the very concept of a trajectory was absent from matrix mechanics. Schrödinger's proposal that this was the trajectory of an electron matter 'wave-packet' was quickly dismissed—Hendrik Lorentz had pointed out that any such wave-packet is unstable and could be expected to disperse very quickly. So the trajectory was absent from wave mechanics, too. And yet it was difficult for anyone who had looked at the track left by an electron not to be convinced of the reality of the electron's particle-like motion.

Heisenberg summarized this period of intense debate as follows:⁴

I remember discussions with Bohr which went through many hours till very late at night and ended almost in despair; and when at the end of the discussion I went alone for a walk in the neighbouring park I repeated to myself again and again the question: Can nature possibly be as absurd as it seemed . . . ?

They reached no real conclusions. Their protracted debate left them exhausted and somewhat tense. When Bohr decided to take a skiing holiday in February 1927, Heisenberg was greatly relieved.

The Dark Point

From the outset, Heisenberg had resolved to eliminate classical space-time pictures involving particles and waves from the quantum mechanics of the atom. He had wanted to focus instead on the properties that were actually observed and recorded in laboratory experiments, such as the positions and intensities of spectral lines. Left alone in Copenhagen in February 1927, he now pondered on the significance and meaning of

* The cloud chamber was invented by Charles Wilson, a student of J. J. Thomson. As an energetic particle passes through such a chamber, it dislodges electrons from atoms in the vapour contained in the chamber, leaving charged ions in its wake. Water droplets condense around the ions, revealing the particle trajectory.

such experimental observables. Somehow, he needed to introduce at least some form of ‘visualizability’ into matrix mechanics.

This train of thought was to lead him to another fundamental discovery. He drew inspiration from dialogues with Pascual Jordan in Göttingen and with Paul Dirac, who had arrived in Copenhagen in September 1926 for a six-month period as a visiting fellow. Pauli, too, provided some significant clues. In a letter Heisenberg had received in October 1926, Pauli made an important observation concerning the relationship between position and momentum in quantum mechanics. He had considered the situation where two electrons collide. When the electrons are far apart, they can be conveniently treated as plane waves, each with a clearly defined position (x) and momentum (p_x). But as they come together they manifest what Pauli called a ‘dark point’, at which things become fuzzy. If the positions are assumed to be controlled, then the momenta are uncontrolled, and vice versa. He wrote: ‘One may view the world with the p -eye and one may view it with the $[x]$ -eye, but if one opens both eyes at the same time one becomes crazy’.⁵

The Position–Momentum Commutation Relation

What was Pauli getting at? In classical mechanics, we know that the product of position and momentum is completely independent of the order in which we perform the multiplication. Returning for a moment to motion only in the x -direction, we would not hesitate to conclude that $x \times p_x$ is identical to $p_x \times x$. Indeed, the question seems faintly ridiculous. In mathematical terms, quantities that behave like this are said to *commute*, and their *commutation relation*, written as $[x, p_x] = xp_x - p_x x$, is quite obviously equal to 0. In such a relation, the term $[x, p_x]$ is called the *commutator*.

But in wave mechanics, the linear momentum is no longer a simple quantity. It has become a differential operator, and operators need a function to operate on. We can immediately deduce the consequences, and we don’t even need to assume a specific form for the wavefunction itself.

If we restrict ourselves once again to the x -direction, then (5.22) reduces to $\hat{p}_x = -i\hbar d/dx$ and we have

$$x\hat{p}_x\psi(x) = -i\hbar x \frac{d}{dx}\psi(x) \quad (7.1)$$

$$\hat{p}_x x\psi(x) = -i\hbar \frac{d}{dx}[x\psi(x)] = -i\hbar\psi(x) - i\hbar x \frac{d}{dx}\psi(x), \quad (7.2)$$

where we have used the product rule to evaluate the differential of $x\psi(x)$ in (7.2). If we subtract (7.2) from (7.1) we can immediately see that

$$x\hat{p}_x\psi(x) - \hat{p}_x x\psi(x) = [x, \hat{p}_x]\psi(x) = i\hbar\psi(x). \quad (7.3)$$

Or

$$[x, \hat{p}_x] = i\hbar. \quad (7.4)$$

In other words, in wave mechanics position and linear momentum no longer commute. The order in which we apply the operations ‘multiply by x ’ and ‘differentiate by x and multiply by $-i\hbar$ ’ makes a difference to the final result. Things become fuzzy. This mirrors precisely what Heisenberg had discovered from matrix mechanics. He later wrote:⁶

In quantum mechanics the relation [Eq. (7.4)] between mass, position and velocity is believed to hold. Therefore we have good reason to become suspicious every time uncritical use is made of the words ‘position’ and ‘velocity’. When one admits that discontinuities are somehow typical of processes that take place in small regions and in short times, then a contradiction between the concepts of ‘position’ and ‘velocity’ is quite plausible.

The Uncertainty Principle

Heisenberg now tried to use matrix mechanics to describe the clearly visible path of an electron in a cloud chamber. He quickly ran into trouble. It was now after midnight, but he set off for a walk in Fælled Park. As he walked in the darkness he asked himself some fairly searching, fundamental questions, such as: What do we actually *mean* when we talk about the *position* of an electron? The track caused by the passage of an electron through a cloud chamber seems real enough—surely it provides an unambiguous measure of the electron’s trajectory through space?

But the track is made visible by the condensation of water droplets around atoms that have been ionized by the electron as it passes through the chamber. The droplets are much larger than the electron they have been used to ‘detect’, suggesting that the instantaneous position and velocity of the electron through the cloud chamber can, in truth, be known only approximately.

He had found the right question: ‘Can quantum mechanics represent the fact that an electron finds itself approximately in a given place and that it moves approximately with a given velocity, and can we make these approximations so close that they do not cause experimental difficulties?’⁷ He returned to his room and quickly demonstrated to himself that he could, indeed, represent such approximations mathematically using matrix mechanics. He had found the connection he was looking for.

Heisenberg had discovered the *uncertainty principle*: the product of the ‘uncertainties’ in certain pairs of variables—called *complementary variables*—such as position and momentum cannot be smaller than Planck’s constant h .^{*} In other words, quantum mechanics places a fundamental limit on the precision with which both position and

^{*} We will soon see that this is modified to $\frac{1}{2}\hbar$.

momentum can be jointly determined in any observation. In an entirely classical world, where h is approximated to be zero, no such constraint exists.

This suggests a deep connection between quantum noncommutation and uncertainty but, as we will see towards the end of this chapter, the reasoning that Heisenberg himself set out in his 1927 paper on the subject proved to be faulty. We will therefore derive one from the other based on an original derivation published by the American physicist Howard Robertson a few years after Heisenberg, in 1929.⁸

But before we can do this, we need to do a little more groundwork.

The Expansion Theorem

The derivation we will consider in this chapter will obviously involve manipulation of quantum-mechanical operators and commutators, such as $[x, \hat{p}_x]$. It will therefore be helpful to take some time here to explore some of the properties of these operators, with the aim of establishing some useful general principles.

Consider an operator \hat{A} which operates on a normalized wavefunction Ψ . The first question we need to consider is this: What do we do if the wavefunction is not an eigenfunction of \hat{A} ? To a certain extent, we already have the answer to this from our partial wave analysis of the scattered electron wave in Chapter 6. Wavefunctions are very flexible things. We can add them together—forming linear combinations or ‘superpositions’—in ways that will reproduce the function we need. If we choose wavefunctions that we already know or can assume to be solutions to the Schrödinger wave equation relevant for our problem, and we combine them in the right way, then we can be confident that the resulting composite wavefunctions will also be valid solutions. This is the *expansion theorem*.

We’re free to choose whatever set of functions can be combined to reproduce the wavefunction Ψ , but it makes sense to expand Ψ in a ‘basis’ formed from all the relevant eigenfunctions of the operator \hat{A} . Or

$$\Psi = \sum_i c_i \Psi_i. \quad (7.5)$$

The functions Ψ_i are eigenfunctions of \hat{A} , such that $\hat{A}\Psi_i = a_i\Psi_i$ where a_i are the corresponding eigenvalues. The c_i are mixing or *expansion coefficients*, representing the proportions of the individual eigenfunctions that must be included in the superposition to reproduce our original wavefunction. We can presume that the eigenfunctions are orthonormal: $\int \Psi_i^* \Psi_j d\tau = \delta_{ij}$.

We don’t have to go too far to find a demonstrative example. As we saw in Chapter 5, when we solve the Schrödinger equation for the electron in a hydrogen atom, we get a series of eigenfunctions, the first few of which are detailed in Eqs. (5.41)–(5.44). Most physicists are perfectly at ease with the expressions for the $2p_0$ and $2p_{\pm 1}$ orbitals, even though the latter are complex (they contain the term $e^{\pm i\phi}$). But chemists prefer to think

about these orbitals in terms of $x, y,$ and z coordinates, as directions in three-dimensional space become relevant as soon as we start to assemble nonlinear molecules.

Of course, the $2p_0$ orbital isn't a problem, as it is noncomplex and directed along the z -axis, so we simply rename it as the $2p_z$ orbital. But to derive the $2p_x$ and $2p_y$ orbitals we need to take the normalized linear combinations

$$\begin{aligned}\psi_{2p_x} &= \frac{1}{\sqrt{2}} (\psi_{21-1} - \psi_{21+1}) = \frac{1}{8\sqrt{2\pi a_0^3}} \rho_n e^{-\rho_n/2} \sin \theta (e^{-i\phi} + e^{i\phi}) \\ &= \frac{1}{4\sqrt{2\pi a_0^3}} \rho_n e^{-\rho_n/2} \sin \theta \cos \phi,\end{aligned}\tag{7.6}$$

where we have used the exponential form for $\cos \phi = \frac{1}{2} (e^{i\phi} + e^{-i\phi})$, and

$$\begin{aligned}\psi_{2p_y} &= \frac{i}{\sqrt{2}} (\psi_{21-1} + \psi_{21+1}) = \frac{i}{8\sqrt{2\pi a_0^3}} \rho_n e^{-\rho_n/2} \sin \theta (e^{-i\phi} - e^{i\phi}) \\ &= \frac{1}{4\sqrt{2\pi a_0^3}} \rho_n e^{-\rho_n/2} \sin \theta \sin \phi,\end{aligned}\tag{7.7}$$

where we have used the exponential form for $\sin \phi = \frac{1}{2} i (e^{i\phi} - e^{-i\phi})$.

You might then ask: Which is the 'correct' form of the wavefunctions for these $2p$ orbitals? The simple answer is there really isn't one—so long as we make use of eigenfunctions of the Schrödinger equation and pay attention to the need for normalization, then we can choose the combinations most appropriate to the problem we're trying to solve.

The Expectation Value

To keep things simple, let's suppose that we need to consider only two eigenfunctions:

$$\Psi = c_m \Psi_m + c_n \Psi_n.\tag{7.8}$$

The next question is this: In the operation of \hat{A} on Ψ , what *result* can we expect to observe? The expansion coefficients are simply numbers, so they can be pulled in front of the operator as follows:

$$\hat{A}\Psi = c_m \hat{A}\Psi_m + c_n \hat{A}\Psi_n = c_m a_m \Psi_m + c_n a_n \Psi_n.\tag{7.9}$$

If we multiply through by the complex conjugate Ψ^* we get

$$\begin{aligned}\Psi^* \hat{A} \Psi &= (c_m^* \Psi_m^* + c_n^* \Psi_n^*) (c_m a_m \Psi_m + c_n a_n \Psi_n) \\ &= |c_m|^2 a_m |\Psi_m|^2 + c_m^* c_n a_n \Psi_m^* \Psi_n + c_n^* c_m a_m \Psi_n^* \Psi_m + |c_n|^2 a_n |\Psi_n|^2.\end{aligned}\quad (7.10)$$

If we now integrate over all coordinates, we know that $\int |\Psi_m|^2 d\tau = \int |\Psi_n|^2 d\tau = 1$ and $\int \Psi_m^* \Psi_n d\tau = \int \Psi_n^* \Psi_m d\tau = 0$, so the integral reduces to

$$\int \Psi^* \hat{A} \Psi d\tau = \langle \hat{A} \rangle = |c_m|^2 a_m + |c_n|^2 a_n.\quad (7.11)$$

This integral is referred to as the *expectation value* of the operator, symbol $\langle \hat{A} \rangle$, and is the mean or average value of the observable corresponding to the operator \hat{A} . As we can see from (7.11), in the simple case we have considered it is indeed the average of the two possible eigenvalues, each weighted by the modulus-squares of their corresponding expansion coefficients. Note that if Ψ is normalized, then

$$\int \Psi^* \Psi d\tau = |c_m|^2 + |c_n|^2 = 1.\quad (7.12)$$

In general, $\langle \hat{A} \rangle = \sum_i |c_i|^2 a_i$, and $\sum_i |c_i|^2 = 1$, and we now see why we used the factor $1/\sqrt{2}$ and $i/\sqrt{2}$ in Eqs. (7.6) and (7.7).

If we want to be stubborn, we might want to insist on a more detailed answer to our second question. The expectation value is a weighted average of the possible eigenvalues, but what result are we going to get with the *next* observation? The answer is *either* a_m *or* a_n , with probabilities $P_m = |c_m|^2$ and $P_n = |c_n|^2$, respectively. There is nothing in Eq. (7.11) to tell us which eigenvalue we will get, just as there is nothing in the equation for the diffracted electron wave to tell us where the next electron will be detected in Fig. 4.3, or where on the spherical wavefront in Fig. 6.3 the next scattered electron will be found.

Be reassured, we'll be coming back to this.

Hermitian Operators

We can see from (7.11) that provided the eigenvalues a_m and a_n are themselves real quantities, then the expectation value will always be exclusively real, as required if it is to represent a physical observable. Operators that yield exclusively real eigenvalues are called *Hermitian* or *self-adjoint operators*, and are *defined* by possession of the property:

$$\begin{aligned}\int \Psi_m^* \hat{A} \Psi_n d\tau &= \left(\int \Psi_n^* \hat{A} \Psi_m d\tau \right)^* \\ &= \int \Psi_n (\hat{A} \Psi_m)^* d\tau = \int (\hat{A} \Psi_m)^* \Psi_n d\tau.\end{aligned}\quad (7.13)$$

The second term in (7.13) is the Hermitian conjugate, or the *complex conjugate transpose* of the integral, formed by transposing the indices m and n and taking the complex conjugate of the result. We can see immediately that when $m = n$,

$$\int \Psi_n^* \hat{A} \Psi_n d\tau = \int \Psi_n^* a_n \Psi_n d\tau = a_n \int \Psi_n^* \Psi_n d\tau = a_n. \quad (7.14)$$

Similarly, for the Hermitian conjugate,

$$\int (\hat{A} \Psi_n)^* \Psi_n d\tau = \int (a_n \Psi_n)^* \Psi_n d\tau = a_n^* \int \Psi_n^* \Psi_n d\tau = a_n^*. \quad (7.15)$$

And so from (7.13) we conclude that $a_n = a_n^*$; any (Hermitian) operator satisfying (7.13) yields eigenvalues that are exclusively real. We can also show quite straightforwardly that the eigenfunctions corresponding to different eigenvalues of a Hermitian operator are orthogonal.

Is the momentum operator \hat{p}_x Hermitian? Let's find out. Suppose our wavefunctions are confined to the x -direction and consider the integral

$$\int_{-\infty}^{+\infty} \psi_m^*(x) \hat{p}_x \psi_n(x) dx = -i\hbar \int_{-\infty}^{+\infty} \psi_m^*(x) \frac{d\psi_n(x)}{dx} dx. \quad (7.16)$$

We can make use of integration by parts, $\int u \frac{dv}{dx} dx = uv - \int v \frac{du}{dx} dx$, to determine that

$$\begin{aligned} -i\hbar \int_{-\infty}^{+\infty} \psi_m^*(x) \frac{d\psi_n(x)}{dx} dx &= -i\hbar [\psi_m^*(x) \psi_n(x)]_{-\infty}^{+\infty} \\ &\quad + i\hbar \int_{-\infty}^{+\infty} \psi_n(x) \frac{d\psi_m^*(x)}{dx} dx. \end{aligned} \quad (7.17)$$

We can presume that the wavefunctions $\psi_m(x)$ and $\psi_n(x)$ vanish at $\pm\infty$, and we note that

$$i\hbar \int_{-\infty}^{\infty} \psi_n(x) \frac{d\psi_m^*(x)}{dx} dx = \int_{-\infty}^{\infty} [\hat{p}_x \psi_m(x)]^* \psi_n(x) dx, \quad (7.18)$$

which is consistent with Eq. (7.13). The linear momentum operator is indeed Hermitian.

Variance and Standard Deviation

The final item of background we need to consider concerns some elements of statistical analysis. We know in our imperfect world that if we make a series of repeated observations or measurements, we will likely get a distribution of results around a mean or average. What follows is entirely classical so, for a moment at least, we can forget all about operators and wavefunctions.

Let's suppose we observe a quantity A , and discover that the mean value of A over a series of observations is $\langle A \rangle$, given by

$$\langle A \rangle = \frac{1}{N} \sum_i A_i N_i = \sum_i A_i \frac{N_i}{N} = \sum_i A_i P_i. \quad (7.19)$$

Here the A_i represent all the different possible values that A may exhibit, and N_i is the number of times each value appears in a series of N observations. Consequently, P_i is the classical(!) probability of the occurrence of A_i in the series, and $\sum_i P_i = 1$. It logically follows from this that any *function* of A_i (such as A_i^2 or $\sqrt{A_i}$) cannot occur with any higher or lower probability than A_i itself. So we can conclude that the mean of the square of the observed values is given by

$$\langle A^2 \rangle = \sum_i A_i^2 \frac{N_i}{N} = \sum_i A_i^2 P_i. \quad (7.20)$$

We presume that each observed value A_i deviates from the mean by an amount ΔA , such that

$$\Delta A = A_i - \langle A \rangle. \quad (7.21)$$

This deviation can obviously be positive or negative, so we choose to define the *variance* from the mean as the average value of the square of the deviations, which is always positive:

$$\sigma^2 = \langle (\Delta A)^2 \rangle = \langle (A_i - \langle A \rangle)^2 \rangle. \quad (7.22)$$

From this, (7.19) and (7.20), we can deduce that

$$\begin{aligned} \sigma^2 &= \langle (\Delta A)^2 \rangle = \sum_i (\Delta A)^2 P_i = \sum_i (A_i - \langle A \rangle)^2 P_i \\ &= \sum_i (A_i^2 - 2A_i \langle A \rangle + \langle A \rangle^2) P_i \\ &= \sum_i A_i^2 P_i - 2\langle A \rangle \sum_i A_i P_i + \langle A \rangle^2 \sum_i P_i. \end{aligned} \quad (7.23)$$

But we know that $\sum_i A_i^2 P_i = \langle A^2 \rangle$, $\sum_i A_i P_i = \langle A \rangle$, and $\sum_i P_i = 1$. Hence

$$\sigma^2 = \langle (\Delta A)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2. \quad (7.24)$$

The *standard deviation* is then calculated as the square root of the variance:

$$\sigma = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}. \quad (7.25)$$

In the above, I've assumed that the values A_i are discrete, like the two sides of a coin or the faces of a die. But the expressions (7.24) and (7.25) hold also for continuous variables. In such situations, the discrete probabilities P_i are replaced by the integral of a probability density, which returns the probability for observing values over the range of the integral.

This might have seemed like an unnecessary detour, but Heisenberg's uncertainty principle concerns the 'uncertainties'—or, less loosely and more precisely, standard deviations—of the values of observable quantities such as position and momentum.

We can now bring all these ingredients together. What follows is based loosely on what Robertson did in 1929.

The Ingredients

1. Expression for the variance, Eq. (7.22).
2. The expression for the quantum-mechanical expectation value, Eq. (7.11).
3. The properties of Hermitian operators, Eq. (7.13).
4. Something called the integral Cauchy–Schwarz inequality.
5. The general properties of complex numbers.
6. The position–momentum commutation relation, Eq. (7.4).

The Recipe

In Step (1) we establish a quantum operator for variance and simplify this by making use of its Hermitian properties. We use this in Step (2) to deduce an expression for the product of the variances related to two different operators, \hat{A} and \hat{B} , and make use of the integral Cauchy–Schwarz inequality. We use some general relationships of complex numbers in Step (3) to recast the expression for the product of the variances in terms of the commutator, $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$. In Step (4) we derive the position–momentum uncertainty relation from the position–momentum commutator, $[x, \hat{p}_x]$. Finally, in Step (5) we derive the uncertainty relation for energy and time.

Step (1): The Operator for Variance

We've learned that in order to determine the values of observable quantities in quantum mechanics, we need to deduce the operator that will return these values from the

wavefunction, and then apply it. From Eq. (7.22), we see that the classical statistical variance is related to the square of the difference between the observed value and the mean. From this it follows that the quantum-mechanical equivalent is the expectation value of the ‘variance operator’. By analogy with Eq. (7.11), we simply replace the observed value with its corresponding quantum-mechanical operator, \hat{A} , and the mean by the expectation value of this operator, $\langle \hat{A} \rangle$, to give

$$\sigma_{\hat{A}}^2 = \langle \Delta \hat{A} \rangle^2 = \int \Psi^* (\hat{A} - \langle \hat{A} \rangle)^2 \Psi d\tau. \quad (7.26)$$

We can reassure ourselves that this is correct by expanding the operator:

$$\begin{aligned} \sigma_{\hat{A}}^2 &= \int \Psi^* (\hat{A} - \langle \hat{A} \rangle) (\hat{A} - \langle \hat{A} \rangle) \Psi d\tau \\ &= \int \Psi^* (\hat{A} - \langle \hat{A} \rangle) (\hat{A}\Psi - \langle \hat{A} \rangle \Psi) d\tau \\ &= \int \Psi^* (\hat{A}^2 \Psi - 2\langle \hat{A} \rangle \hat{A}\Psi + \langle \hat{A} \rangle^2 \Psi) d\tau \\ &= \int \Psi^* \hat{A}^2 \Psi d\tau - 2\langle \hat{A} \rangle \int \Psi^* \hat{A}\Psi d\tau + \langle \hat{A} \rangle^2 \int \Psi^* \Psi d\tau. \end{aligned} \quad (7.27)$$

But $\int \Psi^* \hat{A}^2 \Psi d\tau = \langle \hat{A}^2 \rangle$, $\int \Psi^* \hat{A}\Psi d\tau = \langle \hat{A} \rangle$, and $\int \Psi^* \Psi d\tau = 1$, so

$$\sigma_{\hat{A}}^2 = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2, \quad (7.28)$$

and we see that this logic precisely mirrors that of (7.23) and (7.24), emphasizing once again the role of the modulus-square of the wavefunction as a probability density.

Equation (7.26) is, however, not quite the form we need. The variance is surely a real physical quantity, so we make use of the Hermitian property of $(\hat{A} - \langle \hat{A} \rangle)$ to write

$$\sigma_{\hat{A}}^2 = \int \Psi^* (\hat{A} - \langle \hat{A} \rangle) (\hat{A} - \langle \hat{A} \rangle) \Psi d\tau = \int [(\hat{A} - \langle \hat{A} \rangle) \Psi]^* (\hat{A} - \langle \hat{A} \rangle) \Psi d\tau. \quad (7.29)$$

Or

$$\sigma_{\hat{A}}^2 = \int |(\hat{A} - \langle \hat{A} \rangle) \Psi|^2 d\tau. \quad (7.30)$$

Step (2): The Product of the Variances of Two Operators

We can follow the same reasoning for a second operator, \hat{B} , with variance σ_B^2 :

$$\sigma_B^2 = \int \left| (\hat{B} - \langle \hat{B} \rangle) \Psi \right|^2 d\tau. \quad (7.31)$$

We should note in passing that for this to work, the wavefunction Ψ must be a simultaneous eigenfunction of both \hat{A} and \hat{B} , i.e. $\hat{A}\Psi = a\Psi$ and $\hat{B}\Psi = b\Psi$, where a and b are the corresponding eigenvalues.

The product $\sigma_A^2 \sigma_B^2$ is then given by

$$\sigma_A^2 \sigma_B^2 = \int \left| (\hat{A} - \langle \hat{A} \rangle) \Psi \right|^2 d\tau \cdot \int \left| (\hat{B} - \langle \hat{B} \rangle) \Psi \right|^2 d\tau. \quad (7.32)$$

We can simplify this by setting $f = (\hat{A} - \langle \hat{A} \rangle)\Psi$ and $g = (\hat{B} - \langle \hat{B} \rangle)\Psi$, giving

$$\sigma_A^2 \sigma_B^2 = \int |f|^2 d\tau \cdot \int |g|^2 d\tau. \quad (7.33)$$

At this point we introduce the integral Cauchy–Schwarz inequality:

$$\int |f|^2 d\tau \cdot \int |g|^2 d\tau \geq \left| \int (f^* g) d\tau \right|^2. \quad (7.34)$$

I don't propose to derive this inequality here, but interested readers should consult Appendix 5. So, from (7.33) and (7.34) we have

$$\sigma_A^2 \sigma_B^2 \geq \left| \int (f^* g) d\tau \right|^2. \quad (7.35)$$

Step (3): Introduce the Commutator, $[\hat{A}, \hat{B}]$

In order to take the next step we need to remind ourselves of some general properties of complex numbers. A complex number z consists of a real part, $\text{Re}(z) = x$, and an imaginary part, $\text{Im}(z) = y$, such that

$$z = x + iy \quad \text{and} \quad z^* = x - iy. \quad (7.36)$$

It follows that

$$\operatorname{Re}(z) = x = \frac{1}{2}(z + z^*) \text{ and } \operatorname{Im}(z) = y = \frac{1}{2i}(z - z^*). \quad (7.37)$$

We can further show that $|z|^2 = zz^* = x^2 + y^2$, and so it follows that $|z|^2 \geq y^2$, or $|z|^2 \geq |(z - z^*)/2|^2$. The equality holds only when $x = 0$.

If we now set $z = \int (f^*g) d\tau$, then $z^* = \int (g^*f) d\tau$ and we can conclude that

$$\left| \int (f^*g) d\tau \right|^2 \geq \left| \frac{1}{2} \left(\int (f^*g) d\tau - \int (g^*f) d\tau \right) \right|^2. \quad (7.38)$$

But

$$\begin{aligned} \int (f^*g) d\tau &= \int \left[(\hat{A} - \langle \hat{A} \rangle) \Psi \right]^* (\hat{B} - \langle \hat{B} \rangle) \Psi d\tau \\ &= \int \Psi^* (\hat{A} - \langle \hat{A} \rangle) (\hat{B} - \langle \hat{B} \rangle) \Psi d\tau \\ &= \int \Psi^* (\hat{A}\hat{B} - \langle \hat{A} \rangle \hat{B} - \langle \hat{B} \rangle \hat{A} + \langle \hat{A} \rangle \langle \hat{B} \rangle) \Psi d\tau \\ &= \int \Psi^* \hat{A}\hat{B}\Psi d\tau - \langle \hat{A} \rangle \int \Psi^* \hat{B}\Psi d\tau - \langle \hat{B} \rangle \int \Psi^* \hat{A}\Psi d\tau + \langle \hat{A} \rangle \langle \hat{B} \rangle \int \Psi^* \Psi d\tau, \end{aligned} \quad (7.39)$$

where we have once again made use of the Hermitian property of the operator $(\hat{A} - \langle \hat{A} \rangle)$. We know by now where this is going. Since $\int \Psi^* \hat{A}\hat{B}\Psi d\tau = \langle \hat{A}\hat{B} \rangle$, $\int \Psi^* \hat{A}\Psi d\tau = \langle \hat{A} \rangle$, $\int \Psi^* \hat{B}\Psi d\tau = \langle \hat{B} \rangle$, and $\int \Psi^* \Psi d\tau = 1$, we can simplify this last expression to

$$\int (f^*g) d\tau = \langle \hat{A}\hat{B} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle. \quad (7.40)$$

We can follow the same logic to deduce that

$$\int (g^*f) d\tau = \langle \hat{B}\hat{A} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle. \quad (7.41)$$

Inserting (7.40) and (7.41) into (7.38) gives

$$\left| \int (f^*g) d\tau \right|^2 \geq \left| \frac{1}{2} (\langle \hat{A}\hat{B} \rangle - \langle \hat{B}\hat{A} \rangle) \right|^2. \quad (7.42)$$

But $\langle \hat{A}\hat{B} \rangle - \langle \hat{B}\hat{A} \rangle$ is the expectation value of the commutator, $\langle [\hat{A}, \hat{B}] \rangle$.* Hence, from (7.35) we have

$$\sigma_A^2 \sigma_B^2 \geq \left| \int (f^*g) d\tau \right|^2 \geq \left| \frac{1}{2} \langle [\hat{A}, \hat{B}] \rangle \right|^2. \quad (7.43)$$

Step (4): The Position–Momentum Uncertainty Relation

This next step is almost trivial. Setting $\hat{A} = x$ (multiplication by x), and $\hat{B} = \hat{p}_x$, from Eq. (7.43) we have

$$\sigma_x^2 \sigma_{p_x}^2 \geq \left| \frac{1}{2} \langle [x, \hat{p}_x] \rangle \right|^2. \quad (7.44)$$

But from (7.4) we know that $[x, \hat{p}_x] = i\hbar$, which is a constant. This means $\langle [x, \hat{p}_x] \rangle = i\hbar$ and

$$\sigma_x^2 \sigma_{p_x}^2 \geq \left(\frac{1}{2} \hbar \right)^2. \quad (7.45)$$

Or

$$\sigma_x \sigma_{p_x} \geq \frac{1}{2} \hbar, \quad (7.46)$$

which is the position–momentum uncertainty relation, sometimes expressed rather more loosely in terms of vaguely defined ‘uncertainties’ Δx and Δp_x :

$$\Delta x \Delta p_x \geq \frac{1}{2} \hbar. \quad (7.47)$$

This relation implies that in a quantum-mechanical system, the position of an object (whether we think of it as a particle or a wave) and its momentum cannot both be established *simultaneously* with arbitrary precision. The word ‘simultaneously’ is very important. The uncertainty relation does not prevent us from observing the position with whatever precision we like. In fact, there is nothing in principle preventing us from making an observation in which the uncertainty in position is zero, $\Delta x = 0$. But we can see immediately from (7.47) that the uncertainty in momentum will then be infinite: $\Delta p_x \geq \frac{1}{2} \hbar / \Delta x$. Generally, the greater the precision in the observation of position (the

* Not sure? Just recall that $\langle [\hat{A}, \hat{B}] \rangle = \int \Psi^* [\hat{A}, \hat{B}] \Psi d\tau = \int \Psi^* (\hat{A}\hat{B} - \hat{B}\hat{A}) \Psi d\tau = \int \Psi^* \hat{A}\hat{B} \Psi d\tau - \int \Psi^* \hat{B}\hat{A} \Psi d\tau = \langle \hat{A}\hat{B} \rangle - \langle \hat{B}\hat{A} \rangle$.

smaller Δx), the greater the uncertainty in momentum, and vice versa. This is what Pauli really meant by things getting ‘fuzzy’.

Before we move on to Step (5), I think it’s helpful to make a couple of points. Firstly, apologies for the continued and rather tedious use of integrals in the derivation so far. Those readers already familiar with quantum mechanics will know that there’s a much neater, and simpler, notation we can use which eliminates the need to keep writing integrals over all coordinates, $d\tau$. But I’m holding off introducing this notation until Chapter 10, not only because this is consistent with history, but also because (as my mother always says) a reward is often more satisfying if you’ve had to wait patiently for it.

For my second point I want to question the logic behind Eq. (7.38). Why take only the modulus-square of the imaginary part of $|\int (f^*g) d\tau|^2$? Why not maintain equality and include the square of the real part as well? The result would be

$$\begin{aligned} \left| \int (f^*g) d\tau \right|^2 &= \left[\frac{1}{2} \left(\int (f^*g) d\tau + \int (g^*f) d\tau \right) \right]^2 \\ &\quad + \left[\frac{1}{2} \left(\int (f^*g) d\tau - \int (g^*f) d\tau \right) \right]^2. \end{aligned} \quad (7.48)$$

The rather obvious answer is that this gives us a little more than we might have bargained for. If we follow this through to the end using much the same logic as we used earlier, we arrive at a version of the relation that Schrödinger deduced in 1930,⁹

$$\sigma_A^2 \sigma_B^2 \geq \left[\frac{1}{2} \left(\langle \{\hat{A}, \hat{B}\} \rangle - 2\langle \hat{A} \rangle \langle \hat{B} \rangle \right) \right]^2 + \left| \frac{1}{2} \langle [\hat{A}, \hat{B}] \rangle \right|^2, \quad (7.49)$$

where $\{\hat{A}, \hat{B}\}$ is an *anti-commutator*: $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$. Schrödinger’s version of the general uncertainty relation is the square root of this result.

Equation (7.49) is a bit more complicated than Robertson’s version, (7.43). But although it looks a bit daunting the first term in (7.49) is actually the *covariance* of the operators, which we write as $\text{cov}(\hat{A}, \hat{B})$:

$$\text{cov}(\hat{A}, \hat{B}) = \langle (\Delta\hat{A}\hat{B}) \rangle = \frac{1}{2} \langle \hat{A}\hat{B} + \hat{B}\hat{A} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle. \quad (7.50)$$

We can see that if $\hat{A} = \hat{B}$, then the covariance $\langle (\Delta\hat{A}\hat{B}) \rangle$ becomes the variance $\langle (\Delta\hat{A})^2 \rangle$ and (7.50) reduces to (7.28). If the operators \hat{A} and \hat{B} are completely independent of each other, then their covariance will be zero and (7.49) reduces to (7.43). However, there are practical circumstances in which $\text{cov}(\hat{A}, \hat{B})$ may be greater than 0, in which case the Schrödinger uncertainty relation—the square root of (7.49)—is the more accurate.

Step (5): The Energy–Time Uncertainty Relation

Let’s now complete the picture by considering an equivalent uncertainty relation for energy and time. We first have to confront a rather obvious difficulty. As we’ve seen in Chapter 5, ‘energy’ is not in itself an operator, but is derived as an observable through the operation of the Hamiltonian operator, \hat{H} . Likewise, there is no operator for time in quantum mechanics—time, like space, is assumed to be continuously variable and forms a kind of ‘backdrop’ against which quantum events play out. In fact, the presumption of a background space-time ‘container’ is one of the major problems that must be overcome in any attempt to combine quantum mechanics and Einstein’s *general* theory of relativity, in which space-time emerges as a dynamical variable. If it can be done, the result would be a *quantum theory of gravity*, and today there are several approaches under active investigation, the two leading contenders being superstring theory and loop quantum gravity.¹⁰

Well, okay. But, in the position–momentum commutator, x appears simply as ‘multiplication by x ’, so why not introduce t as ‘multiplication by t ’? Go ahead—be my guest. But you will find that \hat{H} and t commute, $[\hat{H}, t] = 0$, so this can’t be the right way to think about the problem.

Here’s another thought. Equation (5.26) seems to suggest that we could introduce an ‘operator for time’, which we might define as $\hat{t} = i\hbar d/dt$, so that $\hat{t}\Psi = E\Psi$. But then you’ll quickly discover that $[\hat{H}, \hat{t}] = 0$, too. So this is not the right approach, either.

We’ll find it easier to come at this from a tangent. Consider a time-independent operator, which I’ll call $\hat{\Omega}$, with which we can form a commutator with the Hamiltonian operator, $[\hat{H}, \hat{\Omega}]$. For now, it doesn’t matter what $\hat{\Omega}$ represents but, as before, we need to assume that the wavefunction Ψ is a simultaneous eigenfunction of both \hat{H} and $\hat{\Omega}$. From Eq. (7.43), we have

$$\sigma_H^2 \sigma_\Omega^2 \geq \left| \frac{1}{2} \langle [\hat{H}, \hat{\Omega}] \rangle \right|^2 \quad (7.51)$$

The expectation value $\langle \hat{\Omega} \rangle$ is given by

$$\langle \hat{\Omega} \rangle = \int \Psi^* \hat{\Omega} \Psi \, d\tau. \quad (7.52)$$

Somehow, we need to introduce time into this picture, so let’s look closely at the *time dependence* of $\hat{\Omega}$ (note that although the operator $\hat{\Omega}$ is time-independent, this doesn’t mean that its expectation value doesn’t vary with time):

$$\begin{aligned} \frac{d}{dt} \langle \hat{\Omega} \rangle &= \frac{d}{dt} \left[\int \Psi^* \hat{\Omega} \Psi \, d\tau \right] \\ &= \int \left(\frac{\partial \Psi^*}{\partial t} \right) \hat{\Omega} \Psi \, d\tau + \int \Psi^* \left(\frac{\partial \hat{\Omega}}{\partial t} \right) \Psi \, d\tau + \int \Psi^* \hat{\Omega} \left(\frac{\partial \Psi}{\partial t} \right) \, d\tau. \end{aligned} \quad (7.53)$$

But $\hat{\Omega}$ is time-independent, so $\partial\hat{\Omega}/\partial t = 0$ and the middle term disappears. From (5.28), we have the time-dependent Schrödinger equation

$$\hat{H}\Psi = i\hbar\frac{\partial}{\partial t}\Psi. \quad (7.54)$$

Or $\partial\Psi/\partial t = \hat{H}\Psi/i\hbar$. Substituting this into (7.53) gives

$$\frac{d}{dt}\langle\hat{\Omega}\rangle = \int\left(\frac{\hat{H}\Psi}{i\hbar}\right)^*\hat{\Omega}\Psi d\tau + \int\Psi^*\hat{\Omega}\left(\frac{\hat{H}\Psi}{i\hbar}\right)d\tau. \quad (7.55)$$

Or

$$\frac{d}{dt}\langle\hat{\Omega}\rangle = -\frac{1}{i\hbar}\int(\hat{H}\Psi)^*\hat{\Omega}\Psi d\tau + \frac{1}{i\hbar}\int\Psi^*\hat{\Omega}(\hat{H}\Psi)d\tau. \quad (7.56)$$

The minus sign appears because the complex conjugate of $1/i\hbar$ is $1/(-i\hbar)$. We now make use of the Hermitian properties of \hat{H} and replace $1/-i$ with i . This allows us to write

$$\frac{d}{dt}\langle\hat{\Omega}\rangle = \frac{i}{\hbar}\int\Psi^*(\hat{H}\hat{\Omega})\Psi d\tau - \frac{i}{\hbar}\int\Psi^*(\hat{\Omega}\hat{H})\Psi d\tau. \quad (7.57)$$

Or

$$\frac{d}{dt}\langle\hat{\Omega}\rangle = \frac{i}{\hbar}\int\Psi^*(\hat{H}\hat{\Omega} - \hat{\Omega}\hat{H})\Psi d\tau = \frac{i}{\hbar}\langle[\hat{H}, \hat{\Omega}]\rangle. \quad (7.58)$$

We really should note the significance of (7.58) as we pass through. If the operator $\hat{\Omega}$ commutes with \hat{H} , then $[\hat{H}, \hat{\Omega}] = 0$, its expectation value is 0 and so $d\langle\hat{\Omega}\rangle/dt = 0$. In these circumstances the expectation value $\langle\hat{\Omega}\rangle$ is time-independent and *conserved*—it is said to be a *constant of the motion*.

I can't resist making a short detour here to consider the example of linear momentum. If we set $\hat{\Omega} = \hat{p}_x$, then the commutator becomes $[\hat{H}, \hat{p}_x] = \hat{H}\hat{p}_x - \hat{p}_x\hat{H}$. Let's evaluate this in stages. From (5.15) we have

$$\begin{aligned} \hat{H}\hat{p}_x\psi(x) &= \left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V\right)\left(-i\hbar\frac{d}{dx}\psi(x)\right) \\ &= \frac{i\hbar^3}{2m}\frac{d^3}{dx^3}\psi(x) - i\hbar V\frac{d}{dx}\psi(x). \end{aligned} \quad (7.59)$$

Likewise,

$$\begin{aligned}\hat{p}_x \hat{H} \psi(x) &= \left(-i\hbar \frac{d}{dx}\right) \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + V \psi(x)\right) \\ &= \frac{i\hbar^3}{2m} \frac{d^3}{dx^3} \psi(x) - i\hbar \frac{d}{dx} [V \psi(x)].\end{aligned}\quad (7.60)$$

And so, subtracting (7.60) from (7.59), we get

$$[\hat{H}, \hat{p}_x] \psi(x) = -i\hbar \left(V \frac{d}{dx} - \frac{d}{dx} V\right) \psi(x) = -i\hbar \left[V, \frac{d}{dx}\right] \psi(x). \quad (7.61)$$

Alternatively,

$$\begin{aligned}[\hat{H}, \hat{p}_x] \psi(x) &= -i\hbar \left(V \frac{d\psi(x)}{dx} - \psi(x) \frac{dV}{dx} - V \frac{d\psi(x)}{dx}\right) \\ &= -i\hbar \left(\frac{dV}{dx}\right) \psi(x),\end{aligned}\quad (7.62)$$

where we have once again made use of the product rule. Equation (7.62) can be written as an operator equation:

$$[\hat{H}, \hat{p}_x] = -i\hbar \left(\frac{dV}{dx}\right). \quad (7.63)$$

Clearly, $[\hat{H}, \hat{p}_x]$ is zero and linear momentum is conserved only if the potential energy V is uniform and does not change with distance, $dV/dx = 0$. In other words, since the force $F = -dV/dx$ —see Eq. (P.8)—this means that linear momentum is conserved only in the *absence* of a force. *This is Newton’s first law of motion!*

There’s more. From (7.58) and (7.63), we can deduce that

$$\frac{d}{dt} \langle \hat{p}_x \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{p}_x] \rangle = -\left\langle \frac{dV}{dx} \right\rangle = \langle F \rangle. \quad (7.64)$$

The rate of change of the expectation value of the linear momentum is equal to the expectation or mean value of the applied force. *This is Newton’s second law*, in which we have replaced the classical observables with the expectation values of their quantum mechanical operator equivalents. This is *Ehrenfest’s theorem*, named for Austrian physicist Paul Ehrenfest.

I think you’ll agree that this short detour was well worth it. But let’s now get back to the energy–time uncertainty relation. From (7.58) we have $\langle [\hat{H}, \hat{\Omega}] \rangle = -i\hbar d\langle \hat{\Omega} \rangle/dt$ (remember $1/i = -i$). If we now substitute this expression for $\langle [\hat{H}, \hat{\Omega}] \rangle$ into (7.51), we

get

$$\sigma_H^2 \sigma_\Omega^2 \geq \left(-\frac{\hbar}{2} \frac{d\langle \hat{\Omega} \rangle}{dt} \right)^2 \geq \frac{\hbar^2}{4} \left(\frac{d\langle \hat{\Omega} \rangle}{dt} \right)^2. \quad (7.65)$$

Taking the square root gives

$$\sigma_H \sigma_\Omega \geq \frac{1}{2} \hbar \left| \frac{d\langle \hat{\Omega} \rangle}{dt} \right|. \quad (7.66)$$

Now comes a bit of sleight of hand. We choose to *define* the ‘uncertainty’ in energy ΔE as the square root of the variance of the Hamiltonian operator, $\Delta E = \sigma_H$. We also *define* the ‘uncertainty’ in time as

$$\Delta t = \frac{\sigma_\Omega}{|d\langle \hat{\Omega} \rangle/dt|}. \quad (7.67)$$

Of course, these definitions are designed to yield

$$\Delta E \Delta t \geq \frac{1}{2} \hbar, \quad (7.68)$$

which is the energy–time uncertainty relation.

There are arguments that the energy–time uncertainty relation actually doesn’t exist except as a variation of the relation for position and momentum. Early attempts to derive the relation for energy and time from first principles were rather unsatisfactory. The derivation I’ve given here is based on a paper published in 1945 by Soviet physicists Leonid Mandelstam and Igor Tamm, and it very clearly specifies the interpretation of Δt as a time *interval*—actually the amount of time required for the expectation value of $\hat{\Omega}$ to change by one standard deviation.¹¹

A Fresh Round of Difficult Discussions

Having deduced the uncertainty principle, Heisenberg developed a number of hypothetical ‘thought’ experiments designed to illustrate how the principle might manifest itself. These were not necessarily meant to be taken seriously as proposals for real experiments, but were rather imaginary examples that built on the practical logic of the apparatus that would be required and its interactions.

To talk about the position and momentum of any object, he reasoned, requires a clear, operational definition in terms of some experiment designed to measure these quantities. To illustrate this, he recalled a conversation he had had as a student in Göttingen. Supposing we wished to measure the path of an electron—its position and velocity (or

momentum) as it passes through a cloud chamber or orbits a nucleus. The most direct way of doing this would be to follow the electron's motion using a microscope. Now the resolving power of an optical microscope increases with increasing frequency of the radiation, and so a hypothetical gamma-ray microscope would be required to 'see' an electron in this way. The gamma-ray photons bounce off the electron, and some are collected by a lens system and used to produce a magnified image.

But now, Heisenberg reasoned, we have a problem. Gamma rays consist of high-energy photons. Each time a gamma-ray photon bounces off an electron, the Compton effect suggests that the electron is given a severe jolt. This jolt means that the direction of motion and the momentum of the electron are changed in ways that are governed by quantum mechanics. And, as Born had argued, only the probability for scattering in certain directions with certain momenta can be calculated. Although we might be able to obtain a fix on the electron's instantaneous position, the sizeable interaction of the electron with the device we are using to measure its position means that we can say nothing at all about the electron's momentum. The less uncertainty in position, the greater the uncertainty in momentum.

We could use much lower energy photons in an attempt to avoid this problem and so measure the electron's momentum, but the use of lower energy (lower frequency or longer wavelength) photons would mean that we must then give up hope of determining the electron's position. The less uncertainty in momentum, the greater the uncertainty in position.

Heisenberg's basic premise is that when making measurements on quantum scales, we run up against a fundamental limit. These are the same scales of distance and energy at which the primary measurement process itself is happening. It is therefore not possible to make a measurement without *disturbing* the object under study in an essential, unpredictable, way. The discontinuity characteristic of quantum jumps dominates the process. At the quantum level, our techniques of measurement are simply too 'clumsy'. In this interpretation, the uncertainty principle places fundamental limits on what is *measurable*.

'Then Niels Bohr returned from his skiing holiday,' Heisenberg wrote, 'and we had a fresh round of difficult discussions.'¹²

Complementarity

Bohr's thinking, though not yet complete, had reached an important stage of maturity. He had concluded that the contradiction implied by the electron's wave-like and particle-like behaviours was more apparent than real. We reach for classical wave and particle concepts to describe the results of experiments because these are the only kinds of concepts with which we are familiar from our experiences as human beings living in a classical world. Whatever the 'real' nature of the electron, the behaviour it exhibits is conditioned by the kinds of experiments we choose to perform. These, by definition, are experiments requiring apparatus of 'classical' dimensions, resulting in effects substantial enough to be observed and recorded in the laboratory, perhaps in the form of an exposed

photographic plate, or the deflection of a pointer in a voltmeter, or the observation of a track in a cloud chamber.

So, we conclude that in this experiment the electron is a wave. In another kind of experiment we conclude that the electron is a particle. These experiments are mutually exclusive. We cannot conceive an experiment to demonstrate both types of behaviour simultaneously, not because we lack the ingenuity, but because such an experiment is simply *inconceivable*.

We can ask questions concerning the electron's wave-like properties and we can ask mutually exclusive questions concerning the electron's particle-like properties, but we cannot ask what the electron *really is*. Bohr declared that these very different behaviours are not contradictory; they are instead *complementary*.

Heisenberg's uncertainty principle appears entirely consistent with Bohr's reasoning. Indeed, Bohr may have immediately grasped the significance of the uncertainty relations in terms of complementary classical wave and particle concepts. But as he read through Heisenberg's paper, which had by now been submitted for publication, Bohr grew increasingly dismayed. Although the end result was compatible, the logic that Heisenberg had set forth in his paper betrayed a startlingly different philosophy.

In grasping for a purely particulate interpretation (thereby avoiding all reference to the wave concepts associated with his arch-rival, Schrödinger), Heisenberg had traced the origin of uncertainty to the Compton effect, to an essential 'clumsiness' resulting from the substantial, discontinuous interaction between the electron and the gamma-ray photon being used to detect it. But, Bohr now pointed out, in principle the Compton effect gives rise to a precisely calculable recoil and is, in any case, applicable only to 'free' electrons (i.e. electrons that are not bound in an orbit around an atomic nucleus).

The origin of the uncertainty, Bohr argued, should rather be traced to the wave nature of the gamma rays used to probe the properties of the electron. The resolving power of any microscope is limited by the effects of diffraction in the lens aperture. This diffraction results in a blurring of the image; an inability to distinguish objects that are closer than the minimum resolvable distance. Although the resolution increases as shorter and shorter wavelengths are used (thus requiring a microscope based on gamma rays to resolve distances approaching the dimensions of an electron, as Heisenberg had assumed), the simple fact that the aperture must be of finite dimensions means that there remains a fundamental limit on the resolving power of the device. This loss of precision represents a fundamental uncertainty.*

According to Bohr, the uncertainty relations place a fundamental limit not on what is measurable, but on what is in principle *knowable*.

Bohr insisted that Heisenberg's paper—by now in press—should be withdrawn. Heisenberg stubbornly refused to yield. Their conflict was 'very disagreeable', Heisen-

* Heisenberg had almost failed to secure his doctorate at the University of Munich because he had been unable to derive expressions for the resolving power of a microscope, incurring the wrath of his examiner Wilhelm Wien, who had covered all the required background in his lectures. Heisenberg had been mortified by this experience, as had Arnold Sommerfeld, his thesis advisor.

berg later admitted. ‘I remember that it ended with my breaking out in tears because I just couldn’t stand this pressure from Bohr.’¹³

At the suggestion of Oskar Klein, a young Swedish physicist recently arrived in Copenhagen on a visiting fellowship, Heisenberg agreed to add a note in the proof of his paper. The note corrected the misinterpretation of the gamma-ray microscope experiment, and acknowledged a debt of gratitude to Bohr: ‘I owe great thanks to Professor Bohr for sharing with me at an early stage the results of these more recent investigations of his—to appear soon in a paper on the conceptual structure of quantum theory—and for discussing them with me.’¹⁴

NOTES

1. Erwin Schrödinger, quotations based on a reconstruction of the debate by Werner Heisenberg, *Physics and Beyond: Memories of a Life in Science*, Allen & Unwin, London, 1971, pp. 73–5.
2. Niels Bohr, quotation based on a reconstruction of the debate by Heisenberg, *ibid.* 76.
3. Max Born, ‘Zur Quantenmechanik der Stossvorgänge’, *Zeitschrift für Physik*, 37 (1926), 863–7, English translation reproduced in John Archibald Wheeler and Wojciech Hubert Zurek (eds), *Quantum Theory and Measurement*, Princeton University Press, Princeton, NJ, 1983. This quote appears on p. 52.
4. Werner Heisenberg, *Physics and Philosophy: The Revolution in Modern Science*, Penguin, London, 1989 (first published 1958), p. 30.
5. Wolfgang Pauli, letter to Werner Heisenberg, 19 October 1926. Quoted in Charles P. Enz, *No Time to be Brief: A Scientific Biography of Wolfgang Pauli*, Oxford University Press, Oxford, 2002, p. 141.
6. Werner Heisenberg, ‘Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik’, *Zeitschrift für Physik*, 43 (1927), 172–98. English translation reproduced in Wheeler and Zurek (eds), *Quantum Theory and Measurement*. This quote appears on p. 63.
7. Heisenberg, *Physics and Beyond*, p. 78.
8. H. P. Robertson, ‘The Uncertainty Principle’, *Physical Review*, 34 (1929), 163–4.
9. E. Schrödinger, ‘Zum Heisenbergschen Unschärfepinzip’, *Proceedings of the Prussian Academy of Sciences, Physical-Mathematical Section*, 14 (1930), 296–303. An English translation is available in A. Agelow and M. Batoni, *Bulgarian Journal of Physics*, 26 (1999), 193–203, arXiv:quant-ph/9903100v3, 23 May 2008.
10. For a little more background, see Jim Baggott, *Quantum Space: Loop Quantum Gravity and the Search for the Structure of Space, Time, and the Universe*, Oxford University Press, Oxford, 2018.
11. L. Mandelstam and Ig. Tamm, ‘The Uncertainty Relation between Energy and Time in Nonrelativistic Quantum Mechanics’, *Journal of Physics (USSR)* 9 (1945), 249–254.
12. Heisenberg, *Physics and Beyond*, p. 79.
13. Werner Heisenberg, interview by Thomas Kuhn, 25 February 1963, *Archive for the History of Quantum Physics*. Quoted in Max Jammer, *The Philosophy of Quantum Mechanics*, Wiley, New York, 1974, p. 65.
14. Heisenberg, ‘Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik’. English translation reproduced in Wheeler and Zurek (eds), *Quantum Theory and Measurement*. This quote appears on p. 84.

FURTHER READING

- Cassidy, David C., *Uncertainty: The Life and Science of Werner Heisenberg*, W. H. Freeman, New York, 1992, Chapter 12.
- Cushing, James T., *Philosophical Concepts in Physics: The Historical Relation between Philosophy and Scientific Theories*, Cambridge University Press, Cambridge, UK, 1998, pp. 299–301.
- Gamow, George, *Thirty Years That Shook Physics*, Dover, New York, 1966, Chapter 5.
- Griffiths, David J., *Introduction to Quantum Mechanics*, 2nd edn, Cambridge University Press, Cambridge, UK, 2017, Chapter 3.
- Heisenberg, Werner, *Physics and Beyond: Memories of a Life in Science*, Allen & Unwin, London, 1971.
- Heisenberg, Werner, *Physics and Philosophy: The Revolution in Modern Science*, Penguin, London, 1989.
- Jammer, Max, *The Philosophy of Quantum Mechanics*, Wiley, New York, 1974, Chapter 3.
- Moore, Walter, *Schrödinger: Life and Thought*, Cambridge University Press, Cambridge, UK, 1989, pp. 226–9.

Heisenberg's Derivation of the Pauli Exclusion Principle

The Stability of Matter and the Periodic Table

We now have many of the foundational equations of quantum theory to hand, but we're still missing a fundamentally important ingredient—electron spin. The phenomenon of spin wasn't being ignored by those pioneers who assembled the theoretical structure piece by piece in the 1920s, but it wasn't at all clear from the experimental evidence just what spin *is* and how it was supposed to fit into the grand scheme of things. In fact, gaining clarity would require bringing together *two* new pieces of the puzzle—the fact that the electron possesses some kind of curious 'inner rotation' and the establishment of a version of the Schrödinger equation that conforms to the demands of Einstein's special theory of relativity. Out of all this would come not only an accommodation of the phenomenon of electron spin, but a prediction of the existence of antimatter.

We will go on to consider Paul Dirac's derivation of the relativistic wave equation in Chapter 9. Here, I'd like to explore the emergence of spin, the extraordinary implications for the structure of matter that were worked out by Pauli in 1925, and a derivation of the Pauli exclusion principle based on the notion of quantum indistinguishability, presented by Heisenberg a year later.

The Normal and Anomalous Zeeman Effects

To do this we need to wind the clock back a little. Before Schrödinger's breakthrough in late 1925–6, it was possible to 'explain' many features of the spectra of atoms using the 'old' quantum mechanics of Bohr and Sommerfeld, which had been created by shoehorning quantum rules into an essentially classical mechanical structure. Much of this explanation rested on a set of selection rules, based on the (by now) familiar quantum numbers n , l , and m_l which, as Schrödinger would later demonstrate, emerged entirely naturally from his wave mechanics of the hydrogen atom.

Now, angular momentum is such a fundamental topic that it typically occupies entire chapters in textbooks on quantum mechanics.¹ Indeed, it's possible to find whole books

devoted to the subject.* It's actually a rather remarkable phenomenon, and demonstrates how much of the physics that we need can be deduced from the mathematical properties of the operators for angular momentum—and especially from consideration of the commutators—without needing to solve the Schrödinger equation at all. Although it's not central to our discussion here, rather than refer you to other books I've included a very brief summary of the treatment of orbital angular momentum in Appendix 6.

Suffice to say that in a multi-electron atom, we need to take account of the way in which the orbital angular momenta and the orbital magnetic moments of the electrons combine together, governed by total quantum numbers L and M_L (we capitalize these for multi-electron atoms).

Despite the success with the wave mechanical description of the hydrogen atom, it quickly became apparent that the spectrum of the simplest multi-electron atom—helium—could not be so readily explained. And the spectra of certain other types of atoms, such as sodium and the atoms of rare-earth elements such as cerium, showed 'anomalous' splitting when placed in a magnetic field that was, quite simply, baffling.

In the 'normal' Zeeman effect, spectral lines are split in the presence of a magnetic field, the extent of the splitting determined by the magnetic field strength. In a two-electron system, we deduce the total orbital angular momentum quantum number L from the *Clebsch–Gordan series* (named for German mathematicians Alfred Clebsch and Paul Gordan)

$$L = (l_1 + l_2), (l_1 + l_2 - 1), (l_1 + l_2 - 2), \dots, |l_1 - l_2|, \quad (8.1)$$

where l_1 and l_2 are the orbital angular momentum quantum numbers of the two electrons considered separately. In a situation where electron 1 is in an s orbital ($l_1 = 0$) and electron 2 is in a p orbital ($l_2 = 1$), then $L = 1$ and there are three values of M_L , corresponding to a 'multiplicity' ($2L + 1$) of 3: $M_L = -1, 0,$ and $+1$. These give rise to three lines in the magnetic spectrum. For $L = 2$, the multiplicity is 5 (see the discussion of quantum numerology in Chapter 5). According to these rules, as L is always an integer, the multiplicity is obviously always going to be an odd number.

But in the 'anomalous' Zeeman effect, the number of lines in a magnetic spectrum is actually *even*. This is only possible if the quantum number in the expression for the multiplicity is *half-integral*. There was simply no precedent for this.

In an attempt to classify the nature and magnitude of the splitting of the spectral lines of multi-electron atoms, in 1920 Sommerfeld had introduced a fourth quantum number, which he called the 'inner quantum number', together with a new selection rule. Where the first three quantum numbers had evolved from reference to classical conceptual models of the inner workings of the atom, this fourth quantum number was entirely *ad hoc*. Sommerfeld simply assumed that the motions of multi-electron atoms are complex and characterized by some kind of 'hidden rotation'. In 1923 the German

* See, for example, Richard N. Zare, *Angular Momentum: Understanding Spatial Aspects in Chemistry and Physics*, Wiley, New York, 1988.

spectroscopist Alfred Landé suggested that for atoms with a many-electron 'core' and a single outlying, 'valence' electron, the hidden rotation should be associated with the core electrons.

Landé's approach was relatively successful, but it was still all quite confusing and barely comprehensible. To add insult to injury, a year earlier German physicists Otto Stern and Walther Gerlach reported the effects of applying a magnetic field to a beam of silver atoms, produced by effusion from a heated oven. Stern had been looking for evidence for the kind of 'space quantization' predicted by the Bohr–Sommerfeld theory and when they found that the beam was indeed split into two components by the magnetic field, they thought they had found it. Others, including Einstein, were convinced they had not, and were greatly puzzled by these results.

The young Austrian physicist Wolfgang Pauli had begun to wrestle with the problem of the anomalous Zeeman effect during a year spent at Bohr's Institute for Theoretical Physics in Copenhagen in 1923–4. It caused him considerable difficulties and regular periods of anguish and despair. When stopped in the street and asked why he looked so unhappy, he replied: 'How can one look happy when he is thinking about the anomalous Zeeman effect?'² Whilst he was able to improve somewhat on the proposals of Sommerfeld and Landé, he disliked the nature of the theorizing that was involved. He was looking for something more fundamental.

Back in Hamburg towards the end of 1924, Pauli's attention was drawn to the work of Cambridge physicist Edmund Stoner, described in the preface to the fourth edition of Sommerfeld's book *Atombau und Spektrallinien*. In Stoner's original paper, published in October 1924 in the British journal *Philosophical Magazine*, he had set out a scheme describing the relationship between the quantum numbers and the idea of electron 'shells', surrounding the nucleus and imagined to nest one inside the other like a Russian *matryoshka* doll. The energy of each shell is determined by the principal quantum number, n . The number of possible states or 'orbitals' within each shell is then determined by the values of l and m_l that each individual electron can take for a given value of n .

As we saw in Chapter 5, these rules dictate that the number of electron orbitals increases as n^2 . But the pattern reflected in the periodic table of the elements tells a slightly different story. Walther Kossel had earlier argued that the striking stability and inertness of the noble gases (such as helium, neon, argon, and krypton) could be understood in terms of Bohr's atomic theory if these atoms were assumed to have filled, or 'closed', shells. The periodic table could then be understood as a progression of occupancy of the electron shells, forming a pattern in which first 2 (hydrogen, helium), then 8 (lithium through to neon), then another 8 (sodium to argon), then 18 electrons (potassium to krypton) are added in sequence until each successive shell is filled, or closed (see Fig. 8.1).

Stoner had gone one step further in his prescription. Instead of assigning a single electron to each orbit he had chosen to assign two: 'In the classification adopted, the remarkable feature emerges that the number of electrons in each completed level is equal to double the sum of the inner quantum numbers as assigned.'³ In n^2 orbits, Stoner suggested, there should be $2n^2$ electrons. For $n = 1$ there is only one orbit, implying

occupancy up to a total of 2 electrons; for $n = 2$ there are 4 orbits, implying up to 8 electrons; for $n = 3$ there are 9 orbits and up to 18 electrons.*

Pauli's Exclusion Principle

Pauli put two and two together. He believed that Landé's model in which a 'hidden rotation' was ascribed to the core of electrons in a multi-electron atom was wrong. And yet the model appeared to work quite well. Stoner was suggesting a doubling of the electron count ascribed to each orbit, and hence each shell. The answer, Pauli reasoned, was to ascribe the fourth quantum number not to the electron core, but to *each individual electron*. He was led to the inspired conclusion that the electron must have a curious, nonclassical 'two-valuedness' (*Zweideutigkeit*, in German) characterized by a quantum number of $\frac{1}{2}$.

There was more. The shell structure of atoms and the periodic table of the elements implied that each orbit could accommodate two, and only two, electrons. Pauli wrote:⁴

There can never be two or more equivalent electrons in the atom, for which, in strong fields, the values of all quantum numbers . . . coincide. If one electron is present in the atom, for which these quantum numbers (in the external field) have definite values, then this state is 'occupied'.

This is Pauli's *exclusion principle*. In essence it says that no electron in an atom can have the same set of four quantum numbers. Today we give the fourth quantum number the symbol s , fixed at $s = \frac{1}{2}$. We apply the same rules as before. The multiplicity is given by $2s+1$, suggesting just two components which, like the quantum number m_l , we determine from the series $-s, -(s-1), \dots, 0, \dots, (s-1), s$. This gives two quantum numbers that we characterize as $m_s = -\frac{1}{2}$ and $m_s = +\frac{1}{2}$. This means that two electrons (considered to be completely independent of each other) can enter the 1s orbital characterized by $n = 0, l = 0, m_l = 0$ provided that they take different values of m_s . As there are only two possibilities, the argument goes, the orbital can accommodate only two electrons. I'm sure it won't surprise you to learn that it's not quite as simple as this, and we'll see how this plays out later in this chapter, when we consider what happens when the spins of the two electrons couple or combine together.

Pauli was unable to provide any formal explanation for this rule and he certainly wasn't able to derive it from first principles. He had no alternative but to argue that it seemed to

* To make explicit the connection with the 2, 8, 8, 18, . . . pattern of the periodic table it is necessary to jump ahead to our current understanding of the atomic orbitals and their relative energies. The orbital with $n = 1, l = 0$ is spherical and as we've seen this is the 1s orbital. This can accommodate up to 2 electrons, accounting for hydrogen and helium. For $n = 2$ the possible orbitals are 2s and three different 2p orbitals, accommodating up to a total of 8 electrons (lithium to neon). For $n = 3$ the orbitals are 3s, 3p (up to 8 electrons—sodium to argon) and five different 3d orbitals (10 electrons). However, the 4s orbital actually lies somewhat lower in energy than 3d and is filled first. Therefore, the combination 4s, 3d, and 4p (accommodating up to 18 electrons in total) together account for the next row of the periodic table, from potassium to krypton.

‘offer itself automatically in a natural way’. When in December 1924 he sent a copy of the manuscript of a paper describing his results to Bohr and Heisenberg in Copenhagen, he received an exuberant, though barbed, response from Heisenberg. Heisenberg called it a ‘swindle’.⁵

Although its origin in the as-yet undefined quantum mechanics of the atom remained vague (to say the least), Pauli’s exclusion principle accounts for an important feature of multi-electron atoms—the fact that they exist at all. There had been nothing in the earlier atomic theory to provide a reason why all the electrons in a multi-electron atom should not simply collapse down into the lowest-energy orbital.

For neutral atoms, increasing the number of electrons implies an increase in the total positive charge of the nucleus. The greater the nuclear charge, the smaller the radius of the innermost orbital, as the electrons are pulled more tightly towards the nucleus. Now we would expect that the repulsion between increasingly closely packed electrons would tend to resist collapse of the atom into ever-smaller orbitals and hence ever-smaller volumes, but it is relatively straightforward to show that electron–electron repulsion can’t prevent heavier atoms from shrinking dramatically in size as the charge of the central nucleus is increased. The repulsion between neighbouring electrons simply isn’t strong enough to overcome the force of attraction. Atomic volumes, easily calculated from atomic weights and densities of the elements, follow a complex variation with atomic charge but they definitely don’t systematically shrink with increasing charge.

By preventing the electrons from collapsing or condensing into the lowest-energy orbital, the exclusion principle allows complex multi-electron atoms to exist in the pattern described by the periodic table. It enables the existence of a marvellous variety of elements, the multitude of possible chemical substances produced by combining these, and hence all material substance, living and non-living. This was a fantastic achievement.

But, still: why only *two* electrons per orbital?

The Self-rotating Electron

Perhaps, argued a young American physicist named Ralph de Laer Kronig, this is because the electron’s ‘two-valuedness’ is associated with *self-rotation*. Sommerfeld had ascribed the fourth quantum number to a ‘hidden rotation’. But what if this rotation were, in fact, a real self-rotation of individual electrons? If the electron rotates about its own axis, in much the same way that the Earth rotates on its axis as it orbits the Sun, then this would generate a small, local magnetic field. The electron in an atom would behave like a tiny bar magnet. It would possess a magnetic moment that can become aligned with or aligned against the lines of force of an applied external magnetic field, giving two states of different energy which would appear as a splitting of spectral lines.

In the absence of this splitting, there is only one state of the electron and hence only one line in the spectrum. Kronig calculated that the angular momentum of electron self-rotation was required to have the fixed value of $\frac{1}{2}\hbar$. In addition, he was able to show that the ratio of the magnetic moment and the angular momentum due to self-rotation, a characteristic factor known as the Landé ‘*g*-factor’ for the electron, had to have the

value 2. This was rather curious, as the ratio of the electron orbital magnetic moment to the orbital angular momentum is 1, as we would expect from classical mechanics (don't worry—we'll be returning to this in Chapter 9).

There was a further problem. From Kronig's hastily derived expressions he calculated a splitting of spectral lines that was a factor of 2 larger than the splitting observed experimentally. This was not connected with the assumption of $g = 2$, as this was needed to explain other experimental observations. Electron self-rotation was able to resolve two out of three of the problems with Landé's core model, but could not yet resolve them all.

Kronig met with Pauli in December 1924, and he explained his ideas. Pauli remarked: '*Das ist ja ein ganz witziger Einfall*' (this is indeed quite a witty idea), but he did not believe that the suggestion had any connection with reality.⁶

Pauli, noted as much for his biting wit as for his talents as a theoretical physicist, completely dismissed Kronig's suggestion. Kronig discussed it further with Bohr and Heisenberg, who were similarly dismissive. He was in any case troubled by the factor of 2 discrepancy between prediction and experiment, and concerned also that the equator of a spinning sphere of charged matter would be required to move ten times faster than the speed of light, forbidden by Einstein's special theory of relativity. He subsequently dropped the idea.

Ten months can be a long time in physics. When two young Dutch physicists Samuel Goudsmit and George Uhlenbeck, based in Leiden in the Netherlands, independently reached the same conclusion, the climate in the physics community was to prove more temperate, and more willing to forgive. When Goudsmit first told Uhlenbeck about the exclusion principle, Uhlenbeck was quick to see the connection. 'But don't you see what this implies?' he said, 'It means that there is a fourth degree of freedom for the electron. It means that the electron has a spin, that it rotates.' Goudsmit was nonplussed: 'What is a degree of freedom?' he asked.⁷

Their research supervisor at Leiden, Austrian physicist Paul Ehrenfest, thought it was a nice idea, but probably wrong. However, as they had yet to establish any kind of reputation as physicists, he figured they had nothing to lose. They summarized their arguments in favour of electron self-rotation in a paper that Ehrenfest agreed to submit to the journal *Naturwissenschaften*. But Uhlenbeck talked to Hendrik Lorentz, who advised him that their proposal was impossible in classical electron theory. Fearing they had made a significant error, Uhlenbeck asked Ehrenfest not to submit the paper. It was too late. Ehrenfest had already sent it off.

The paper was published in November 1925. It initially provoked the same concerns that Kronig's proposal, now almost forgotten, had attracted. Travelling to Leiden in December 1925, Bohr was met at the railway station by Pauli and Stern and asked for his opinion about the Dutch proposal. Bohr may have said that he found it 'very interesting', which was Bohr-speak implying that it was probably wrong. On arrival in Leiden he was met by Einstein and Ehrenfest, who asked him the same question. When Einstein went on to explain how some of his objections to the proposal could be overcome, Bohr began to have a change of heart.

From Leiden Bohr travelled to Göttingen, where he was met by Heisenberg and Pascual Jordan, who asked the question again. This time Bohr was enthusiastic, although

Heisenberg vaguely recalled having heard a similar proposal some time before. On his way back to Copenhagen Bohr's train stopped in Berlin, where he was met by Pauli, who had travelled from Hamburg expressly to ask Bohr once again what he thought about electron self-rotation. Bohr now said it represented a great advance. Pauli, still unable to get past the fact that the classical picture of a spinning bit of charged matter made absolutely no sense in the context of atomic physics, called it 'a new Copenhagen heresy'.⁸

Bohr became a strong advocate of electron self-rotation and may have been the first to use the term 'electron spin'. In a postscript to a follow-up paper by Uhlenbeck and Goudsmit published in *Nature* early in 1926, Bohr bemoaned the problems that beset the interpretation of atomic spectra and declared that⁹

The situation seems, however, to be somewhat altered by the introduction of the hypothesis of the spinning electron which, in spite of the incompleteness of the conclusions that can be derived from models, promises to be a very welcome supplement to our ideas of atomic structure.

Electron spin is a term that has stuck, despite the fact that its meaning in quantum mechanics is considerably far removed from its classical interpretation. Today we think of the two possible orientations of the electron spin corresponding to $m_s = -\frac{1}{2}$ and $m_s = +\frac{1}{2}$ as 'spin down', \downarrow , and 'spin up', \uparrow .

As the idea gained wider acceptance, it became clear that Pauli (and Bohr, and Heisenberg) had discouraged Kronig from developing his own proposal further and from therefore becoming the 'discoverer' of electron spin.* Kronig tended to play down the matter, but could not help feeling some bitterness: 'I should not have mentioned the matter at all,' he wrote later to Bohr, 'if it were not to take a fling at the physicists of the preaching variety, who are always so damned sure of, and inflated with, the correctness of their own opinion.'¹⁰

The mysterious factor of 2 discrepancy between predicted and observed splitting of spectral lines was resolved satisfactorily. English physicist Llewellyn Hilleth Thomas subsequently showed that recasting the problem in the proper rest frame of the electron changed the expression for the splitting, replacing the appearance of g in the expression with $g - 1$. Assuming $g = 2$ effectively reduces the predicted splitting by half.

It was English theorist Paul Dirac who subsequently suggested that if the electron can be considered to possess two possible 'spin' orientations then this, perhaps, explains why each atomic orbital can accommodate only two electrons. The two electrons must be of opposite spin to 'fit' in the same orbital. An orbital can hold a maximum of two electrons provided their spins are *paired*.

The problem of the anomalous Zeeman effect could now be resolved by coupling together the orbital and spin angular momenta of the electron. In the limit of weak

* A verse penned some time later summarized the situation: '*Der Kronig hätt' den Spin entdeckt, hätt' Pauli ihn nicht abgeschreckt.*' (Kronig would have discovered the spin if Pauli had not discouraged him.) See Charles P. Enz, *No Time to be Brief: A Scientific Biography of Wolfgang Pauli*, Oxford University Press, Oxford, 2002, p. 117.

coupling between the electron spin and orbital angular momenta, we deduce the total orbital angular momentum quantum number L from (8.1), and the total spin quantum number S from the equivalent series:

$$S = (s_1 + s_2), (s_1 + s_2 - 1), (s_1 + s_2 - 2), \dots, |s_1 - s_2|. \quad (8.2)$$

In the limit of weak magnetic field strength, we can combine L and S to produce a total angular momentum quantum number \mathcal{J} :

$$\mathcal{J} = (L + S), (L + S - 1), (L + S - 2), \dots, |L - S|. \quad (8.3)$$

This particular recipe for combining the total orbital and spin angular momentum quantum numbers is called L - S or Russell–Saunders coupling, named for American astronomer Henry Norris Russell and Frederick Saunders.* There is no net contribution to \mathcal{J} from the closed shells of multi-electron atoms (i.e. for which every orbital within the shell is filled with two spin-paired electrons), so we just need to consider the orbital and spin angular momenta of any outlying ‘valence’ electrons.

We see immediately from Eqs. (8.1), (8.2), and (8.3) how the total angular momentum quantum number \mathcal{J} can acquire half-integral values. Atoms with a single unpaired outer electron will contribute a total spin angular momentum characterized by $S = \frac{1}{2}$. So, a state with $L=1$ and $S=\frac{1}{2}$ has, from (8.3), values of $\mathcal{J} = \frac{3}{2}, \frac{1}{2}$. In a magnetic field, the quantum state corresponding to $\mathcal{J} = \frac{1}{2}$ splits into two, giving two lines in the spectrum, corresponding to $M_{\mathcal{J}} = \pm \frac{1}{2}$. The state corresponding to $\mathcal{J} = \frac{3}{2}$ splits into four, corresponding to $M_{\mathcal{J}} = \pm \frac{1}{2}, \pm \frac{3}{2}$. In both cases, the splitting is even, not odd.

This was great progress, but there were still many, many puzzles. A classical spinning object is not constrained in principle to only two positions of alignment of its magnetic moment. It was reasoned that this restriction must be somehow due to the quantum nature of the electron.

It was just not clear how.

Of course, none of the above explains *why* the introduction of electron spin should automatically restrict each orbital to two electrons. After all, what’s wrong with three electrons all in one orbit, with two in spin up orientations and one in a spin down orientation? The answer to this question would come in 1926—more than a year after Pauli had articulated the exclusion principle—in a couple of papers published by Heisenberg. Although Heisenberg approached the problem using matrix mechanics, he used arguments derived from the spectrum of helium to trace the exclusion principle back to the *symmetry* properties of the wavefunctions of multi-electron atoms.† Although

* Russell is also well known for his work with Ejnar Hertzsprung on the relationship between stellar luminosity and surface temperature, summarized in a Hertzsprung–Russell diagram.

† Yes, Heisenberg was prepared to talk about the properties of the *wavefunctions* of the atomic states of helium but applied them in constructing matrix elements for use in matrix mechanics.

I will here take a different (and more conventional) approach, Heisenberg's logic forms the basis for the following derivation.

The Ingredients

1. The Schrödinger equation for the helium atom.
2. The orbital approximation.
3. The wavefunctions of multiple, independent Hamiltonian operators.
4. Quantum indistinguishability and spatial wavefunctions.
5. The electric dipole operator and selection rules for electric dipole transitions.
6. Electron spin wavefunctions.
7. The spectrum of helium.

The Recipe

We begin in Step (1) by writing down the Schrödinger equation for a two-electron system, for which the helium atom serves as a useful example. As we will see, the presence of a term accounting for electron–electron repulsion means that this equation is impossible to solve analytically. Undaunted, we simply neglect this term, allowing us to approximate the Schrödinger equation for helium as the sum of the Hamiltonians for each ‘hydrogen-like’ electron, considered independently of each other (this is called the orbital approximation). As this is the first time we’ve had to deal with a system like this, in Step (2) we consider the most appropriate form for the helium atom wavefunction.

In Step (3) we acknowledge that the electrons cannot be distinguished one from another, and work out what this means for our expression for the spatial component of multi-electron wavefunctions. We introduce the electric dipole operator in Step (4) and deduce some simple selection rules for electric dipole transitions based on the symmetry properties of the spatial wavefunctions. In Step (5) we introduce the notion of electron spin wavefunctions and combine these with the spatial components to give expressions for a set of possible total wavefunctions.

We pull all these threads together in Step (6) by studying the spectrum of helium. Simply by observing which transitions are allowed and which are forbidden, we can discover expressions for the wavefunctions of the states and deduce the generalized *Pauli principle*. The exclusion principle then follows as a direct consequence.

Step (1): The Schrödinger Equation for Helium

Let's label the electrons using the subscripts 1 and 2. The full time-dependent wavefunction for the two-electron system is then $\Psi_{12}(r_1, r_2, t)$. Again, we suppose that we can separate the spatial and time variables:

$$\Psi_{12}(r_1, r_2, t) = \psi_{12}(r_1, r_2)\phi(t) = \psi_{12}(r_1, r_2)e^{-iE_{12}t/\hbar}, \quad (8.4)$$

where E_{12} is the energy of the two-electron system. Once again, we focus our attention on the spatial functions $\psi_{12}(r_1, r_2)$. In devising the Hamiltonian for the system, we suppose that we can ascribe kinetic energy terms to each electron separately (\hat{T}_1 and \hat{T}_2) where, as usual, $\hat{T}_1 = -\hbar^2\nabla_1^2/2m_e$, etc. Similarly, we ascribe a Coulomb potential separately to each electron (V_1 and V_2), and include a potential term to account for their mutual repulsion, V_{12} . This gives

$$\hat{H}_{12}\psi_{12}(r_1, r_2) = \left[\frac{-\hbar^2}{2m_e} (\nabla_1^2 + \nabla_2^2) - \frac{2e^2}{4\pi\epsilon_0 r_1} - \frac{2e^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 r_{12}} \right] \psi_{12}(r_1, r_2) \quad (8.5)$$

In (8.5), r_1 is the distance of electron 1 measured from the helium nucleus, r_2 the distance of electron 2. r_{12} is then the distance between the electrons. Note that the larger nuclear charge ($2e, Z = 2$) is reflected in the Coulomb potentials for each individual electron, which are twice the size of the potential for the electron in a hydrogen atom, and that the repulsive potential term is positive. Finally, you should note that we're continuing to use the electron mass m_e instead of the nucleus–electron reduced mass.

If we try to pursue the same strategy we used in Chapter 5, which involved transforming to a system of spherical polar coordinates, we'll quickly discover that we now have a partial differential equation with six independent variables, r_1, θ_1, ϕ_1 and r_2, θ_2, ϕ_2 . Good luck with that.

One approach to finding approximate solutions to Eq. (8.5) is to treat the repulsion term V_{12} as a 'perturbation' to an otherwise solvable equation involving two independent Hamiltonian operators, $\hat{H}_1 = \hat{T}_1 + V_1$ and $\hat{H}_2 = \hat{T}_2 + V_2$. However, although perturbation theory is a staple of any introductory course on quantum mechanics, it is beyond the scope of this book.

As it happens, I'm really not all that interested in solving Eq. (8.5), even approximately. Its purpose here is to serve as a 'sandpit' in which we can play around with aspects of the quantum mechanics of a two-electron system, allowing us to deduce some important principles. So let's simplify things by dropping the repulsion term altogether. This *orbital approximation* allows us to disentangle the two electrons and treat them entirely separately:

$$\begin{aligned} \hat{H}_{12}\psi_{12}(r_1, r_2) &\approx \hat{H}_1\psi_{12}(r_1, r_2) + \hat{H}_2\psi_{12}(r_1, r_2) \\ &\approx \left(\frac{-\hbar^2}{2m_e} \nabla_1^2 - \frac{2e^2}{4\pi\epsilon_0 r_1} \right) \psi_{12}(r_1, r_2) + \left(\frac{-\hbar^2}{2m_e} \nabla_2^2 - \frac{2e^2}{4\pi\epsilon_0 r_2} \right) \psi_{12}(r_1, r_2). \end{aligned} \quad (8.6)$$

The two Hamiltonian operators \hat{H}_1 and \hat{H}_2 are obviously hydrogen-like (look back at Eq. (5.17)), and as we already have solutions to hand from Chapter 5 this suggests a potential shortcut. But we now have an equation involving the sum of two independent

Hamiltonian operators, $\hat{H}_1 + \hat{H}_2$, acting on the wavefunction $\psi_{12}(r_1, r_2)$, and before going any further we need to decide what this means for the wavefunction itself.

Step (2): The Wavefunctions of Linear Composite Systems

We might be tempted to assume that the total wavefunction for the two-electron system can be written as a superposition of the wavefunctions of the independent electrons, according to $\psi_{12}(r_1, r_2) = \psi_1(r_1) + \psi_2(r_2)$. But, in fact, in a composite system involving the linear combination of two independent Hamiltonians, the system wavefunction is actually the *product* of the two wavefunctions, $\psi_{12}(r_1, r_2) = \psi_1(r_1) \psi_2(r_2)$. Let's just quickly prove this.

From (8.6) we have

$$\hat{H}_{12}\psi_{12}(r_1, r_2) \approx (\hat{H}_1 + \hat{H}_2)\psi_{12}(r_1, r_2) = (\hat{H}_1 + \hat{H}_2)\psi_1(r_1) \psi_2(r_2). \quad (8.7)$$

Because in this approximation the two electrons are regarded to be completely independent of one another, we can safely assume that \hat{H}_1 operates only on $\psi_1(r_1)$, \hat{H}_2 on $\psi_2(r_2)$, so that

$$\begin{aligned} \hat{H}_{12}\psi_{12}(r_1, r_2) &\approx (\hat{H}_1 + \hat{H}_2)\psi_1(r_1) \psi_2(r_2) \\ &\approx \hat{H}_1\psi_1(r_1) \psi_2(r_2) + \hat{H}_2\psi_1(r_1) \psi_2(r_2) \\ &\approx \psi_2(r_2) \hat{H}_1\psi_1(r_1) + \psi_1(r_1) \hat{H}_2\psi_2(r_2). \end{aligned} \quad (8.8)$$

But we know that $\hat{H}_1\psi_1(r_1) = E_1\psi_1(r_1)$, and $\hat{H}_2\psi_2(r_2) = E_2\psi_2(r_2)$, where E_1 and E_2 are the energies of the independent electron states described by $\psi_1(r_1)$ and $\psi_2(r_2)$, respectively. So

$$\begin{aligned} \hat{H}_{12}\psi_{12}(r_1, r_2) &\approx \psi_2(r_2) E_1\psi_1(r_1) + \psi_1(r_1) E_2\psi_2(r_2) \\ &\approx (E_1 + E_2) \psi_1(r_1) \psi_2(r_2) = (E_1 + E_2) \psi_{12}(r_1, r_2). \end{aligned} \quad (8.9)$$

And so we can see that

$$\hat{H}_{12}\psi_{12}(r_1, r_2) = E_{12}\psi_{12}(r_1, r_2), \text{ where } E_{12} \approx E_1 + E_2. \quad (8.10)$$

In other words, in this approximation the energy of the two-electron system E_{12} is the sum of the energies of the independent electrons, as we would expect.

Obviously, we can generalize (8.9) and (8.10) to a system of n independent, non-interacting particles:

$$(\hat{H}_1 + \hat{H}_2 + \dots + \hat{H}_n)\psi = (E_1 + E_2 + \dots + E_n)\psi, \text{ where } \psi = \psi_1\psi_2 \dots \psi_n. \quad (8.11)$$

Thus, we can write the spatial wavefunction $\psi_{12}(r_1, r_2)$ as the product of two 'hydrogenic' wavefunctions:

$$\psi_{12}(r_1, r_2) = \psi_{n_1 l_1 m_{l_1}}(r_1) \psi_{n_2 l_2 m_{l_2}}(r_2). \quad (8.12)$$

Here the quantum numbers n_1, l_1, m_{l_1} refer to the orbital occupied by electron 1; n_2, l_2, m_{l_2} refer to the orbital occupied by electron 2. From Eqs. (5.46) and recalling from Chapter 3 that $R_Z = Z^2 R_H$, we can deduce that the total energy is given by

$$E_{12} = E_1 + E_2 = -4R_H \left(\frac{1}{n_1^2} + \frac{1}{n_2^2} \right) = -4 \frac{m_e e^4}{8\epsilon_0^2 h^2} \left(\frac{1}{n_1^2} + \frac{1}{n_2^2} \right). \quad (8.13)$$

This suggests that if both electrons occupy the lowest-energy 1s orbital— $\psi_{100}(r_1) \psi_{100}(r_2)$ —then the energy of this 'ground state' configuration, which we write as $1s^2$, should be eight times that of the ground state of the hydrogen atom. Alas, the ground state of the hydrogen atom is about -13.6 electron volts, and that of the helium atom is about -79 electron volts, or about 5.8 times that of hydrogen, a discrepancy of nearly 40%. This simply demonstrates that electron–electron repulsion can't be ignored without consequences.

This is all well and good. But now we must own up to the fact that this still can't be the whole story. Equation (8.12) assumes we know which electron is in which orbital. But all electrons possess identical properties of mass and charge. There is absolutely no way of telling one electron from another, and we have to work out how we deal with this kind of *quantum indistinguishability*.

Step (3): Introducing Quantum Indistinguishability

In acknowledging that electrons are indistinguishable quantum particles, we are confronted with an inability to choose between two perfectly valid expressions for the two-electron wavefunction:

$$\psi_{12}(r_1, r_2) = \psi_{n_1 l_1 m_{l_1}}(r_1) \psi_{n_2 l_2 m_{l_2}}(r_2) \quad (8.14)$$

and

$$\psi_{21}(r_1, r_2) = \psi_{n_2 l_2 m_{l_2}}(r_1) \psi_{n_1 l_1 m_{l_1}}(r_2). \quad (8.15)$$

In other words, even though the two sets of values of n , l , and m_l might be distinctly different, we can't say which electron is in which orbital. It's also pretty clear that swapping the electrons doesn't affect the energies of the two-electron state: $E_{21} = E_{12}$. In situations like this, the correct way to proceed is to form a normalized linear combination

of the two possibilities. There are two ways of doing this:

$$\psi_+(r_1, r_2) = \frac{1}{\sqrt{2}} [\psi_{n_1 l_1 m_{l_1}}(r_1) \psi_{n_2 l_2 m_{l_2}}(r_2) + \psi_{n_2 l_2 m_{l_2}}(r_1) \psi_{n_1 l_1 m_{l_1}}(r_2)] \quad (8.16)$$

and

$$\psi_-(r_1, r_2) = \frac{1}{\sqrt{2}} [\psi_{n_1 l_1 m_{l_1}}(r_1) \psi_{n_2 l_2 m_{l_2}}(r_2) - \psi_{n_2 l_2 m_{l_2}}(r_1) \psi_{n_1 l_1 m_{l_1}}(r_2)]. \quad (8.17)$$

In (8.16) and (8.17) the normalization factor $1/\sqrt{2}$ reflects the fact that both $\psi_{12}(r_1, r_2)$ and $\psi_{21}(r_1, r_2)$ are equally probable such that, when we take the modulus-squares of the wavefunctions, they will be equally weighted by a factor of $\frac{1}{2}$.

At this stage we don't need to choose between these possibilities, but it will shortly prove helpful to note the properties of these combinations when we interchange the two electrons. Swopping the labels r_1 and r_2 in Eq. (8.16) gives

$$\begin{aligned} \psi_+(r_2, r_1) &= \frac{1}{\sqrt{2}} [\psi_{n_1 l_1 m_{l_1}}(r_2) \psi_{n_2 l_2 m_{l_2}}(r_1) + \psi_{n_2 l_2 m_{l_2}}(r_2) \psi_{n_1 l_1 m_{l_1}}(r_1)] \\ &= \psi_+(r_1, r_2), \end{aligned} \quad (8.18)$$

and we see that $\psi_+(r_1, r_2)$ is *symmetric* (it doesn't change sign) on the interchange of the two electrons. In contrast, swopping the labels in (8.17) gives

$$\begin{aligned} \psi_-(r_2, r_1) &= \frac{1}{\sqrt{2}} [\psi_{n_1 l_1 m_{l_1}}(r_2) \psi_{n_2 l_2 m_{l_2}}(r_1) - \psi_{n_2 l_2 m_{l_2}}(r_2) \psi_{n_1 l_1 m_{l_1}}(r_1)] \\ &= -\psi_-(r_1, r_2). \end{aligned} \quad (8.19)$$

And we see that $\psi_-(r_1, r_2)$ is *antisymmetric* (it changes sign) on the interchange of the two electrons.

Step (4): Selection Rules for Electric Dipole Transitions

These symmetry properties have important consequences for the spectrum of helium, which we can understand by considering the selection rules associated with so-called *electric dipole transitions*. Think of it this way. In any absorption or emission transition in which an electron is promoted or demoted from one orbital to another, the strength of the transition depends on the electric dipole moment—a measure of the extent by which electric charge is spatially redistributed and separated, or the extent of the *polarity*. The greater the redistribution, the stronger the transition and the stronger the corresponding line in an atomic spectrum.

The transition dipole moment is governed by the electron dipole operator, $\boldsymbol{\mu} = -er$, where r is the dipole length or charge separation distance. In a transition between some

initial state described by the wavefunction ψ_i and some final state described by ψ_f , the dipole moment $\mu_{f \leftarrow i}$ is given by

$$\mu_{f \leftarrow i} = -e \int \psi_f^* r \psi_i d\tau. \quad (8.20)$$

So let's examine what happens in a transition involving an initial antisymmetric wavefunction such as $\psi_-(r_1, r_2)$ and a final symmetric wavefunction such as $\psi_+(r_1, r_2)$. In this case the electric dipole operator is $\mu = -er_1 - er_2 = -e(r_1 + r_2)$, and

$$\mu_{+ \leftarrow -} = -e \int \psi_+^*(r_1, r_2) (r_1 + r_2) \psi_-(r_1, r_2) d\tau_1 d\tau_2. \quad (8.21)$$

We can see immediately that swapping the labels of the electrons changes the sign of the antisymmetric function $\psi_-(r_1, r_2)$ and hence it changes the sign of the integrand in (8.21):

$$\begin{aligned} \mu_{+ \leftarrow -} &= -e \int \psi_+^*(r_2, r_1) (r_1 + r_2) \psi_-(r_2, r_1) d\tau_1 d\tau_2 \\ &= +e \int \psi_+^*(r_1, r_2) (r_1 + r_2) \psi_-(r_1, r_2) d\tau_1 d\tau_2. \end{aligned} \quad (8.22)$$

This isn't physically acceptable. Like any measurable quantity, the transition dipole moment can't depend on the way we *choose* to label the electrons. We conclude that this integral must actually be zero—no such transition is possible.

The end result is that transitions between states described by symmetric and antisymmetric spatial wavefunctions are *forbidden*. The *allowed* transitions are therefore between initial and final states which are either both symmetric or both antisymmetric.

Step (5): Electron Spin Wavefunctions

In this next step we need to account for electron spin and devise expressions for the wavefunctions for the different possible spin configurations of the two electrons. From our earlier discussion we know that there are only two possible spin orientations, which we call spin down, \downarrow , and spin up, \uparrow . If the electron spins in a two-electron system could be considered to be completely independent of one another, we might anticipate four possible configurations, $\uparrow_1 \uparrow_2$, $\uparrow_1 \downarrow_2$, $\downarrow_1 \uparrow_2$, and $\downarrow_1 \downarrow_2$. However, we know that the spins couple together and combine like vectors.

The spin quantum number s of both electrons is fixed at $\frac{1}{2}$. These combine to produce a total spin quantum number S according to Eq. (8.2). From this it is apparent that there are only two possibilities for S —1 or 0—corresponding to $s_1 + s_2$ and $|s_1 - s_2|$. Configurations with a total spin angular momentum quantum number of 0 have a multiplicity $(2S + 1) = 1$ and are called *singlet* states. As the name implies, there is

Table 8.1 Clebsch–Gordan Coefficients for the Coupling of Two Electron Spins

Configuration	$S = 1$	$S = 1$	$S = 0$	$S = 1$
	$M_S = +1$	$M_S = 0$	$M_S = 0$	$M_S = -1$
$\uparrow_1\uparrow_2$	1			
$\uparrow_1\downarrow_2$		$1/\sqrt{2}$	$1/\sqrt{2}$	
$\downarrow_1\uparrow_2$		$1/\sqrt{2}$	$-1/\sqrt{2}$	
$\downarrow_1\downarrow_2$				1

only one state, corresponding to $S = 0, M_S = 0$. Those configurations with a total spin angular momentum quantum number of 1 have a multiplicity $(2S + 1) = 3$ and are called *triplet* states, corresponding to $S = 1, M_S = -1, 0, +1$.

The extent of coupling of the four different possible electron spin configurations is governed by the relevant coupling coefficients or *Clebsch–Gordan coefficients*. Their derivation is quite a complicated business but the good news for all students studying quantum mechanics is that these have been worked out for all cases likely to be of interest and can be simply looked up in a table. For the specific case of the coupling of two electron spins, the coefficients are summarized in Table 8.1.

If we denote the spin wavefunctions as $\chi_{12}(\uparrow_1\uparrow_2)$, etc., then we can now write the wavefunctions for all four possibilities as follows:

$$S = 0, M_S = 0, \chi_-^0(1, 2) = \frac{1}{\sqrt{2}} [\chi_{12}(\uparrow_1\downarrow_2) - \chi_{12}(\downarrow_1\uparrow_2)] \quad (8.23)$$

and

$$S = 1, M_S = +1, \chi_+^{+1}(1, 2) = \chi_{12}(\uparrow_1\uparrow_2) \quad (8.24)$$

$$S = 1, M_S = 0, \chi_+^0(1, 2) = \frac{1}{\sqrt{2}} [\chi_{12}(\uparrow_1\downarrow_2) + \chi_{12}(\downarrow_1\uparrow_2)] \quad (8.25)$$

$$S = 1, M_S = -1, \chi_+^{-1}(1, 2) = \chi_{12}(\downarrow_1\downarrow_2). \quad (8.26)$$

As before, the symmetry of each spin wavefunction is denoted by the use of the subscripts $-$ and $+$. The superscript indicates the value of M_S associated with the wavefunction. We note from this that the singlet spin wavefunction is antisymmetric with respect to the interchange of the labels for electrons 1 and 2: $\chi_-^0(2, 1) = -\chi_-^0(1, 2)$. All the wavefunctions for the triplet states are symmetric to the interchange: $\chi_+^{M_S}(2, 1) = \chi_+^{M_S}(1, 2)$.

We can now pull all these threads together. The total two-electron wavefunctions for the helium atom are products of the spatial *and* spin component functions, and we can quickly deduce that there are eight different ways of writing these, corresponding to different singlet and triplet states:

Singlet wavefunctions	Overall symmetry	Triplet wavefunctions	Overall symmetry
$\psi_+(r_1, r_2) \chi_-^0(1, 2)$	Antisymmetric	$\psi_+(r_1, r_2) \chi_+^{+1}(1, 2)$	Symmetric
$\psi_-(r_1, r_2) \chi_-^0(1, 2)$	Symmetric	$\psi_+(r_1, r_2) \chi_+^0(1, 2)$	Symmetric
		$\psi_+(r_1, r_2) \chi_+^{-1}(1, 2)$	Symmetric
		$\psi_-(r_1, r_2) \chi_+^{+1}(1, 2)$	Antisymmetric
		$\psi_-(r_1, r_2) \chi_+^0(1, 2)$	Antisymmetric
		$\psi_-(r_1, r_2) \chi_+^{-1}(1, 2)$	Antisymmetric

In this analysis we judge the symmetry properties of the total wavefunctions on the understanding that this is based on the *simultaneous* interchange of the labels for both spatial and spin components.

Step (6): Comparison with the Spectrum of Helium

We're almost there. All we do now is anticipate what this might mean for different configurations of the electrons in specific orbitals (we'll find we can get everything we need just by considering the lowest-energy s, p, and d orbitals), and compare this with what we actually see in the spectrum of helium.

We start with the ground state, in which both electrons occupy the 1s orbital, $1s^2$. We characterize the states of multi-electron atoms in terms of the values of the angular momentum quantum numbers L , S , and \mathcal{J} . By analogy with the hydrogen atomic orbitals, states with $L = 0$ are labelled S (states with $L = 1$ are labelled P; $L = 2$, D, etc.). The value of S appears as a superscript, as in 1S (singlet-S) or 3P (triplet-P). The value of \mathcal{J} then appears as a subscript, as in 1S_0 (singlet-S-nought) or $^3P_0, ^3P_1, ^3P_2$ (triplet-P-nought, triplet-P-one, triplet-P-two).*

For the configuration $1s^2$, both electrons are in the lowest-energy s orbital ($l_1 = l_2 = 0 = L$), and the spins are paired (singlet state). The 'term symbol' for this ground state of helium is therefore 1S_0 .

But we also know that in the ground state both electrons occupy the same orbital, $n_1, l_1, m_{l1} = n_2, l_2, m_{l2}$. We can see immediately from Eq. (8.17) that in these circumstances the antisymmetric function vanishes: $\psi_-(r_1, r_2) = 0$. This means that the

* Remember from Eq. (8.3) that for $L = 1, S = 1$ (3P), the possible values of \mathcal{J} are 2 ($L + S$), 1, and 0 ($|L - S|$). Hence $^3P_0, ^3P_1$, and 3P_2 .

spatial wavefunction for the ground state *must be symmetric*— $\psi_+(r_1, r_2)$ —and the total wavefunction, including spatial and spin components, *must therefore be antisymmetric overall*, $\psi_+(r_1, r_2) \chi_-^0(1, 2)$. This is quite important, and useful, as we know from Step (4) that the selection rule for electric dipole transitions dictates that the allowed transitions to and from the ground state will involve only excited states with symmetric spatial wavefunctions. *States with antisymmetric spatial wavefunctions cannot be reached from the ground state.*

Things get rather interesting when we organize the possible electronic states into singlets and triplets. If we excite an electron from the ground $1s^2(^1S_0)$ state into the $1s^12p^1$ configuration, we have two possibilities. In the first, we preserve the orientations of the electron spins so that they remain ‘paired’. The result is the $1s^12p^1(^1P_1)$ state. As the term symbol suggests, this is a singlet state. What happens when the spin of the excited electron is inverted? Then we have a configuration $1s^12p^1(^3P_0, ^3P_1, ^3P_2)$, which are all triplet states.

Let’s now inspect the spectrum of helium and use it to tell us which kinds of transitions are allowed and which are forbidden. A selection of some of the most intense lines in the spectrum is included in Table 8.2.¹¹ The first three lines listed in this table originate with the ground state, from which we conclude that the singlet states $1s^12p^1(^1P_1)$, $1s^13p^1(^1P_1)$ and $1s^14p^1(^1P_1)$ must all have symmetric spatial wavefunctions. From lines 6, 8, and 10, which all involve $1s^12p^1(^1P_1)$, we conclude the same for $1s^12s^1(^1S_0)$, $1s^13s^1(^1S_0)$ and $1s^13d^1(^1D_2)$. From this we can generalize: *all singlet states have symmetric spatial wavefunctions.*

Table 8.2 *A Selection of Transitions in the Absorption Spectrum of Helium*

#	Initial state	Final state	Wavelength (nm)	Relative intensity
1	$1s^2(^1S_0)$	$1s^14p^1(^1P_1)$	52.2	100
2	$1s^2(^1S_0)$	$1s^13p^1(^1P_1)$	53.7	400
3	$1s^2(^1S_0)$	$1s^12p^1(^1P_1)$	58.4	1000
4	$1s^12s^1(^3S_1)$	$1s^13p^1(^3P_0, ^3P_1, ^3P_2)$	388.9	560
5	$1s^12p^1(^3P_0, ^3P_1, ^3P_2)$	$1s^13d^1(^3D_1, ^3D_2, ^3D_3)$	587.6	870
6	$1s^12p^1(^1P_1)$	$1s^13d^1(^1D_2)$	667.8	200
7	$1s^12p^1(^3P_0, ^3P_1, ^3P_2)$	$1s^13s^1(^3S_1)$	706.5	180
8	$1s^12p^1(^1P_1)$	$1s^13s^1(^1S_0)$	728.1	50
9	$1s^12s^1(^3S_1)$	$1s^12p^1(^3P_0, ^3P_1, ^3P_2)$	1083.0	1650
10	$1s^12s^1(^1S_0)$	$1s^12p^1(^1P_1)$	2058.1	500

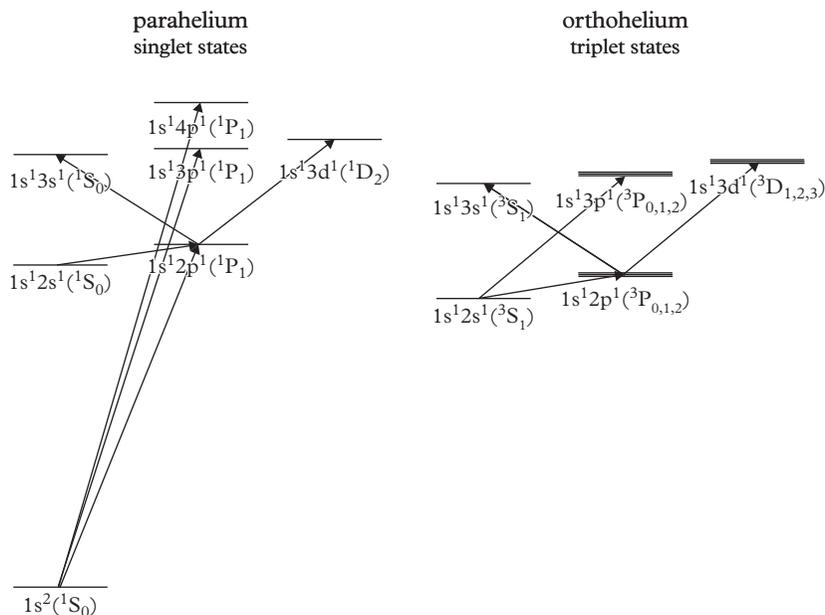


Figure 8.2 The energy levels of parahelium and orthohelium (not to scale), showing the allowed transitions featured in Table 8.2.

Now we must acknowledge something quite remarkable. There are *no* lines in the spectrum corresponding to singlet–triplet or triplet–singlet transitions. This becomes abundantly clear when we draw these transitions on an energy-level diagram, as shown in Fig. 8.2. This distinction between the collection of singlet and triplet states of helium led to their classification as distinct forms, named parahelium and orthohelium.

From this we can deduce a spin selection rule, $\Delta S=0$, and conclude that the absence of a transition from $1s^2(^1S_0)$ to $1s^12p^1(^3P_0, ^3P_1, ^3P_2)$ must mean that the spatial wavefunction for the triplet state is antisymmetric (and so forbidden for electric dipole transitions). From lines 4, 5, 7, and 9 we generalize this to: *all triplet states have antisymmetric spatial wavefunctions*.

We should note in passing that all the transitions included in Table 8.2 conform to the selection rule $\Delta l = \pm 1$ (see the discussion in Chapter 5), which allows transitions between s and p orbitals and between p and d orbitals but forbids transitions between s–s, p–p, d–d, and s–d orbitals, and so on. Don’t worry—we’ll soon have an explanation for this, too.

Where does all this leave us? Well, determining that singlets have exclusively symmetric spatial wavefunctions and triplets have exclusively antisymmetric spatial wavefunctions allows us to strike out half of the possible combinations from our initial list, above, leaving us with the following:

Singlet wavefunctions	Overall symmetry	Triplet wavefunctions	Overall symmetry
$\psi_+(r_1, r_2) \chi_-^0(1, 2)$	Antisymmetric	$\psi_-(r_1, r_2) \chi_+^{+1}(1, 2)$	Antisymmetric
		$\psi_-(r_1, r_2) \chi_+^0(1, 2)$	Antisymmetric
		$\psi_-(r_1, r_2) \chi_+^{-1}(1, 2)$	Antisymmetric

This is the generalized *Pauli principle*:

The total (spatial \times spin) wavefunction for a multi-electron atom must be exclusively antisymmetric with respect to the pairwise interchange of electrons.

We can deduce one important consequence of the Pauli principle by considering what happens to the total wavefunction if we try to push two electrons into the same orbital (n, l, m_l) with the same spin orientations:

$$\psi_+(r_1, r_2) \chi_-^0(1, 2) = \frac{1}{\sqrt{2}} \psi_+(r_1, r_2) [\chi_{12}(\uparrow_1 \uparrow_2) - \chi_{12}(\uparrow_1 \uparrow_2)] = 0. \quad (8.27)$$

If we try to do the same with the triplet states, then we already know from Eq. (8.17) that $\psi_-(r_1, r_2)$ will likewise vanish. This is *Pauli's exclusion principle*: no two electrons can occupy the same quantum state (i.e. no two electrons can possess the same values of the four quantum numbers n, l, m_l, m_s).

Fermions and Bosons

Although this seems like something of a breakthrough, it's important to realize that whilst we have some understanding of how the structures of the spatial wavefunctions come about—as 'hydrogenic' solutions of Schrödinger's wave equation—there is so far no such understanding for the spin wavefunctions. These have been introduced into the above derivation in a rather *ad hoc* manner, with no real accounting for where the different spin orientations \downarrow and \uparrow arise, or why the spin quantum number is restricted to the single value of $\frac{1}{2}$.

Consequently, Heisenberg offered his work to Pauli 'with doubtful feelings and no satisfaction. The calculations are all so imprecise and incomplete'.¹² The understanding would come only in a properly relativistic treatment of the wave equation, which we will go on to consider in Chapter 9.

But we shouldn't leave this chapter without acknowledging that the Pauli principle applies not just to electrons, but to *all* particles with half-integral spin quantum numbers, which includes protons, neutrons, and their constituent quarks. Such particles are classified as *fermions* (named for Enrico Fermi). Our present understanding of elementary particle physics—summarized in the so-called 'standard model'—is that all matter

particles are fermions, and all multifermion systems have wavefunctions antisymmetric to the pairwise interchange of particles.

In contrast bosons (named for Satyendra Nath Bose) have integral spin quantum numbers (0, 1, 2) and wavefunctions that are symmetric to pairwise interchange. Particles that in the standard model are responsible for 'carrying' forces between matter particles are all bosons. These include the photon (which carries the electromagnetic force between electrically charged particles), the W and Z particles (so-called 'heavy photons', responsible for the weak force), and the gluons (which carry the strong or 'colour' force between differently coloured quarks). All these particles possess a spin quantum number of 1.

We note in passing that a photon with $s = 1$ must impart a unit \hbar of angular momentum to any atom that absorbs it (or must carry away a unit \hbar of angular momentum from any atom that emits it). Angular momentum must be conserved in such processes, so it must be picked up (or given up) by the excited electron. Of course, the only place it can go (or come from) is the *orbital* angular momentum of the excited electron. Hence the selection rule $\Delta l = \pm 1$.

Unlike fermions, bosons are not restricted in terms of the quantum numbers they can possess, and a large number can therefore accumulate in the same quantum state. This is called 'Bose condensation' or 'Bose–Einstein condensation', the most familiar example being laser light. Other examples include superfluidity and superconductivity.

.....

NOTES

1. See, for example, David J. Griffiths, *Introduction to Quantum Mechanics*, 2nd edn, Cambridge University Press, Cambridge, UK, 2017, Chapter 4, and Peter Atkins and Ronald Friedman, *Molecular Quantum Mechanics*, 5th edn, Oxford University Press, Oxford, 2011, Chapter 4.
2. Charles P. Enz, *No Time to be Brief: A Scientific Biography of Wolfgang Pauli*, Oxford University Press, Oxford, 2002, p. 92.
3. E. C. Stoner, 'The Distribution of Electrons among Atomic Levels', *Philosophical Magazine*, 48 (1924), 719, quoted in Enz, *No Time to be Brief*, p. 122.
4. W. Pauli, Jr., 'Über den Einfluß der Geschwindigkeitsabhängigkeit der Elektronenmasse auf den Zeemaneffekt', *Zeitschrift für Physik*, 31 (1925), 373. An English translation of this quote is provided by Enz, *No Time to be Brief*, p. 122.
5. Werner Heisenberg, postcard to Wolfgang Pauli, 15 December 1924. Quoted in Enz, *No Time to be Brief*, p. 124.
6. R. Kronig, 'The Turning Point', in M. Fierz and V. F. Weisskopf (eds), *Theoretical Physics in the Twentieth Century: A Memorial Volume to Wolfgang Pauli*, Wiley Interscience, New York, 1960, p. 21. Quoted by Enz, *No Time to be Brief*, p. 111.
7. Quotations taken from S. A. Goudsmit, 'The Discovery of the Electron Spin', lecture delivered at the golden jubilee of the Dutch Physical Society, April 1971. English translation by J. H. van der Waals, available at <https://www.lorentz.leidenuniv.nl/history/spin/goudsmit.html>.
8. Wolfgang Pauli, quoted in Abraham Pais, *Niels Bohr's Times, in Physics, Philosophy and Polity*, Clarendon Press, Oxford, 1991, p. 243.

9. N. Bohr, in G. E. Uhlenbeck and S. A. Goudsmit, ‘Spinning Electrons and the Structure of Spectra’, *Nature*, **117** (1926), 264–5.
10. Ralph de Laer Kronig, letter to Niels Bohr, 8 April 1926. Quoted in Pais, *Niels Bohr’s Times, in Physics, Philosophy and Polity*, p. 244.
11. Persistent lines of neutral helium (He I), data from the online *Handbook of Basic Atomic Spectroscopic Data*, US National Institute of Standards and Technology (NIST), available at <https://www.physics.nist.gov/PhysRefData/Handbook/Tables/heliumtable3.htm>.
12. Werner Heisenberg, letter to Wolfgang Pauli, 28 July 1926, quoted by Michela Massimi, *Pauli’s Exclusion Principle: The Origin and Validation of a Scientific Principle*, Cambridge University Press, Cambridge, UK, 2005, p. 119.

.....

FURTHER READING

- Atkins, Peter, and Friedman, Ronald, *Molecular Quantum Mechanics*, 5th edn, Oxford University Press, Oxford, 2011, Chapter 7, especially pp. 221–9.
- Cassidy, David C., *Uncertainty: The Life and Science of Werner Heisenberg*, W. H. Freeman, New York, 1992, Chapter 11.
- Enz, Charles P., *No Time to be Brief: A Scientific Biography of Wolfgang Pauli*, Oxford University Press, Oxford, 2002, Chapter 5, especially pp. 119–29.
- Gamow, George, *Thirty Years That Shook Physics*, Dover, New York, 1966, Chapter 3.
- Griffiths, David J., *Introduction to Quantum Mechanics*, 2nd edn, Cambridge University Press, Cambridge, UK, 2017, Chapter 5.
- Massimi, Michela, *Pauli’s Exclusion Principle: The Origin and Validation of a Scientific Principle*, Cambridge University Press, Cambridge, UK, 2005, Chapters 2 and 4.
- Pais, Abraham, *Niels Bohr’s Times, in Physics, Philosophy and Polity*, Oxford University Press, Oxford, 1991, Chapter 11.
- Tomonaga, Sin-Itiro, *The Story of Spin* (translated by Takeshi Oka), University of Chicago Press, Chicago, 1997, Lecture 2.

Dirac's Derivation of the Relativistic Wave Equation

Electron Spin and Antimatter

Although considerable progress had been made in the development of what was by now a substantial body of quantum theoretical principles and equations, these suffered from the limitation that they did not meet the requirements of Einstein's special theory of relativity. Pauli was, in turn, frustrated by the lack of progress on the proper incorporation of electron spin and the exclusion principle into the main structure. He called it an 'aesthetic failure'.¹ There was a growing realization that the problem of electron spin was in some way connected with the problem of finding a fully relativistic expression for the wave equation.

Swedish theoretical physicist Oskar Klein independently rediscovered a relativistic version of Schrödinger's wave equation in the spring of 1926. With some further modifications by Hamburg theorist Walter Gordon, this came to be known as the *Klein-Gordon equation*. In essence, this is the relativistic equation that Schrödinger himself had discovered and subsequently abandoned in January 1926 when he found that it did not yield predictions in agreement with the spectroscopy of the hydrogen atom.

Meanwhile, English physicist Paul Dirac had become all too aware that whilst his major contributions to the mathematical structure of quantum mechanics were worthy of international recognition, he had on several occasions been pipped to the post by his European rivals. He felt strongly that he needed to discover something original, something that he could be the first to report and claim as his own.

Electron spin was an obvious target for his attention. Towards the end of 1926, Dirac agreed a bet with Heisenberg on how soon spin would be understood within the framework of quantum mechanics. Heisenberg bet on three years at the least. Dirac rashly bet three months. Three months went by, but the problem remained unresolved. But Dirac, who had become completely absorbed by the theory of relativity when he had first heard about it as an engineering student in 1919, now turned his attention to finding a fully relativistic version of the quantum mechanics of an electron.

What he would find would simultaneously solve the riddle of electron spin. But it would also yield much more than anyone had bargained for.

The Klein–Gordon Equation

I mentioned in passing in Chapter 5 that the full time-dependent Schrödinger Eq. (5.28) is a first-order differential in time and a second-order differential in spatial coordinates. This means that space and time are not treated on an ‘equal footing’, and the equation is therefore not consistent with the demands of Einstein’s special theory of relativity.

An alternative approach is to start out with the correct classical, relativistic expression for the energy, which we have already encountered as Eq. (4.7):

$$E^2 = p^2 c^2 + m_0^2 c^4. \quad (9.1)$$

This is valid for a freely moving particle (i.e. in circumstances in which the potential energy $V=0$). If we simply replace the classical momentum p with its quantum-mechanical operator equivalent $\hat{p} = -i\hbar\nabla$, and introduce a general electron wavefunction Ψ (which, let’s not forget, depends on both r and t), we get

$$E^2\Psi = c^2\hat{p}^2\Psi + m_0^2c^4\Psi. \quad (9.2)$$

Or

$$E^2\Psi = c^2\left(\hat{p}^2 + m_0^2c^2\right)\Psi. \quad (9.3)$$

We know from Eq. (5.28) that

$$i\hbar\frac{\partial}{\partial t}\Psi = E\Psi \quad (9.4)$$

We’re now faced with two options. In Eq. (9.3), energy appears as E^2 whereas in (9.4) it appears as E . We can choose to bring these together by taking the square root of (9.3), the implication being that applying the resulting square root operator once returns the eigenvalue E , applying it twice returns the eigenvalue E^2 :

$$i\hbar\frac{\partial}{\partial t}\Psi = c\sqrt{\hat{p}^2 + m_0^2c^2}\Psi. \quad (9.5)$$

But the square root operator in (9.5) looks distinctly unpromising. Instead, we could choose a second option, which involves applying the operator $i\hbar\partial/\partial t$ in Eq. (9.4) a second time, so that (remember, $i^2 = -1$)

$$-\hbar^2\frac{\partial^2}{\partial t^2}\Psi = E^2\Psi \quad (9.6)$$

Combining (9.3) and (9.6) now gives

$$\begin{aligned} -\hbar^2 \frac{\partial^2}{\partial t^2} \Psi &= c^2 (\hat{p}^2 + m_0^2 c^2) \Psi \\ &= c^2 (-\hbar^2 \nabla^2 + m_0^2 c^2) \Psi. \end{aligned} \quad (9.7)$$

Or

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \Psi - \nabla^2 \Psi + \frac{m_0^2 c^2}{\hbar^2} \Psi = 0. \quad (9.8)$$

Now both time and spatial coordinates are treated equivalently in second-order differentials. This is a version of the Klein–Gordon equation. However, although this is a perfectly valid relativistic equation, it doesn't quite give us what we're looking for. It doesn't feature electron spin in any obvious way. It turns out that the Klein–Gordon equation is applicable only in the *absence* of a magnetic field or for bosons with spin quantum number 0.

Dirac's fascination with relativity and his already burgeoning reputation as a quantum physicist made the search for a fully relativistic form of the new quantum mechanics irresistible. Dirac didn't like the Klein–Gordon equation. Neither did Heisenberg. 'Herr Pauli,' Hungarian physicist Johann Kudar reported to Dirac in a letter dated 21 December 1926, 'regards the relativistic wave equation of *second* order with much suspicion'.²

Dirac had made several abortive attempts at a relativistic version of the wave equation over the previous two years. At a conference in Brussels sponsored by the Belgian industrialist Ernest Solvay in October 1927 he raised the problem with Bohr, only to be told by Bohr that Klein had already solved it. Dirac tried to explain why Klein's equation was unsatisfactory but was cut off when the lecture they had gathered to hear commenced.³

Though short, the conversation with Bohr convinced Dirac that a proper solution had to be found as a matter of some urgency. He once again set to work on the problem as soon as he got back from Brussels. He worked alone, in relative isolation, consulting nobody. It was an approach that suited his personality.

An important clue was available in a treatment of electron spin that Pauli had published in May 1927.⁴

Pauli Spin Matrices

Although we don't yet know how they arise, there's a further set of manipulations of the electron spin wavefunctions that will prove useful in what follows. If we assume that the electron spin angular momentum can be considered in the same way as orbital angular momentum, then we can make use of all the results captured in Appendix 6.

If, analogous to the orbital angular momentum operator \hat{L} , we define a spin operator as \hat{S} , with components \hat{S}_x , \hat{S}_y , and \hat{S}_z , then from Eq. (A6.13) we have

$$[\hat{S}_x, \hat{S}_y] = i\hbar\hat{S}_z \quad [\hat{S}_y, \hat{S}_z] = i\hbar\hat{S}_x \quad [\hat{S}_z, \hat{S}_x] = i\hbar\hat{S}_y. \quad (9.9)$$

Obviously, the spin wavefunctions, which we write as $\chi(\uparrow)$ and $\chi(\downarrow)$, can no longer be considered in terms of the spherical harmonic functions used to describe orbital motion, but we can still assume that the magnitude of the spin vector is given by (cf. Eq. (A6.19))

$$S = \sqrt{s(s+1)}\hbar \quad (9.10)$$

From Eq. (A6.24) we can further assume that

$$\begin{aligned} \hat{S}_z\chi(\uparrow) &= m_s\hbar\chi(\uparrow) = +\frac{1}{2}\hbar\chi(\uparrow) \\ \hat{S}_z\chi(\downarrow) &= m_s\hbar\chi(\downarrow) = -\frac{1}{2}\hbar\chi(\downarrow). \end{aligned} \quad (9.11)$$

The unique ‘two-valuedness’ of the spin wavefunctions suggests an interesting alternative representation. Instead of writing the wavefunctions as $\chi(\uparrow)$ and $\chi(\downarrow)$, we write them instead as simple column matrices:

$$\chi(\uparrow) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (9.12)$$

$$\chi(\downarrow) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (9.13)$$

The operator \hat{S}_z can then be written as a square 2×2 matrix, such that when we apply the rules of matrix multiplication we recover the same results (a very brief introduction to matrices can be found in Appendix 7). For example, if we set up \hat{S}_z as follows:

$$\hat{S}_z = \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (9.14)$$

then

$$\hat{S}_z\chi(\uparrow) = \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\hbar \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\hbar\chi(\uparrow), \quad (9.15)$$

and

$$\hat{S}_z\chi(\downarrow) = \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2}\hbar \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -\frac{1}{2}\hbar\chi(\downarrow). \quad (9.16)$$

Although I don't propose to derive them here, we can determine equivalent matrix representations for \hat{S}_x and \hat{S}_y :

$$\hat{S}_x = \frac{1}{2}\hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \hat{S}_y = \frac{1}{2}\hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \hat{S}_z = \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (9.17)$$

Let's just prove that this is all consistent by considering the commutator $[\hat{S}_x, \hat{S}_y]$:

$$\begin{aligned} [\hat{S}_x, \hat{S}_y] &= \hat{S}_x \hat{S}_y - \hat{S}_y \hat{S}_x \\ &= \frac{1}{4}\hbar^2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \frac{1}{4}\hbar^2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \frac{1}{4}\hbar^2 \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} - \frac{1}{4}\hbar^2 \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \\ &= \frac{1}{4}i\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{1}{4}i\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= i\hbar \cdot \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = i\hbar \hat{S}_z. \end{aligned} \quad (9.18)$$

Because each of these spin angular momentum operators carry the common factor $\frac{1}{2}\hbar$, it is convenient to express them in terms of the *Pauli spin matrices*:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (9.19)$$

such that $\hat{S}_x = \frac{1}{2}\hbar\sigma_x$, etc.

The spin matrices have some interesting properties that we will find most useful in what follows. The only difference between them and their corresponding spin angular momentum operators is a scalar factor $\frac{1}{2}\hbar$, so the spin matrices have very similar commutation relations:

$$[\sigma_x, \sigma_y] = 2i\sigma_z \quad [\sigma_y, \sigma_z] = 2i\sigma_x \quad [\sigma_z, \sigma_x] = 2i\sigma_y. \quad (9.20)$$

We can easily show this for the case of $[\sigma_x, \sigma_y]$ as follows:

$$\begin{aligned} [\sigma_x, \sigma_y] &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} - \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \\ &= i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= 2i\sigma_z. \end{aligned} \quad (9.21)$$

It's also worth taking a quick look at the anti-commutators, such as $\{\sigma_x, \sigma_y\}$:

$$\begin{aligned}
 \{\sigma_x, \sigma_y\} &= \sigma_x\sigma_y + \sigma_y\sigma_x \\
 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
 &= \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} + \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \\
 &= i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
 &= 0.
 \end{aligned} \tag{9.22}$$

From (9.22) it follows that $\sigma_x\sigma_y = -\sigma_y\sigma_x$, and we can similarly deduce that $\sigma_y\sigma_z = -\sigma_z\sigma_y$, and $\sigma_z\sigma_x = -\sigma_x\sigma_z$. But, if this is the case, it then follows that

$$[\sigma_x, \sigma_y] = \sigma_x\sigma_y - \sigma_y\sigma_x = \sigma_x\sigma_y + \sigma_x\sigma_y = 2\sigma_x\sigma_y. \tag{9.23}$$

And so from (9.21) and (9.23) we have

$$\sigma_x\sigma_y = i\sigma_z. \tag{9.24}$$

Likewise,

$$\sigma_y\sigma_z = i\sigma_x \tag{9.25}$$

$$\sigma_z\sigma_x = i\sigma_y. \tag{9.26}$$

Of course, we could have arrived at these expressions directly from the definitions of the Pauli spin matrices given in Eq. (9.19), but this wouldn't have afforded quite so much fun.

I'm going to ask you to bear with me for this next bit, as its relevance to the discussion is not immediately obvious. But I assure you that what follows will lead us to an essential ingredient in Dirac's derivation of the relativistic wave equation.

A quick glance at Appendix 6 will convince you that the spin matrices behave rather like the components of a three-dimensional vector, which we can write as $\boldsymbol{\sigma}$. This shouldn't come as much of a surprise, as the spin matrices are based on the spin angular momentum operators written as 2×2 matrices, and the quantum mechanical operators have classical (vector) counterparts. So let's look at the scalar or 'dot' product of $\boldsymbol{\sigma}$ with an arbitrary three-dimensional vector \mathbf{P} with which it commutes:

$$\boldsymbol{\sigma} \cdot \mathbf{P} = \sigma_x P_x + \sigma_y P_y + \sigma_z P_z. \tag{9.27}$$

Likewise, if \mathbf{Q} is another vector which also commutes with $\boldsymbol{\sigma}$, then $\boldsymbol{\sigma} \cdot \mathbf{Q} = \sigma_x Q_x + \sigma_y Q_y + \sigma_z Q_z$. Thus,

$$\begin{aligned}
(\boldsymbol{\sigma} \cdot \mathbf{P}) \cdot (\boldsymbol{\sigma} \cdot \mathbf{Q}) &= (\sigma_x P_x + \sigma_y P_y + \sigma_z P_z) (\sigma_x Q_x + \sigma_y Q_y + \sigma_z Q_z) \\
&= \sigma_x^2 P_x Q_x + \sigma_x \sigma_y P_x Q_y + \sigma_x \sigma_z P_x Q_z \\
&\quad + \sigma_y \sigma_x P_y Q_x + \sigma_y^2 P_y Q_y + \sigma_y \sigma_z P_y Q_z \\
&\quad + \sigma_z \sigma_x P_z Q_x + \sigma_z \sigma_y P_z Q_y + \sigma_z^2 P_z Q_z.
\end{aligned} \tag{9.28}$$

But we know that $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1$, $\sigma_y \sigma_x = -\sigma_x \sigma_y$, $\sigma_y \sigma_z = -\sigma_z \sigma_y$, and $\sigma_z \sigma_x = -\sigma_x \sigma_z$. If we make these substitutions and gather terms together, we have

$$\begin{aligned}
(\boldsymbol{\sigma} \cdot \mathbf{P}) \cdot (\boldsymbol{\sigma} \cdot \mathbf{Q}) &= P_x Q_x + P_y Q_y + P_z Q_z \\
&\quad + \sigma_x \sigma_y (P_x Q_y - P_y Q_x) + \sigma_y \sigma_z (P_y Q_z - P_z Q_y) \\
&\quad + \sigma_z \sigma_x (P_z Q_x - P_x Q_z).
\end{aligned} \tag{9.29}$$

We note that $P_x Q_x + P_y Q_y + P_z Q_z = \mathbf{P} \cdot \mathbf{Q}$. If we now substitute for $\sigma_x \sigma_y$, $\sigma_y \sigma_z$, and $\sigma_z \sigma_x$ using Eqs. (9.24), (9.25), and (9.26), we get

$$\begin{aligned}
(\boldsymbol{\sigma} \cdot \mathbf{P}) \cdot (\boldsymbol{\sigma} \cdot \mathbf{Q}) &= \mathbf{P} \cdot \mathbf{Q} + i\sigma_x (P_y Q_z - P_z Q_y) + i\sigma_y (P_z Q_x - P_x Q_z) \\
&\quad + i\sigma_z (P_x Q_y - P_y Q_x).
\end{aligned} \tag{9.30}$$

But the components in brackets in (9.30) represent the *cross product* of the vectors \mathbf{P} and \mathbf{Q} —see Appendix 6, Eq. (A6.7). So the terms involving the components $i\sigma_x$, $i\sigma_y$, and $i\sigma_z$ represent the *scalar product* of $i\boldsymbol{\sigma}$ with $\mathbf{P} \times \mathbf{Q}$:

$$(\boldsymbol{\sigma} \cdot \mathbf{P}) \cdot (\boldsymbol{\sigma} \cdot \mathbf{Q}) = \mathbf{P} \cdot \mathbf{Q} + i\boldsymbol{\sigma} \cdot (\mathbf{P} \times \mathbf{Q}). \tag{9.31}$$

This looks much like some kind of vector identity, but is, in fact, based entirely on the properties of commutators and anti-commutators of the spin matrices.⁵ We will learn of the full significance of this relation when we come to study Dirac's recipe.

Momentum and Energy in Relativistic Electrodynamics

As we will soon see, although there's much to be gained by formulating a relativistic wave equation for a freely moving electron, to tease out the relation between electron spin and a magnetic field we need to place the electron—considered as a moving point charge—in an electromagnetic field. This means involving ourselves not just in aspects of Maxwell's electrodynamics, but in the version of electrodynamics as modified by the demands of special relativity.

I've decided that this is beyond the scope of the present book. So, if you will indulge me, I propose simply to present the ingredients we will need and invite the more curious among readers to delve into a course on relativistic electrodynamics or consult a suitable text.⁶

Now, there are many different ways of expressing Maxwell’s equations for the electromagnetic field. The ‘standard’ equations are written in terms of the electric and magnetic field vectors \mathbf{E} and \mathbf{B} as follows:

$$\begin{aligned}
 \operatorname{div}\mathbf{E} &= \nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} && \text{Gauss’s law (deduced from Coulomb’s law)} \\
 \operatorname{div}\mathbf{B} &= \nabla \cdot \mathbf{B} = 0 && \text{Gauss’s law for magnetism (there are no magnetic} \\
 &&& \text{monopoles)} \\
 \operatorname{curl}\mathbf{E} &= \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} && \text{Faraday’s law of induction} \\
 \operatorname{curl}\mathbf{B} &= \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \varepsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}, && \text{Ampere’s law (as modified by Maxwell)}
 \end{aligned}
 \tag{9.32}$$

where ρ is the charge density and \mathbf{J} is the current, given by the charge density multiplied by the volume. The constants ε_0 and μ_0 are, respectively, the relative permittivity and permeability of free space, and $c = 1/\sqrt{\varepsilon_0 \mu_0}$.

In most practical situations, the charge density and current are known, and Maxwell’s eight equations (there are three each for $\operatorname{curl}\mathbf{E}$ and $\operatorname{curl}\mathbf{B}$) are used to deduce the six unknowns—the field vectors \mathbf{E} and \mathbf{B} , each of which have three spatial components. But these equations can be simplified through the introduction of the so-called scalar and vector *potentials*, ϕ and \mathbf{A} , defined such that

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}
 \tag{9.33}$$

and

$$\mathbf{B} = \operatorname{curl}\mathbf{A} = \nabla \times \mathbf{A}.
 \tag{9.34}$$

If Maxwell’s equations are now rewritten based on these relations, there are only four equations and four unknowns: the scalar potential ϕ and the three spatial components of the vector potential, \mathbf{A} .

We’re obviously interested in variables associated with the *motion* of a point charge in an electromagnetic field, such as momentum and energy. It turns out that for our purposes here we do not need to work out specific formulations for these quantities; we just need to anticipate how they will be modified in relativistic electrodynamics.

The derivations are rather convoluted, so I’ll move directly to the results:

$$p \rightarrow p - e\mathbf{A}
 \tag{9.35}$$

and

$$E \rightarrow E - e\phi,
 \tag{9.36}$$

where e is the charge on the electron, and the arrow should be interpreted to mean ‘becomes’ or ‘is replaced by’.⁷

The additional terms in the expressions for both relativistic momentum and energy result from the *interaction* between the point charge and the electromagnetic field. The field acts to store momentum and energy that can be exchanged with the particle's momentum and kinetic energy, whenever the particle encounters gradients in the particle–field interaction. The implicit assumption here is that the field is static, $\partial\phi/\partial t = 0$.

We now have almost everything we need. It's worthwhile noting that Dirac was an astute mathematician, and placed considerable emphasis on mathematical rigour, often at the expense of physical interpretation. It is rare in the history of physics for such faith in the mathematics to be rewarded with new, unlooked-for physical insights, especially insights that are subsequently found to be consistent with new discoveries. As we've seen so far in this book, it's more typically the case that the mathematics is tortured and twisted to conform to the physics, in an often-desperate attempt to explain what we see.

Aside from some changes in notation, the derivation provided below follows that in Dirac's original 1928 paper fairly closely.⁸ The derivation is reproduced in Dirac's *The Principles of Quantum Mechanics*, first published in 1930 and now in a fourth edition.⁹ This will prove to be one of the most complex of the recipes in this book, as you can probably judge just from the extended list of ingredients. The Dirac equation typically does not feature in an introductory course on quantum mechanics, and advanced courses tend to assume rather more mathematical competence. This doesn't shake my firm belief that the details of this derivation can be grasped by any student comfortable with a little vector algebra and calculus.

And even though what I present here is greatly simplified compared with Dirac's original, I believe that, whether you are able to follow it or not, you can at least sit back and admire his mathematical wizardry.

The Ingredients

1. A version of the Klein–Gordon equation, (9.3).
2. The Pauli spin matrices, Eq. (9.19).
3. Equation (7.58): $d\langle\hat{\Omega}\rangle/dt = i[\hat{H}, \hat{\Omega}]/\hbar$, which allows us to determine whether the property $\langle\hat{\Omega}\rangle$ is time-independent and conserved, or a *constant of the motion*.
4. Commutation relations for the spatial components of orbital angular momentum and linear momentum, such as $[\hat{L}_z, \hat{p}_x] = i\hbar\hat{p}_y$, $[\hat{L}_z, \hat{p}_y] = -i\hbar\hat{p}_x$, and $[\hat{L}_z, \hat{p}_z] = 0$. These are given in Appendix 6, Eq. (A6.13). In his 1928 paper, Dirac referred to these as the 'Vertauschungs' (exchange or permutation) relations.
5. Exchange relations for the spin matrices, Eqs. (9.24), (9.25), and (9.26).
6. Expressions for momentum and energy in relativistic electrodynamics, Eqs. (9.35) and (9.36).
7. The spin matrix identity, Eq. (9.31).

8. A couple of vector calculus identities: $\nabla \times \nabla f = 0$; $\nabla \times \mathbf{A}f = (\nabla \times \mathbf{A})f + \nabla f \times \mathbf{A}$, where f is a scalar function.
9. The relation between the magnetic field vector, \mathbf{B} , and the vector potential, \mathbf{A} , $\mathbf{B} = \nabla \times \mathbf{A}$, Eq. (9.34).
10. Expressions for the orbital and spin magnetic moments, and the energy of the interactions between these and a magnetic field.

The Recipe

In Step (1) we start with the Klein–Gordon equation and take the square root, which is then rewritten in a linear form. This is achieved by introducing 4×4 matrices (the ‘ α -matrices’) and assuming that the electron wavefunction Ψ can be written as a four-component column matrix. In Step (2) we convert the α -matrices into new 4×4 matrices that resemble the Pauli spin matrices, but with twice the dimensionality. This allows us to deduce a relativistic wave equation for a freely moving electron.

To establish the connection with electron spin, in Step (3) we place the electron in a central field with a spherical potential, and determine the constant of the motion from the relations between the Hamiltonian operator \hat{H} and the operators for angular momentum using Eq. (7.58). To get at the result we need to apply the ‘Vertauschungs’ relations and the exchange relations for the spin matrices.

To explore the physics of the interaction of the spinning electron with a magnetic field, in Step (4) we place the relativistic free electron in an electromagnetic field, and then in Step (5) take the *non-relativistic limit* of the result. This might seem a rather curious thing to do, but we will see that the result is very different to what we get if we start with a non-relativistic electron. On the way we will make use of the spin matrix identity (9.31), a couple of vector calculus identities, and the relation between the magnetic field vector \mathbf{B} and the vector potential \mathbf{A} , $\mathbf{B} = \nabla \times \mathbf{A}$. We close this out in Step (6) by evaluating the spin magnetic moment of the electron.

Step (1): Linearize the Square Root of the Klein–Gordon Equation

If we take the square root of both sides of Eq. (9.3), we get

$$E\Psi = c\sqrt{\hat{p}^2 + m_0^2 c^2}\Psi. \quad (9.37)$$

Expanding \hat{p}^2 in (9.37) into its components then gives

$$E\Psi = c\sqrt{\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2 + m_0^2 c^2}\Psi. \quad (9.38)$$

We assume that this rather ugly square root operator can be linearized* by introducing coefficients α_1 - α_4 , such that

$$E\Psi = c(\alpha_1\hat{p}_x + \alpha_2\hat{p}_y + \alpha_3\hat{p}_z + \alpha_4m_0c)\Psi. \quad (9.39)$$

Let's note—for the record—that this expression treats space and time in an equivalent fashion, as required if we are to arrive at a result which meets the demands of Einstein's special theory of relativity. The appearance of E indicates that it is a first-order differential equation in time (Eq. (9.4)), and the components of the linear momentum operator are first-order differentials in the three spatial coordinates.

Obviously, in order for this to work, the α coefficients must possess a very specific set of relationships among themselves. To discover what these are, let's just square the term that multiplies the electron wavefunction on the right-hand side of (9.39):

$$\begin{aligned} &(\alpha_1\hat{p}_x + \alpha_2\hat{p}_y + \alpha_3\hat{p}_z + \alpha_4m_0c)(\alpha_1\hat{p}_x + \alpha_2\hat{p}_y + \alpha_3\hat{p}_z + \alpha_4m_0c) \\ &= \alpha_1^2\hat{p}_x^2 + \alpha_1\alpha_2\hat{p}_x\hat{p}_y + \alpha_1\alpha_3\hat{p}_x\hat{p}_z + \alpha_1\alpha_4\hat{p}_x m_0c \\ &+ \alpha_2\alpha_1\hat{p}_y\hat{p}_x + \alpha_2^2\hat{p}_y^2 + \alpha_2\alpha_3\hat{p}_y\hat{p}_z + \alpha_2\alpha_4\hat{p}_y m_0c \\ &+ \alpha_3\alpha_1\hat{p}_z\hat{p}_x + \alpha_3\alpha_2\hat{p}_z\hat{p}_y + \alpha_3^2\hat{p}_z^2 + \alpha_3\alpha_4\hat{p}_z m_0c \\ &+ \alpha_4\alpha_1m_0c\hat{p}_x + \alpha_4\alpha_2m_0c\hat{p}_y + \alpha_4\alpha_3m_0c\hat{p}_z + \alpha_4^2m_0^2c^2. \end{aligned} \quad (9.40)$$

We should here recall that \hat{p}_x, \hat{p}_y , etc., commute, so $\hat{p}_x\hat{p}_y = \hat{p}_y\hat{p}_x$, and so on. This allows us to simplify (9.40) and gather some terms together:

$$\begin{aligned} &(\alpha_1\hat{p}_x + \alpha_2\hat{p}_y + \alpha_3\hat{p}_z + \alpha_4m_0c)(\alpha_1\hat{p}_x + \alpha_2\hat{p}_y + \alpha_3\hat{p}_z + \alpha_4m_0c) \\ &= \alpha_1^2\hat{p}_x^2 + \alpha_2^2\hat{p}_y^2 + \alpha_3^2\hat{p}_z^2 + \alpha_4^2m_0^2c^2 \\ &+ (\alpha_1\alpha_2 + \alpha_2\alpha_1)\hat{p}_x\hat{p}_y + (\alpha_1\alpha_3 + \alpha_3\alpha_1)\hat{p}_x\hat{p}_z \\ &+ (\alpha_1\alpha_4 + \alpha_4\alpha_1)\hat{p}_x m_0c + (\alpha_2\alpha_3 + \alpha_3\alpha_2)\hat{p}_y\hat{p}_z \\ &+ (\alpha_2\alpha_4 + \alpha_4\alpha_2)\hat{p}_y m_0c + (\alpha_3\alpha_4 + \alpha_4\alpha_3)\hat{p}_z m_0c. \end{aligned} \quad (9.41)$$

We can see from this that if we are to recover Eq. (9.3) from (9.39), then

$$\alpha_1^2 = \alpha_2^2 = \alpha_3^2 = \alpha_4^2 = 1 \quad (9.42)$$

* On the face of it, it's not at all clear why this might be a good assumption, as a square root operator typically implies a differential equation of infinite order. But Dirac convinced himself by 'playing around' with the maths and, in particular, by discovering that $\sqrt{\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2} = \sigma_x\hat{p}_x + \sigma_y\hat{p}_y + \sigma_z\hat{p}_z$. See Helge S. Kragh, *Dirac: A Scientific Biography*, Cambridge University Press, Cambridge, UK, 1990, p. 58. You might want to prove this relation for yourself: just remember that $\sigma_x^2 = 1$ (and so on) and $\sigma_x\sigma_y = -\sigma_y\sigma_x$ (and so on).

and, to eliminate all the cross-terms, we need

$$\begin{aligned} \alpha_1\alpha_2 &= -\alpha_2\alpha_1 & \alpha_1\alpha_3 &= -\alpha_3\alpha_1 & \alpha_2\alpha_3 &= -\alpha_3\alpha_2 \\ \alpha_1\alpha_4 &= -\alpha_4\alpha_1 & \alpha_2\alpha_4 &= -\alpha_4\alpha_2 & \alpha_3\alpha_4 &= -\alpha_4\alpha_3. \end{aligned} \tag{9.43}$$

Dirac had initially presumed that simply extending the Pauli spin matrices to a fourth matrix would satisfy these requirements, but it is clear that 2×2 matrices can't be the answer.

Dirac was arguably a mathematician first, a physicist second. What confronted him was a problem in mathematics, and his principal concern was to solve the mathematical problem, and only then worry about the physical interpretation of the solution. As he played around with the equations, he was struck by an insight: 'I suddenly realized that there was no need to stick to quantities which can be represented by matrices with only two rows and columns,' he later explained. 'Why not go to four rows and columns?'¹⁰

And this, of course, is the answer. I don't propose to derive the forms of the α -matrices here, but will instead demonstrate how their properties meet all the requirements of (9.42) and (9.43):

$$\begin{aligned} \alpha_1 &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \\ \alpha_2 &= \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \\ \alpha_3 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \\ \alpha_4 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \end{aligned} \tag{9.44}$$

Let's just quickly check this for α_1^2 and $\alpha_1\alpha_3$:

$$\alpha_1^2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathbf{1} \tag{9.45}$$

and

$$\begin{aligned}
 \alpha_1 \alpha_3 &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\
 \alpha_3 \alpha_1 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \\
 &= -\alpha_1 \alpha_3.
 \end{aligned} \tag{9.46}$$

I leave you to prove to yourself the consistency of the rest of the α -matrices in (9.44) with the requirements summarized in (9.42) and (9.43).

The introduction of 4×4 matrices implies that the electron wavefunction must now be considered as a four-component column matrix,

$$\Psi = \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix}, \tag{9.47}$$

and, for now, we set aside any concerns we have for the fact that we seem to have twice as many solutions on our hands than we might have bargained for.

Step (2): Introduce the Pauli Spin Matrices

As written, the first three α -matrices already feature the Pauli spin matrices as components (look back at Eq. (9.19)), but these are ‘off-diagonal’ in the sense that they do not sit along the leading diagonals of each matrix:

$$\alpha_1 = \begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix} \quad \alpha_2 = \begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix} \quad \alpha_3 = \begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix}. \tag{9.48}$$

We can fix this simply by factoring the first three α -matrices as products of the ‘double’ spin matrices and a new 4×4 matrix, ρ ,*

$$\alpha_1 = \rho \sigma'_x \quad \alpha_2 = \rho \sigma'_y \quad \alpha_3 = \rho \sigma'_z, \tag{9.49}$$

* In his original derivation, Dirac introduced three ρ -matrices, ρ_1 , ρ_2 , and ρ_3 , but did not make use of ρ_2 and ρ_3 . Consequently, I've chosen to continue with $\rho = \rho_1$.

where

$$\rho = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad (9.50)$$

and

$$\sigma'_x = \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix} \quad \sigma'_y = \begin{pmatrix} \sigma_y & 0 \\ 0 & \sigma_y \end{pmatrix} \quad \sigma'_z = \begin{pmatrix} \sigma_z & 0 \\ 0 & \sigma_z \end{pmatrix}. \quad (9.51)$$

It is safe to assume that all the relationships we established above for the commutators and anti-commutators of the 2×2 spin matrices σ_x , σ_y , and σ_z —and the exchange relations—will apply also for their 4×4 counterparts σ'_x , σ'_y , and σ'_z .

We can now substitute for the α 's in (9.39) as follows:

$$E\Psi = c\rho (\sigma'_x \hat{p}_x + \sigma'_y \hat{p}_y + \sigma'_z \hat{p}_z) \Psi + \beta m_0 c^2 \Psi, \quad (9.52)$$

where $\beta = \alpha_4$.

This can be written a little more succinctly and in a more familiar 'Hamiltonian' form as

$$c\rho (\boldsymbol{\sigma}' \cdot \hat{\mathbf{p}}) \Psi + \beta m_0 c^2 \Psi = E\Psi, \quad (9.53)$$

where $(\boldsymbol{\sigma}' \cdot \hat{\mathbf{p}})$ is the scalar product of the 4×4 'double' spin matrices and the linear momentum operator. This is the relativistic Dirac equation for a freely moving electron.

Step (3): The Connection with Electron Spin

In the absence of a potential field, then the energy E in (9.53) represents the electron *kinetic energy*, \hat{T} , such that $\hat{T}\Psi = E\Psi$, or

$$\hat{T}\Psi = c\rho (\boldsymbol{\sigma}' \cdot \hat{\mathbf{p}}) \Psi + \beta m_0 c^2 \Psi. \quad (9.54)$$

If we now place the electron in a central force field with spherical potential $V(r)$, then we have

$$\hat{H}\Psi = \hat{T}\Psi + V(r)\Psi = c\rho (\boldsymbol{\sigma}' \cdot \hat{\mathbf{p}}) \Psi + \beta m_0 c^2 \Psi + V(r)\Psi. \quad (9.55)$$

Clearly, if we set $V(r) = -e^2/4\pi\epsilon_0 r$, then we have a relativistic version of the wave equation for the hydrogen atom.

We're interested in the angular momentum of the electron in this system which, if it is to be time-independent and conserved, should be a constant of the motion. So let's consider the z -component of the orbital angular momentum, \hat{L}_z (for details see Appendix 6). To be a constant of the motion, then the expectation value $\langle \hat{L}_z \rangle$ must be time-independent. We know from our discussion of Step (5) of the recipe in Chapter 7 that

$$\frac{d}{dt}\langle \hat{L}_z \rangle = \frac{i}{\hbar}[\hat{H}, \hat{L}_z] \quad \text{or} \quad i\hbar \frac{d}{dt}\langle \hat{L}_z \rangle = -[\hat{H}, \hat{L}_z] = [\hat{L}_z, \hat{H}]. \quad (9.56)$$

Since $\beta m_0 c^2$ and $V(r)$ are scalar functions and so commute with \hat{L}_z , from (9.55), we have

$$\begin{aligned} [\hat{L}_z, \hat{H}] &= c\rho [\hat{L}_z, (\boldsymbol{\sigma}' \cdot \hat{\boldsymbol{p}})] = c\rho \left\{ \hat{L}_z (\boldsymbol{\sigma}' \cdot \hat{\boldsymbol{p}}) - (\boldsymbol{\sigma}' \cdot \hat{\boldsymbol{p}}) \hat{L}_z \right\} \\ &= c\rho \left\{ \boldsymbol{\sigma}' \cdot (\hat{L}_z \hat{\boldsymbol{p}} - \hat{\boldsymbol{p}} \hat{L}_z) \right\} \\ &= c\rho \left\{ \boldsymbol{\sigma}' \cdot [\hat{L}_z, \hat{\boldsymbol{p}}] \right\} \\ &= c\rho \left\{ \boldsymbol{\sigma}' \cdot ([\hat{L}_z, \hat{p}_x] + [\hat{L}_z, \hat{p}_y] + [\hat{L}_z, \hat{p}_z]) \right\} \\ &= c\rho \left\{ \sigma'_x [\hat{L}_z, \hat{p}_x] + \sigma'_y [\hat{L}_z, \hat{p}_y] + \sigma'_z [\hat{L}_z, \hat{p}_z] \right\}. \end{aligned} \quad (9.57)$$

From Eq. (A6.8), we know that $\hat{L}_z = x\hat{p}_y - y\hat{p}_x$, and we can readily deduce the 'Vertauschungs' relations involving \hat{L}_z : $[\hat{L}_z, \hat{p}_x] = i\hbar\hat{p}_y$, $[\hat{L}_z, \hat{p}_y] = -i\hbar\hat{p}_x$, and $[\hat{L}_z, \hat{p}_z] = 0$. So, from (9.57) we have

$$i\hbar \frac{d}{dt}\langle \hat{L}_z \rangle = [\hat{L}_z, \hat{H}] = i\hbar c\rho (\sigma'_x \hat{p}_y - \sigma'_y \hat{p}_x), \quad (9.58)$$

and we see immediately that $d\hat{L}_z/dt \neq 0$, so the z -component of the orbital angular momentum is *not* a constant of the motion in this system.

So let's now consider the z -component of the 'double' spin matrix, σ'_z . We follow the same logic:

$$\begin{aligned} i\hbar \frac{d}{dt}\sigma'_z &= [\sigma'_z, \hat{H}] = c\rho [\sigma'_z, (\boldsymbol{\sigma}' \cdot \hat{\boldsymbol{p}})] \\ &= c\rho \left\{ \sigma'_z (\boldsymbol{\sigma}' \cdot \hat{\boldsymbol{p}}) - (\boldsymbol{\sigma}' \cdot \hat{\boldsymbol{p}}) \sigma'_z \right\} \\ &= c\rho \left\{ (\sigma'_z \boldsymbol{\sigma}' - \boldsymbol{\sigma}' \sigma'_z) \cdot \hat{\boldsymbol{p}} \right\} \\ &= c\rho \left\{ [(\sigma'_z \sigma'_x + \sigma'_z \sigma'_y + \sigma'_z \sigma'_z) - (\sigma'_x \sigma'_z + \sigma'_y \sigma'_z + \sigma'_z \sigma'_z)] \cdot \hat{\boldsymbol{p}} \right\} \\ &= c\rho \left\{ [(\sigma'_z \sigma'_x - \sigma'_x \sigma'_z) + (\sigma'_z \sigma'_y - \sigma'_y \sigma'_z) + (\sigma'_z \sigma'_z - \sigma'_z \sigma'_z)] \cdot \hat{\boldsymbol{p}} \right\} \\ &= c\rho \left\{ (\sigma'_z \sigma'_x - \sigma'_x \sigma'_z) \hat{p}_x + (\sigma'_z \sigma'_y - \sigma'_y \sigma'_z) \hat{p}_y + (\sigma'_z \sigma'_z - \sigma'_z \sigma'_z) \hat{p}_z \right\}. \end{aligned} \quad (9.59)$$

By analogy with the Pauli spin matrices and particularly the exchange relations summarized in (9.24)–(9.26), we have $\sigma'_z \sigma'_x = -\sigma'_x \sigma'_z = i\sigma'_y$, and $\sigma'_z \sigma'_y = -\sigma'_y \sigma'_z = -i\sigma'_x$, and so

$$i\hbar \frac{d}{dt} \sigma'_z = [\sigma'_z, \hat{H}] = 2ic\rho (\sigma'_y \hat{p}_x - \sigma'_x \hat{p}_y). \quad (9.60)$$

And we see that σ'_z is not a constant of the motion either. But it's now really rather obvious that we can combine these as follows:

$$\frac{d}{dt} \langle \hat{L}_z \rangle + \frac{1}{2} \hbar \frac{d}{dt} \sigma'_z = c\rho (\sigma'_x \hat{p}_y - \sigma'_y \hat{p}_x) + c\rho (\sigma'_y \hat{p}_x - \sigma'_x \hat{p}_y) = 0. \quad (9.61)$$

There is only one conclusion. *The electron has an intrinsic spin angular momentum which must be added to the orbital angular momentum.* It is the total (orbital plus spin) angular momentum which is time-independent and conserved, and which is a constant of the motion. Of course, if we define \hat{S}'_z —the z -component of the spin angular momentum written as a 4×4 matrix—as $\hat{S}'_z = \frac{1}{2} \hbar \sigma'_z$, then Eq. (9.61) becomes $d(\langle \hat{L}_z \rangle + \langle \hat{S}'_z \rangle)/dt = 0$.

We can obviously repeat this logic for \hat{L}_x , \hat{L}_y , σ'_x , and σ'_y , so we conclude in general that $d\hat{L}/dt + \frac{1}{2} \hbar d\sigma'/dt = 0$. The electron has an intrinsic spin angular momentum of $\frac{1}{2} \hbar \sigma'$.

Step (4): Place the Free Electron in a Relativistic Electromagnetic Field. . .

This result demonstrates that electron spin arises naturally in a proper relativistic treatment of the wave equation, but it doesn't yet suggest how the spin might respond in the presence of a magnetic field. To explore the consequences, we return to the equation for a freely moving electron, (9.53), but we now place the electron in a relativistic electromagnetic field.

From Eqs. (9.35) and (9.36), we know to replace \hat{p} with $\hat{p} - e\mathbf{A}$, and E with $E - e\phi$:

$$c\rho [\sigma' \cdot (\hat{p} - e\mathbf{A})] \Psi + \beta m_0 c^2 \Psi = (E - e\phi) \Psi. \quad (9.62)$$

Step (5): . . . And Take the Non-relativistic Limit

This next step might seem a little strange, but again I'm going to ask you to bear with me. Let's swap (9.62) around and take the square of both sides of the equation:

$$(E - e\phi)^2 \Psi = c^2 \rho^2 [\sigma' \cdot (\hat{p} - e\mathbf{A})]^2 \Psi + (\rho\beta + \beta\rho) [\sigma' \cdot (\hat{p} - e\mathbf{A})] m_0 c^3 \Psi + \beta^2 m_0^2 c^4 \Psi. \quad (9.63)$$

But $\rho^2 = \beta^2 = \mathbf{1}$, and $\rho\beta = -\beta\rho$ (check it out), leaving us with

$$(E - e\phi)^2 \Psi = c^2 [\boldsymbol{\sigma}' \cdot (\hat{\mathbf{p}} - e\mathbf{A})]^2 \Psi + m_0^2 c^4 \Psi. \quad (9.64)$$

This is just a version of Eq. (9.2) with E replaced by $E - e\phi$, and the linear momentum operator $\hat{\mathbf{p}}$ replaced by $\boldsymbol{\sigma}' \cdot (\hat{\mathbf{p}} - e\mathbf{A})$.

In the *non-relativistic limit*, in which the electron is assumed to be moving slowly with respect to the speed of light, we can use the analogy with the classical Eq. (4.12) to write

$$\frac{\hat{\mathbf{p}}^2}{2m_e} \Psi + m_0 c^2 \Psi = E \Psi. \quad (9.65)$$

By comparison with (9.64), we replace E with $E - e\phi$, and $\hat{\mathbf{p}}$ with $\boldsymbol{\sigma}' \cdot (\hat{\mathbf{p}} - e\mathbf{A})$. After a little rearrangement, this gives us access to a non-relativistic limiting form of (9.62):

$$\frac{[\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - e\mathbf{A})]^2}{2m_e} \Psi + m_0 c^2 \Psi + e\phi \Psi = E \Psi. \quad (9.66)$$

Notice that I've dropped the prime on $\boldsymbol{\sigma}$ because in this non-relativistic limit we've returned to a system of 2×2 Pauli spin matrices, and a two-component column matrix for Ψ .

Let's now evaluate the rather forbidding term in square brackets in (9.66):

$$[\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - e\mathbf{A})]^2 \Psi = \{[\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - e\mathbf{A})] \cdot [\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - e\mathbf{A})]\} \Psi. \quad (9.67)$$

If we set $\mathbf{P} = \hat{\mathbf{p}} - e\mathbf{A}$, then the right-hand side of (9.67) becomes $[(\boldsymbol{\sigma} \cdot \mathbf{P}) \cdot (\boldsymbol{\sigma} \cdot \mathbf{P})] \Psi$ and we can make use of the spin matrix identity, Eq. (9.31), to write this as

$$[(\boldsymbol{\sigma} \cdot \mathbf{P}) \cdot (\boldsymbol{\sigma} \cdot \mathbf{P})] \Psi = \mathbf{P}^2 \Psi + i\boldsymbol{\sigma} \cdot (\mathbf{P} \times \mathbf{P}) \Psi. \quad (9.68)$$

We might be tempted to conclude that the second term on the right-hand side of (9.68) must be zero, since the cross product of identical vectors is zero. But notice that I've stubbornly insisted on carrying through the electron wavefunction, Ψ . The vector \mathbf{P} contains the gradient operator ∇ (because $\hat{\mathbf{p}} = -i\hbar\nabla$), and ∇ operates on Ψ , and this makes a world of difference.

Let's now evaluate this in stages, starting with $(\mathbf{P} \times \mathbf{P}) \Psi$. We have

$$\begin{aligned} (\mathbf{P} \times \mathbf{P}) \Psi &= [(\hat{\mathbf{p}} - e\mathbf{A}) \times (\hat{\mathbf{p}} - e\mathbf{A})] \Psi \\ &= (\hat{\mathbf{p}} \times \hat{\mathbf{p}}) \Psi - e[(\hat{\mathbf{p}} \times \mathbf{A}) + (\mathbf{A} \times \hat{\mathbf{p}})] \Psi + e^2 (\mathbf{A} \times \mathbf{A}) \Psi. \end{aligned} \quad (9.69)$$

This is where it becomes really important not to lose sight of the electron wavefunction. The first term on the right-hand side of (9.69) is

$$(\hat{p} \times \hat{p}) \Psi = -i\hbar(\nabla \times \nabla) \Psi = -i\hbar(\nabla \times \nabla \Psi) = 0, \quad (9.70)$$

since $\nabla \times \nabla f = 0$, where f is any scalar function. Similarly, for the third term,

$$e^2(\mathbf{A} \times \mathbf{A}) \Psi = 0, \quad (9.71)$$

since $(\mathbf{A} \times \mathbf{A}) = 0$. We can expand the second term in (9.69) as follows,

$$\begin{aligned} -e[(\hat{p} \times \mathbf{A}) + (\mathbf{A} \times \hat{p})] \Psi &= i\hbar e[(\nabla \times \mathbf{A} \Psi) + (\mathbf{A} \times \nabla \Psi)] \\ &= i\hbar e(\nabla \times \mathbf{A}) \Psi, \end{aligned} \quad (9.72)$$

since $\nabla \times \mathbf{A}f = (\nabla \times \mathbf{A})f + \nabla f \times \mathbf{A} = (\nabla \times \mathbf{A})f - \mathbf{A} \times \nabla f$, which means that $\nabla \times \mathbf{A}f + \mathbf{A} \times \nabla f = (\nabla \times \mathbf{A})f$.

But we know from Eq. (9.34) that $\nabla \times \mathbf{A} = \mathbf{B}$. So, we now have

$$(\mathbf{P} \times \mathbf{P}) \Psi = [(\hat{p} - e\mathbf{A}) \times (\hat{p} - e\mathbf{A})] \Psi = i\hbar e \mathbf{B} \Psi. \quad (9.73)$$

From (9.67) and (9.68), we have (remember, $i^2 = -1$)

$$[\boldsymbol{\sigma} \cdot (\hat{p} - e\mathbf{A})]^2 \Psi = (\hat{p} - e\mathbf{A})^2 \Psi - \hbar e (\boldsymbol{\sigma} \cdot \mathbf{B}) \Psi. \quad (9.74)$$

We're now in a position to put this result into Eq. (9.66):

$$\frac{(\hat{p} - e\mathbf{A})^2}{2m_e} \Psi - \frac{\hbar e}{2m_e} (\boldsymbol{\sigma} \cdot \mathbf{B}) \Psi + m_0 c^2 \Psi + e\phi \Psi = E \Psi. \quad (9.75)$$

But we know from our earlier discussion that the electron spin angular momentum is given by $\hat{S} = \frac{1}{2} \hbar \boldsymbol{\sigma}$, or $\boldsymbol{\sigma} = 2\hat{S}/\hbar$. On substituting for $\boldsymbol{\sigma}$ in (9.75), we get

$$\frac{(\hat{p} - e\mathbf{A})^2}{2m_e} \Psi - 2 \frac{e}{2m_e} (\hat{S} \cdot \mathbf{B}) \Psi + m_0 c^2 \Psi + e\phi \Psi = E \Psi. \quad (9.76)$$

You'll notice that I've deliberately not cancelled the factor of 2 in the second term on the left of (9.76). The reason will soon become obvious.

Step (6): The Magnetic Dipole Moment of a Spinning Electron

To complete this final part of the story, we need to back-fill on some of the details of the interaction of an electron with a magnetic field.

The electron is a charged particle and, when placed in a magnetic field \mathbf{B} , we expect the electron to become aligned with the magnetic 'lines of force', re-orientating like a

compass needle as it enters the field. This is the *magnetic dipole moment*, $\boldsymbol{\mu}$, which depends on the angular momentum of the electron and is manifested as a torque given by $\boldsymbol{\mu} \times \mathbf{B}$. The *energy* associated with this torque is $-\boldsymbol{\mu} \cdot \mathbf{B}$.

For the case of electron *orbital* angular momentum a relation between $\boldsymbol{\mu}_{\text{orbital}}$ and the angular momentum vector \mathbf{L} can be deduced from classical mechanics (see Appendix 6, Eq. (A6.15)),

$$\boldsymbol{\mu}_{\text{orbital}} = \gamma_e \mathbf{L} \quad (9.77)$$

where $\gamma_e = -e/2m_e$ is the *gyromagnetic ratio*, the ratio of the electron magnetic dipole moment to its angular momentum. The associated energy is then given by $-\gamma_e (\mathbf{L} \cdot \mathbf{B})$.

We can follow precisely the same logic with the electron's intrinsic spin angular momentum. Unlike orbital angular momentum, there is no classical counterpart to electron spin but, by analogy with (9.77), we can write

$$\boldsymbol{\mu}_{\text{spin}} = \gamma_e \mathbf{S}, \quad (9.78)$$

where \mathbf{S} is the spin angular momentum vector. This implies an associated energy of $-\gamma_e (\mathbf{S} \cdot \mathbf{B})$.

Again, we replace \mathbf{S} with the quantum-mechanical operator \hat{S} , so that $\boldsymbol{\mu}_{\text{spin}} = \gamma_e \hat{S}$ and the associated energy is given by $-\gamma_e (\hat{S} \cdot \mathbf{B})$. But if we now compare this with the relevant term in (9.76), we see immediately that *the ratio is twice as large* as we might expect from our consideration of orbital angular momentum. This really shouldn't come as too much of a surprise, since the eigenvalues of \hat{L} are $m_l \hbar$, and m_l takes only *integer* values, whereas the eigenvalues of \hat{S} are $m_s \hbar$, and for an electron m_s can take only the values $+\frac{1}{2}$ and $-\frac{1}{2}$.

We accommodate this difference by introducing a new numerical factor, g_e , called variously the '*g-factor*', '*Landé factor*' (named for Alfred Landé), or the '*Landé g-factor*'. For the orbital magnetic moment, $\boldsymbol{\mu}_{\text{orbital}} = g_e \gamma_e \hat{L}$, and $g_e = 1$. For the spin magnetic moment, $\boldsymbol{\mu}_{\text{spin}} = g_e \gamma_e \hat{S}$, and $g_e = 2$.

As I mentioned in Chapter 8, this distinction between orbital and spin magnetic moments had been known from experiment for some time. But Dirac was now able to demonstrate how this difference arises directly from a relativistic treatment of the wave equation for the electron.

News of Dirac's breakthrough spread quickly. Dirac wrote a letter to Max Born in Göttingen ahead of publication of a paper setting out his approach. Léon Rosenfeld described its reception: 'It was immediately seen as *the* solution. It was regarded really as an absolute wonder.'¹¹

The Dream of Philosophers

Dirac had been obliged to pay a price for his solution, however. His use of 4×4 matrices meant that he had twice as many solutions on his hands. The two possible orientations of the electron spin angular momentum account for half of these. The other

two are characterized by their *negative energy*, a consequence most easily demonstrated by bringing to rest the free electron described by Eq. (9.53). In these circumstances $c\rho(\boldsymbol{\sigma}'\cdot\hat{\mathbf{p}})\Psi = 0$ and

$$\beta m_0 c^2 \Psi = E \Psi. \tag{9.79}$$

From (9.44) (remember, $\beta = \alpha_4$) and (9.47), we have

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} m_0 c^2 \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix} = E \begin{pmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{pmatrix}$$

$$\begin{pmatrix} m_0 c^2 \Psi_1 \\ m_0 c^2 \Psi_2 \\ -m_0 c^2 \Psi_3 \\ -m_0 c^2 \Psi_4 \end{pmatrix} = \begin{pmatrix} E_1 \Psi_1 \\ E_2 \Psi_2 \\ E_3 \Psi_3 \\ E_4 \Psi_4 \end{pmatrix}. \tag{9.80}$$

The negative energy solutions are an inevitable consequence of using the correct relativistic expression for the total energy, (9.3), which involves E^2 .

These solutions have some bizarre properties. Where the ‘ordinary’, positive-energy electrons will accelerate under an applied force, the particles described by the negative-energy solutions actually slow down the harder they are pushed. The temptation of most physicists when faced with such unreasonable results is to dismiss them as ‘unphysical’ and continue only with the positive-energy solutions. This might have been acceptable when considering problems in classical physics, in which energy changes are assumed to be continuous and systems that start with positive energy cannot suddenly jump to negative-energy states.

However, quantum mechanics admits exactly this kind of sudden, discontinuous jump and it was perfectly feasible for a positive-energy electron to jump to a negative-energy state. This would appear in the laboratory as a jump from a ‘conventional’ state characterized by the familiar negative electron charge, $-e$, to a state characterized by a very unfamiliar positive electron charge, $+e$. No such transition had ever been observed.

This unusual behaviour left nagging doubts. Heisenberg wrote to Bohr: ‘I am much more unhappy about the question of the relativistic formulation and about the inconsistency of the Dirac theory. Dirac was here [in Leipzig, June 1928] and gave a very fine lecture about his ingenious theory. But he has no more idea than we do about how to get rid of the difficulty $e \rightarrow -e$ ’.¹²

The extra solutions had to be taken seriously, and they were a worry. What could they represent?

Dirac wrestled with what became known as the ‘ \pm ’ problem for the next two years. In December 1929 he outlined a proposed solution. Suppose, he said, that the universe is filled with a ‘sea’ of negative-energy states all occupied by spin-paired electrons. We would have no way of knowing of the existence of such a sea because, when filled,

it wouldn't interact with anything and would merely serve as a kind of backdrop against which positive-energy changes would be measured.

However, if an electron were to be promoted out of the sea—to become an observable, positive-energy electron—it would leave behind a 'hole'. The negative-energy hole would behave exactly like a positive-energy, positively charged particle.

Dirac suggested that the positively charged particle created by the hole in the negative-energy sea was, in fact, a proton. His logic was not without precedent. Rutherford had referred to the nucleus of the hydrogen atom as a 'positive electron' for six years before introducing the term 'proton' in 1920. And making protons out of 'electron holes' had a nice symmetry to it that fed Dirac's desire to find a unitary description of the fundamental constituents of matter.

As he reported to a meeting of the British Association for the Advancement of Science in September 1930:¹³

It has always been the dream of philosophers to have all matter built from one fundamental kind of particle, so that it is not altogether satisfactory to have two in our theory, the electron and the proton.* There are, however, reasons for believing that the electron and the proton are really not independent, but are just two manifestations of one elementary kind of particle.

But Dirac had reached for the dream too soon. His proposal was roundly criticized on all sides because, among other things, it demanded that the masses of the electron and 'hole'-derived proton should be the same. It was already well known that there is a substantial difference in the masses of these particles, the proton heavier by a factor of almost 2000. The debate continued.

Dirac eventually gave up on his dream, finally accepting in 1931 that the hole should have the same mass as the electron. Instead he suggested the existence of the positive electron, 'a new kind of particle, unknown to experimental physics, having the same mass and opposite charge to an electron'.¹⁴ Such particles would come to be known as *antiparticles* which, when combined, produce *antimatter*. American physicist Carl Anderson found evidence for Dirac's positive electron, which he named the *positron*, in cosmic ray experiments in 1932–3.[†]

Quantum Electrodynamics

It would be remiss of me to close this chapter without mentioning that Dirac's 'absolute wonder' would eventually be overtaken by experiment, some 19 years later, and shown to be not quite adequate.

* The neutron had not yet been discovered.

† Anderson initially thought that the particle tracks he had observed in these experiments were due to protons. But protons are too heavy, and in May 1933 he changed his mind and suggested that the tracks were, in fact, due to positrons. These particles were subsequently confirmed to be Dirac's positive electrons by Patrick Blackett and Giuseppe Occhialini.

There were two sources of experimental discomfort. For the configurations $2s^1(^2S_{1/2})$ and $2p^1(^2P_{1/2})$, Dirac's relativistic theory of the hydrogen atom spectrum predicts that these should be degenerate. However, detailed microwave spectroscopy studies by Willis Lamb and Robert Retherford showed in 1947 that there is a small energy difference between these states. This came to be known as the *Lamb shift*.

The second source of trouble lay in the g -factor for electron spin. As we saw above, Dirac's theory predicts $g_e = 2$. But experimental results reported by Isidor Rabi, John Nafe, Edward Nelson, Polykarp Kusch, and H. M. Foley suggested that it is actually slightly larger, more like 2.00244. To be sure, this is a small difference; a little over 0.1 per cent. But such differences were now well outside the bounds of experimental error.

The answer would emerge in the form of *quantum electrodynamics* (QED), a quantum version of relativistic electrodynamics in which wave-particle duality is replaced by the duality of *field* and particle. This was devised in distinctly different mathematical formulations by Richard Feynman, Julian Schwinger, and Sin-Itiro Tomonaga. Freeman Dyson showed in early 1949 that these formulations are equivalent.

The result was a theory that predicts the results of experiments to astonishing levels of accuracy and precision. The g -factor for the electron is predicted by QED to have the value 2.00231930476, with an uncertainty of plus or minus 0.00000000052. The comparable experimental value is 2.00231930482, with an experimental uncertainty of plus or minus 0.00000000040.* 'To give you a feeling for the accuracy of these numbers,' Feynman wrote, 'it comes out something like this: If you were to measure the distance from Los Angeles and New York to this accuracy, it would be exact to the thickness of a human hair.'¹⁵

NOTES

1. Wolfgang Pauli, *Atti del Congresso Internazionale dei Fisici, 11–20 Settembre 1927, Como, Pavia, Roma*, 1928. Quoted in Charles P. Enz, *No Time to be Brief: A Scientific Biography of Wolfgang Pauli*, Oxford University Press, Oxford, 2002, p. 160.
2. Johann Kudar, letter to Paul Dirac, 21 December 1926. Quoted in Helge S. Kragh, *Dirac: A Scientific Biography*, Cambridge University Press, Cambridge, UK, 1990, p. 54.
3. P. A. M. Dirac, *Directions in Physics*, edited by Heinrich Hora and J. R. Shepanski, Wiley, New York, 1978, p. 15. Quoted in Kragh, *Dirac*, p. 56. Dirac's recollections varied, and some sources suggest that his interrupted conversation with Bohr may have taken place in Copenhagen.

* These numbers are subject to constant refinement, both experimental and theoretical. The values quoted here are taken from G. D. Coughlan and J. E. Dodd, *The Ideas of Particle Physics: An Introduction for Scientists*. Cambridge University Press, Cambridge, UK, 1991, p. 34. The value 2.00231930436256(35), where the numbers in brackets represent the uncertainty in the last two digits, is a 2018 figure recommended by the CODATA Task Group—see <http://physics.nist.gov>. The value 2.00231930436146(56) was reported in 2008 by D. Hanneke, S. Fogwell and G. Gabrielse, 'New Measurement of the Electron Magnetic Moment and the Fine Structure Constant', *Physical Review Letters*, **100** (2008), 120801.

4. W. Pauli, Jr., 'On the Quantum Mechanics of the Magnetic Electron', *Zeitschrift für Physik*, 37 (1927), 243.
5. Phil Tee, 'Proof of $(\boldsymbol{\sigma} \cdot \mathbf{A}) \cdot (\boldsymbol{\sigma} \cdot \mathbf{B}) = \mathbf{A} \cdot \mathbf{B} + i\boldsymbol{\sigma} \cdot (\mathbf{A} \times \mathbf{B})$ ', personal communication, 3 December 2018. Thanks, Phil.
6. See, for example, Fulvio Melia, *Electrodynamics*, University of Chicago Press, Chicago, 2001, Chapters 5 and 6.
7. Readers familiar with Dirac's derivation might query Eq. (9.35). In many papers and books, this appears as $p \rightarrow p - (e/c)\mathbf{A}$, and it's right to wonder what has therefore happened to the speed of light which appears in this version. It took me longer than it really should to discover the reason, which simply concerns the choice of *units*. In 'electrostatic units', which are based on the CGS system, the electron charge is 4.803×10^{-10} statcoulombs or Franklins (Fr). We recover the more familiar electron charge by (you guessed it) dividing by the speed of light in centimetres per second ($2.997 \times 10^{10} \text{ cms}^{-1}$). This gives 1.602×10^{-20} abcoulombs (or 'electromagnetic units'), for which 1 abcoulomb = 10 coulombs. So dividing by c returns the electron charge as the much more familiar 1.602×10^{-19} coulombs. In (9.35), I've simply assumed that e is expressed in coulombs, so dividing by c is unnecessary.
8. P. A. M. Dirac, 'The Quantum Theory of the Electron', *Proceedings of the Royal Society of London, Series A*, 117 (1928), 610–24.
9. P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th edn, Clarendon Press, Oxford, 1958, pp. 254–67.
10. Paul Dirac, in C. Weiner (ed.), *Exploring the History of Nuclear Physics*, American Institute of Physics, New York, pp. 109–46. Quoted in Kragh, *Dirac*, p. 59.
11. Léon Rosenfeld, interview with Thomas Kuhn and John Heilbron, July 1963, *Archive for the History of Quantum Physics*. Quoted in Kragh, *Dirac*, p. 62.
12. Werner Heisenberg, letter to Niels Bohr, 23 July 1928. Quoted in Kragh, *Dirac*, p. 66.
13. Paul Dirac, 'The Proton', *Nature*, 126 (1930), 605–6. Quoted in Kragh, *Dirac*, p. 97.
14. Paul Dirac, 'Quantised Singularities in the Electromagnetic Field', *Proceedings of the Royal Society of London, Series A*, 133 (1931), 60–72. Quoted in Kragh, *Dirac*, p. 103.
15. Richard P. Feynman, *QED: The Strange Theory of Light and Matter*, Penguin, London, 1985, p. 7.

.....

FURTHER READING

- Dirac, P. A. M., *The Principles of Quantum Mechanics*, 4th edn, Clarendon Press, Oxford, 1958, Chapter 11.
- Enz, Charles P., *No Time to be Brief: a Scientific Biography of Wolfgang Pauli*, Oxford University Press, Oxford, 2002, Chapter 5, especially pp. 145–6.
- Farmelo, Graham, *The Strangest Man: The Hidden Life of Paul Dirac, Quantum Genius*, Faber & Faber, London, 2009, Chapter 11.
- Griffiths, David J., *Introduction to Quantum Mechanics*, 2nd edn, Cambridge University Press, Cambridge, UK, 2017, Chapter 4, especially pp. 173–7.
- Kragh, Helge S., *Dirac: A Scientific Biography*. Cambridge University Press, 1990, Chapter 6.
- Massimi, Michela, *Pauli's Exclusion Principle: The Origin and Validation of a Scientific Principle*, Cambridge University Press, Cambridge, UK, 2005, pp. 119–22.
- Tomonaga, Sin-Itiro, *The Story of Spin*, translated by Takeshi Oka, University of Chicago Press, Chicago, 1997, Lecture 3.

Dirac, Von Neumann, and the Derivation of the Quantum Formalism

State Vectors in Hilbert Space

The evolution of quantum mechanics through the 1920s was extraordinarily exciting but it was also profoundly messy. As physicists scrambled to make sense of what experiment was telling them about the mechanics of electrons and the inner workings of atoms, they had grabbed hold of whatever classical concepts seemed to fit and had twisted and tortured these to produce a quantum description. The simple truth is that quantum mechanics is like nothing we've ever seen before, yet the only tools that are available to make sense of it come from classical descriptions wholly unsuited to the task.

Planck had somehow managed to wring $E = h\nu$ from a classical description of the entropy of black body radiation. Bohr had made some spectacularly erroneous assumptions to derive the Rydberg formula and establish the notions of quantum numbers and quantum jumps. De Broglie had fused Einstein's special relativity with the Planck-Einstein relation to derive $\lambda = h/p$ and establish the notion of wave-particle duality. Schrödinger had taken de Broglie's idea of 'matter waves' seriously and had derived a non-relativistic wave equation which, when applied to the electron in a hydrogen atom, demonstrated how the quantum numbers emerge naturally as the eigenvalues of the Hamiltonian.

But what were the electron wavefunctions supposed to *represent*? Schrödinger wanted to take them literally as descriptive of an essentially wave-like matter, but Born insisted instead that they are merely calculation devices: the modulus-squares of the wavefunctions give us *probabilities* for different outcomes of quantum events or probabilities for 'finding' the electron in specific physical states. Heisenberg and Bohr interpreted the uncertainty principle as a limitation on what can be *known* about a quantum system.

Through the work of Pauli, Heisenberg, and Dirac, the phenomenon of electron spin was finally accommodated in a proper relativistic treatment of the wave equation, the stability of matter was explained, and antimatter was discovered. But the problem of interpretation remained.

To say that the formulation of quantum mechanics was confused is to indulge in grand understatement. Those physicists attending the Solvay conference in Brussels in 1927 witnessed the opening exchanges of what was to become one of the most famous debates in the entire history of science, as Einstein and Bohr squared off on the interpretation of quantum mechanics.

I promise to return to this debate in Chapters 11 and 12, but for now I want to confront the elephant in the room. By the late 1920s, quantum mechanics was a conceptual mess. What was to be done?

Some physicists believed that in order to make progress it was necessary to throw out much of the conceptual baggage that early quantum mechanics tended to carry around with it and re-establish the theory on much firmer mathematical foundations. It was at this critical stage, perhaps, that the search for deeper insights into the underlying physical reality was set aside in favour of mathematical expediency. Pragmatism demanded a coherent mathematical framework which *worked*. To these physicists, it didn't matter so much that the deeper meaning of the theory's concepts appeared to become increasingly disconnected from the reality that the theory was trying to describe. And yet reality itself, it could be argued, had taken a profoundly bizarre turn.

The heated arguments about the nature and interpretation of the wavefunctions and the need to leave behind the conceptions of classical wave motion hinted at the need for a new structure that did not promote this tendency. In his *Principles of Quantum Mechanics*, first published in 1930, Dirac set out a powerful new structure designed deliberately to look nothing like its intellectual predecessors. He wrote:¹

[Quantum mechanics] requires the states of a dynamical system and the dynamical variables to be interconnected in quite strange ways that are unintelligible from a classical standpoint. The states and dynamical variables have to be represented by mathematical quantities of different natures from those ordinarily used in physics.

State Vectors and Vector Spaces

What was needed was a mathematical structure which captured all the quantum-mechanical properties and relationships that had thus far been uncovered. Although rather novel, there was no real issue with the notion of deriving physical observables as the expectation values of their corresponding quantum-mechanical operators, Eq. (7.11), and physically real observables would still imply Hermitian operators according to (7.13). The non-commutativity of the operators for complementary observables would still imply an uncertainty relation according to Eq. (7.43).

The operators were not the problem. Rather, all the conceptual problems were coming from the substrate: from the things that the operators were supposed to be operating *on*.

But if the concept of the wavefunction was to be abandoned—at least in the way that I have used it so far in this book—then whatever was to replace it still needed to possess all the properties and exhibit all the relationships that had so far been discovered. This included the ability to form superpositions (7.5), and the properties of orthogonality and

normalization (6.7). The modulus-square of whatever was to replace the wavefunction must still relate to a quantum probability of some kind.

There were some important clues in the relationship between Schrödinger's wave mechanics and Heisenberg's matrix mechanics.

In 1926 the Hungarian mathematician John von Neumann had attended a lecture on matrix mechanics given by Heisenberg in Göttingen. The great mathematician David Hilbert had also been in the audience, and after the lecture he expressed personal difficulties in comprehending the significance of matrix mechanics and its relationships with wave mechanics. Von Neumann volunteered an explanation in a language that Hilbert could understand. Whilst the matrix and wave mechanical formalisms could be connected by a set of rules within a general 'transformation theory' (as Schrödinger himself had demonstrated), von Neumann believed this equivalence belied a deeper, more fundamental mathematical framework.

The problem was that the two quantum descriptions were so very different in their approaches. Matrix mechanics is a theory in a 'space' of index values identifying the locations of elements within the theory's abstract tables of numbers.* Wave mechanics is a theory in an abstract multidimensional 'configuration space'. The theories could not easily be brought together. Von Neumann wrote:²

That this cannot be achieved without some violence to the formalism and to mathematics is not surprising. The spaces . . . are in reality very different, and every attempt to relate the two must run into great difficulties.

However, von Neumann argued, the spaces themselves are not so important. It is the *relation* between functions in the two spaces that is essential for the physical interpretation of the theory's predictions. Von Neumann showed that these functions are *isomorphic*—it is possible to establish a direct, one-to-one correspondence between them: they 'are identical in their intrinsic structure (they realise the same abstract properties in different mathematical forms)—and since they . . . are the real analytical substrata of the matrix and wave theories, this isomorphism means that the two theories must always yield the same numerical results'.³

We should recall from Chapter 9 that we introduced two- and four-dimensional column matrices to represent the spin states of an electron (and positron). These column matrices and the matrices of matrix mechanics act in much the same way as *vectors* in classical physics.

And here is at least part of the answer. The wavefunctions should be replaced not with another set of functions or matrices, but rather with a particular set of *state vectors* which represent the physical states of quantum objects. Obviously, classical vectors operating in ordinary three-dimensional Euclidean space can't meet our requirements.

* I've put 'space' in inverted commas because we need to be clear that this does *not* refer to the familiar, three-dimensional Euclidean space of everyday experience. The term is being used here as a shorthand for all the *dimensions* within which a set of mathematical objects (such as matrix elements) can be said to 'exist' and—more meaningfully—within which they can be manipulated.

Instead, the quantum-mechanical state vectors operate in an abstract, possibly complex, multidimensional, or even infinite-dimensional *vector space*.

The paper that von Neumann drafted in 1926 to explain the relationships between matrix and wave mechanics to Hilbert was subsequently to become the *Mathematical Foundations of Quantum Mechanics*, first published in German in 1932.

Thus, both Dirac and von Neumann realized that a structure based on vector spaces frees us from the mental prison of ordinary three-dimensional space. Such a structure allows us to define an abstract space with whatever number and nature of dimensions we need to suit our purposes, including complex dimensions. By definition, these dimensions are not constrained to be visualizable, for example in a Cartesian coordinate system.

A vector space is defined by the rules for combining the vectors that constitute it, specifically those for forming inner, scalar, or 'dot' products, and for multiplying the vectors by scalar quantities. The combining rules chosen to create a new quantum mechanics mirror those of the wavefunctions of standard wave mechanics but cannot be derived from them. Rather, they are *assumed* and later proved to be consistent through agreement between prediction and experiment. The vector space can be complex, yet it is intended to provide a very simple, straightforward, and intuitive connection with quantum probabilities. All the manipulations required to yield the values of observable quantities in matrix and wave mechanics can be reproduced in a quantum mechanics based on vector spaces.

Despite the apparent flexibility achieved through the introduction of these concepts, the requirements of quantum mechanics and the connection with real quantities measurable in the real world must still be satisfied. This is particularly important in situations where an infinite number of dimensions is required. In other words, the results of calculations involving an infinite series of terms must *converge*. Vector spaces with the right kinds of convergence properties are singled out for special attention and are referred to as *Hilbert spaces*.

It is easy to get lost in this kind of mathematical abstraction, but here's how I try to think about it. In the Prologue I explained how switching to a phase space description in Hamiltonian mechanics allows us to represent the motions of an entire system consisting of many moving objects as the single trajectory of a point in a multidimensional configuration space. Now, phase space is a mathematical abstraction, and in this sense it is 'unreal'. But it is constructed from the real positions and momenta of multiple, physically real objects.

It might therefore be helpful to acknowledge that Hilbert space is to quantum mechanics what phase space is to classical Hamiltonian mechanics.*

Hilbert's Sixth Problem

If we are to replace the wavefunctions of Schrödinger's wave mechanics with state vectors in an abstract vector space, then to formulate a new quantum representation we still need

* Alas, this should not be taken to imply that there exists a more fundamental reality for which Hilbert space is a mathematical abstraction.

a second key ingredient. We need some rules for how these vectors should be manipulated, how they should be combined, how they are operated on to deduce observables, how they relate to quantum probability, and how we expect them to evolve in time.

The source of this second ingredient can also be traced back to the influence of Hilbert. In a lecture delivered to the International Congress of Mathematicians in Paris in 1900, Hilbert outlined a long list of key problems that he believed would occupy the ‘leading mathematical spirits of coming generations’. The list has become known as *Hilbert’s problems*.⁴

The sixth of these concerns the mathematical treatment of the axioms of physics. Drawing on the foundations of geometry as an example, Hilbert stated that an important goal for future mathematicians would be ‘to treat in the same manner, by means of axioms, those physical sciences in which mathematics plays an important part’. He went on to say:

If geometry is to serve as a model for the treatment of physical axioms, we shall try first by a small number of axioms to include as large a class as possible of physical phenomena, and then by adjoining new axioms to arrive gradually at the more special theories.

Axioms are self-evident truths that are not themselves in need of further proof, and represent the foundation stones of the mathematical structure that can be derived from them. The introduction of the axiomatic method by Hilbert and his disciples was manifested as an almost pathological drive to eliminate any form of intuitive reasoning from mathematics, arguing that the subject was far too important for its truths to be anything other than ‘hard-wired’ from its axioms through to its theorems (or true statements).

This drive for mathematical rigour and formality inevitably resulted in a disconcerting increase in obscure symbolism and mathematical abstraction contributing, according to the philosopher Roland Omnès, to the ‘fracture’: the near impossibility for anyone of average intelligence but without formal training in mathematics or logic to fully comprehend aspects of modern fundamental science.⁵

So, here was the complete answer. To derive a new way of formulating quantum mechanics, we replace the wavefunctions with a substrate of state vectors in an abstract vector space and formally embed all the most important definitions and relations we need within a system of axioms. Now, Dirac and von Neumann took conspicuously different approaches in their classic texts, so here I propose to present the axioms as you would expect to find them in a contemporary textbook.

For the purposes of this chapter, the list of axioms replaces the usual list of mathematical ingredients, and a brief exploration of the formalism replaces the recipe.

The Axioms

1. The state of a quantum-mechanical system is completely defined by its state vector.
2. Observables are represented by a specific class of mathematical operators that act on the relevant Hilbert space.

3. The mean value of an observable is given by the expectation value of its corresponding operator.
4. The probability of obtaining the value of an observable is given by the modulus-square of the state vector (this is the Born rule).
5. In a closed system with no external influences, the state vector evolves in time according to the time-dependent Schrödinger equation.

I know what you're thinking. These are not like the axioms of Euclidean geometry, which are concerned with the self-evident properties of straight lines, circles, right angles, and parallel lines. The axioms of quantum mechanics are hardly self-evident. But we can accept them as axioms if we adopt a modern interpretation of the term, as propositions belonging to a formal language (a set of symbols and a set of rules for combining the symbols to form proposition statements) which are assumed to be true by hypothesis, and which are to be justified through agreement between prediction and experience.

Incidentally, Axiom 1 is quite provocative. The debate that was begun by Einstein and Bohr in 1927 continues to this day, and largely hinges on whether quantum mechanics is to be regarded as a complete theory. Clearly, if it is to be self-consistent and mathematically rigorous, any new quantum formalism *must* be considered to be complete, and Axiom 1 makes this a given. We'll be revisiting this question with a vengeance in Chapter 12.

The Formalism

I still think it's worth taking this in small steps. In Step (1), I will introduce a notation that was devised by Dirac in 1939 and which greatly simplifies the new formalism, and I will use this to explore a few of the basic mathematical properties of the state vectors. In Step (2) we explore the actions of operators corresponding to the physical observables, and the rules we need to use for combining the state vectors will start to become apparent as we manipulate them. In Step (3), we explore the properties of superposition states and use these to derive so-called projection amplitudes and projection operators. This leads us in Step (4) to the introduction of the Born rule through Axiom 4, which allows us to deduce quantum probabilities from the modulus-squares of the projection amplitudes and to define something called the probability density matrix.

If all this seems really rather too abstract, we can try to bring it back down to earth in Step (5) by making a direct comparison between the state vectors of quantum mechanics and classical unit vectors. Finally, in Step (6) we see how to handle the time dependence of the state vectors through the introduction of the time evolution operator.

Step (1): Introduce the Dirac 'Bracket' Notation

Dirac first published *The Principles of Quantum Mechanics* in 1930, and nine years later he devised a notation that would greatly simplify the new quantum formalism

and which has since become a staple of undergraduate courses on the subject. His motivation was based on the recognition that ‘a good notation can be of great value in helping the development of a theory, by making it easy to write down those quantities or combinations of quantities that are important, and difficult or impossible to write down those that are unimportant’.⁶ He seems initially to have kept this notation pretty much to himself and it did not receive wider acknowledgement until the publication of the third edition of *The Principles* in 1947.

In this new notation, Dirac replaced the wavefunctions of wave mechanics with a very different set of mathematical objects, which he called *ket vectors*, which is just a particular form for the more general state vector. The ket vector is assumed to represent the state of a quantum dynamical system at a particular time. So, for example, instead of the spatial wavefunction ψ_n , we write the ket vector as $|n\rangle$.^{*} Here the use of n can be taken to indicate that the state depends on one or more quantum numbers, or it can simply be used as a convenient label to distinguish different states. Many textbooks suggest that $|n\rangle$ can be thought to be ‘equivalent to’ ψ_n , but this is somewhat misleading. For one thing, ψ_n is generally interpreted as a scalar function, not a vector.[†]

So, according to Axiom 1, the state of a quantum-mechanical system is completely described by the ket vector $|n\rangle$.

Remember that we don’t really *deduce* the mathematical properties of the ket vectors directly from the physics. Instead, we *assume* those properties that prove to be consistent with the way that the formalism needs to work if it is to produce the right kinds of predictions for the physics.

Although von Neumann used a different notation in *Mathematical Foundations*, he opens Chapter 2 with a *definition* of abstract Hilbert space in terms of the properties of the vectors themselves. We’ll do the same here, but using Dirac’s notation. So, if we add together two different ket vectors $|m\rangle$ and $|n\rangle$, we get another ket vector:

$$|m\rangle + |n\rangle = |o\rangle. \quad (10.1)$$

Such addition is both *commutative*,

$$|m\rangle + |n\rangle = |n\rangle + |m\rangle, \quad (10.2)$$

and *associative*,

$$(|m\rangle + |n\rangle) + |o\rangle = |m\rangle + (|n\rangle + |o\rangle). \quad (10.3)$$

We can define a ‘null vector’, $|0\rangle$, such that $|m\rangle + |0\rangle = |m\rangle$, and $|0\rangle - |m\rangle = -|m\rangle$.

^{*} Don’t be put off by the strange use of the vertical line and angle bracket to symbolize the vector—the reasons for this particular notation will soon become clear.

[†] Inside the $| \rangle$ notation we can use whatever symbol we think is most appropriate for the problem we’re trying to solve, and many physicists will quite happily use the Greek symbol ψ , writing the vector as $|\psi\rangle$. I’ve done this myself in other books and on many other occasions. However, I’m going to avoid this kind of notation in this book, for the simple reason that we often can’t help thinking of ψ as a *wavefunction*, and it’s always helpful to avoid confusion.

The product of a scalar quantity (which can be complex) and a ket vector is another ket vector, for example,

$$|zm\rangle = z|m\rangle = |m\rangle z = |n\rangle, \quad (10.4)$$

where z can be a complex number.

Multiplication by a scalar quantity is *distributive*:

$$\begin{aligned} z(|m\rangle + |n\rangle) &= z|m\rangle + z|n\rangle \text{ and} \\ (z + z')|n\rangle &= z|n\rangle + z'|n\rangle. \end{aligned} \quad (10.5)$$

So far, this doesn’t seem so very different from the properties of familiar classical vectors, but there is one important exception. Classical vectors are defined in a *real* three-dimensional vector space (typically defined along Cartesian coordinates x, y, z). There is no prescription for multiplying them by complex scalar quantities, because such a prescription simply isn’t needed. In contrast, we *require* the vector space in quantum mechanics to accommodate the possibility of complex vectors.

The formalism starts to get really interesting when we look at how we need to form inner, scalar, or ‘dot’ products of these vectors (which for simplicity I will call inner products in what follows). To do this we define a *bra vector*, which we write as $\langle n|$. It’s tempting to think of this simply as the complex conjugate of the ket vector $|n\rangle$, but we will soon see that there’s a little more to it than this.

The inner product is then formed by multiplying the bra vector from the left, thus combining bra and ket vectors into a ‘bracket’. For example, multiplying $\langle m|$ and $|n\rangle$ gives the bracket $\langle m|n\rangle$. Interchanging the labels is then equivalent to taking the complex conjugate of the product: $\langle n|m\rangle = \langle m|n\rangle^*$.

The bra vectors have the same mathematical properties as their associated ket vectors—summarized earlier in Eqs. (10.1)–(10.5)—with one important exception. The bra corresponding to the vector $z|n\rangle$ is not $\langle n|z$, but rather $\langle n|z^*$, where z^* is the complex conjugate of z .

Multiplication by bra vectors is also distributive:

$$\begin{aligned} \langle o|(|m\rangle + |n\rangle) &= \langle o|m\rangle + \langle o|n\rangle \text{ and} \\ (\langle m| + \langle n|)|o\rangle &= \langle m|o\rangle + \langle n|o\rangle. \end{aligned} \quad (10.6)$$

A ket vector is said to be *normalized* if the inner product formed with its bra is equal to 1, e.g. $\langle n|n\rangle = 1$. Two different ket vectors are said to be *orthogonal* if their inner product is equal to 0, e.g. $\langle m|n\rangle = 0$. Thus, the vectors are *orthonormal* if

$$\langle m|n\rangle = \delta_{mn}, \quad (10.7)$$

where the Kronecker delta $\delta_{mn} = 1$ when $m = n$ and 0 when $m \neq n$.

We can see that this reflects the properties of the quantum wavefunctions that we discovered in Chapters 5–9, and in his 1939 paper Dirac wrote:⁷

Two general rules in connexion with the new notation may be noted, namely, *any quantity in brackets $\langle \ \rangle$ is a number, and any expression containing an unclosed bracket symbol \langle or \rangle is a vector in Hilbert space, of the nature of [a wavefunction].*

But if we persist with the notion that $|n\rangle$ is ‘equivalent to’ or ‘of the nature of’ ψ_n , then we must be careful to admit that $\langle n|$ cannot be equivalent to ψ_n^* . Rather, its equivalence with quantities derived in wave mechanics is only manifest when we combine the bra and ket vectors together. The inner products $\langle m|n\rangle$ and $\langle n|n\rangle$ are then ‘equivalent to’ the integrals $\int \psi_m^* \psi_n d\tau$ and $\int \psi_n^* \psi_n d\tau$, respectively.

I don’t want to say any more about this here, but readers interested in the significance and formal definition of $\langle n|$ should consult Chris Isham’s *Lectures on Quantum Theory*.⁸

Step (2): Explore the Actions of Operators Corresponding to Observables

So let’s now apply an operator, \hat{A} , which we assume relates to some physical observable, such as linear momentum or energy. If $|n\rangle$ is an *eigenstate* of \hat{A} , we know from Schrödinger’s wave mechanics that the action of \hat{A} *must* be to return an eigenvalue, which we will write as a_n . We obviously want the same from the state vectors, so

$$\hat{A}|n\rangle = a_n|n\rangle. \quad (10.8)$$

Multiplying from the left by the bra vector $\langle n|$ then gives

$$\langle n| \cdot (\hat{A}|n\rangle) = \langle n|\hat{A}|n\rangle = \langle n|a_n|n\rangle = a_n \langle n|n\rangle. \quad (10.9)$$

The product $\langle n|\hat{A}|n\rangle$ (which is ‘equivalent to’ $\int \psi_n^* \hat{A} \psi_n d\tau$) is the expectation value $\langle \hat{A}_n \rangle$, so we can see from (10.9) that, if $|n\rangle$ is normalized, then the expectation value is *exactly* the eigenvalue a_n .

But if a_n is to represent a physical observable, then it must be an exclusively real quantity. To see what this means for the operator \hat{A} we take the Hermitian conjugate of (10.9) (cf. Eq. (7.15)):

$$\begin{aligned} [\langle n| \cdot (\hat{A}|n\rangle)]^* &= (\hat{A}|n\rangle)^* \cdot |n\rangle \\ &= (a_n|n\rangle)^* |n\rangle = (\langle n| a_n^*) |n\rangle = a_n^* \langle n|n\rangle = a_n^*. \end{aligned} \quad (10.10)$$

If a_n is to be exclusively real, then $a_n = a_n^*$ and so

$$\langle n| \cdot (\hat{A}|n\rangle) = [\langle n| \cdot (\hat{A}|n\rangle)]^*. \quad (10.11)$$

In other words, \hat{A} must be an Hermitian or self-adjoint operator. This is what is meant by ‘special class’ in Axiom 2.

Of course, in this formalism we need to retain everything we’ve learned about the operators themselves from the preceding chapters. Specifically, the operators must satisfy all the various commutation relations (which can also be taken as axioms and assumed without proof), such as $[x, \hat{p}_x] = i\hbar$, satisfied by assuming $\hat{p}_x = -i\hbar\partial/\partial x$. This means that the derivation of Heisenberg’s uncertainty principle is preserved (Chapter 7) and very much part of the new structure.

Step (3): Explore the Properties of Superposition States

In order to appreciate the need for the term ‘mean value’ in Axiom 3, let’s form a superposition of two ket vectors $|m\rangle$ and $|n\rangle$, both of which are eigenstates of \hat{A} with eigenvalues a_m and a_n .*

$$|p\rangle = c_m|m\rangle + c_n|n\rangle. \quad (10.12)$$

Here $|p\rangle$ is the resultant state and c_m and c_n are expansion coefficients (cf. Eq. (7.8)). Applying the operator now gives

$$\begin{aligned} \hat{A}|p\rangle &= c_m\hat{A}|m\rangle + c_n\hat{A}|n\rangle \\ &= c_ma_m|m\rangle + c_na_n|n\rangle. \end{aligned} \quad (10.13)$$

We now define the bra vector, $\langle p|$, as

$$\langle p| = \langle m|c_m^* + \langle n|c_n^*, \quad (10.14)$$

such that

$$\begin{aligned} \langle p|\hat{A}|p\rangle &= (\langle m|c_m^* + \langle n|c_n^*) (c_ma_m|m\rangle + c_na_n|n\rangle) \\ &= c_m^*c_ma_m\langle m|m\rangle + c_m^*c_na_n\langle m|n\rangle \\ &\quad + c_n^*c_ma_m\langle n|m\rangle + c_n^*c_na_n\langle n|n\rangle. \end{aligned} \quad (10.15)$$

If $|m\rangle$ and $|n\rangle$ are eigenstates of \hat{A} , then $\langle m|n\rangle = \delta_{mm}$ —Eq. (10.7). The product $\langle p|\hat{A}|p\rangle$ is the expectation value $\langle \hat{A}_p \rangle$ so, just as before (Eq. (7.11)), we have

$$\langle \hat{A}_p \rangle = \langle p|\hat{A}|p\rangle = |c_m|^2a_m + |c_n|^2a_n \quad (10.16)$$

* For simplicity, I’m limiting this to just two ket vectors. In general, $|p\rangle = \sum_j c_j |j\rangle$, where the sum extends over all vectors required to represent the superposition state.

If $|p\rangle$ is normalized, so that $\langle p|p\rangle = 1$, then

$$\begin{aligned}\langle p|p\rangle &= (\langle m|c_m^* + \langle n|c_n^*) (c_m|m\rangle + c_n|n\rangle) \\ &= |c_m|^2 \langle m|m\rangle + c_m^* c_n \langle m|n\rangle + c_n^* c_m \langle n|m\rangle + |c_n|^2 \langle n|n\rangle \\ &= |c_m|^2 + |c_n|^2 = 1.\end{aligned}\tag{10.17}$$

And so we see that the expectation value in (10.16) really is the mean of the two eigenvalues, each weighted by the modulus-square of the corresponding expansion coefficient.

At this point it's interesting to note what happens if we form the inner product of the vectors $\langle m|$ and $|p\rangle$, from Eq. (10.12):

$$\langle m|p\rangle = c_m \langle m|m\rangle + c_n \langle m|n\rangle = c_m.\tag{10.18}$$

Likewise,

$$\langle n|p\rangle = c_m \langle n|m\rangle + c_n \langle n|n\rangle = c_n.\tag{10.19}$$

This means that we can rewrite (10.12) as*

$$\begin{aligned}|p\rangle &= (\langle m|p\rangle) |m\rangle + (\langle n|p\rangle) |n\rangle \\ &= |m\rangle \langle m|p\rangle + |n\rangle \langle n|p\rangle.\end{aligned}\tag{10.20}$$

The inner products $\langle m|p\rangle$ and $\langle n|p\rangle$ are sometimes referred to as *projection amplitudes*. They provide a measure of the extent of the projection of one vector onto another, in this case of $|p\rangle$ on $|m\rangle$ and $|p\rangle$ on $|n\rangle$.

The inner products, just like the coefficients themselves, are scalar quantities ('numbers', as Dirac explained in the quote earlier), allowing us to switch the order of multiplication in Eq. (10.20). I've done this specifically to highlight the 'outer' products $|m\rangle \langle m|$ and $|n\rangle \langle n|$, which are called 'butterfly' operators (for obvious reasons) or, more meaningfully, *projection operators*. For example, when we apply $|n\rangle \langle n|$ to an arbitrary vector, the leading bra vector $\langle n|$ forms an inner product (or 'dots into' the vector), giving the projection amplitude of the arbitrary vector on $|n\rangle$. The trailing $|n\rangle$ then turns this back into a vector, and so it 'maps' the arbitrary vector onto $|n\rangle$.

The projection operators thus map a vector onto a subspace of the Hilbert space determined by the eigenstates ($|m\rangle$ and $|n\rangle$ in this example), and possess just two eigenvalues, 0 or 1.

To see how this works, let's take a quick look at the action of the projection operator $\hat{P}_n = |n\rangle \langle n|$ on some arbitrary vector $|a\rangle$,

$$\hat{P}_n |a\rangle = |n\rangle \langle n|a\rangle = p_n |a\rangle,\tag{10.21}$$

* Again, more generally $|p\rangle = \sum_j |j\rangle \langle j|p\rangle$.

where p_n is assumed to be the corresponding eigenvalue. If we now apply \hat{P}_n a second time, we get

$$\hat{P}_n^2|a\rangle = |n\rangle\langle n|n\rangle\langle n|a\rangle = |n\rangle\langle n|a\rangle = p_n|a\rangle. \quad (10.22)$$

Obviously, if $\hat{P}_n^2|a\rangle = \hat{P}_n|a\rangle$, then the only possible values for p_n are 0 or 1.

Step (4): Introduce the Born Rule (Axiom 4)

Importing the Born rule through Axiom 4 allows us to translate the modulus-squares of the expansion coefficients or projection amplitudes into quantum probabilities. If we ask what is the probability of finding the state $|n\rangle$ in the superposition $|p\rangle$, the answer is $|c_n|^2$ or $|\langle n|p\rangle|^2$. These are called *projection probabilities*.

In fact, Eqs. (10.15) and (10.17) are suggestive of a matrix formulation (though you shouldn't imagine that we've somehow drifted into matrix mechanics). If we define the *probability density matrix* as

$$\hat{\rho} = \begin{pmatrix} |c_m|^2 & c_m^*c_n \\ c_n^*c_m & |c_n|^2 \end{pmatrix} = \begin{pmatrix} |\langle m|p\rangle|^2 & \langle m|p\rangle^*\langle n|p\rangle \\ \langle n|p\rangle^*\langle m|p\rangle & |\langle n|p\rangle|^2 \end{pmatrix}, \quad (10.23)$$

then the *trace* of this matrix (see Appendix 7, Eq. (A7.2)) sums to 1: $\text{Tr}(\hat{\rho}) = |c_m|^2 + |c_n|^2 = 1$. Furthermore, if we set up the product $\langle p|\hat{A}|p\rangle$ also as a matrix with elements $\langle m|\hat{A}|m\rangle$, $\langle m|\hat{A}|n\rangle$, etc., the expectation value $\langle\hat{A}_p\rangle$ can be derived as the trace of the product of $\hat{\rho}$ and this operator matrix:

$$\langle\hat{A}_p\rangle = \text{Tr}(\hat{\rho}\hat{A}_p) = \text{Tr}\left[\begin{pmatrix} |c_m|^2 & c_m^*c_n \\ c_n^*c_m & |c_n|^2 \end{pmatrix} \begin{pmatrix} \langle m|\hat{A}|m\rangle & \langle m|\hat{A}|n\rangle \\ \langle n|\hat{A}|m\rangle & \langle n|\hat{A}|n\rangle \end{pmatrix}\right]. \quad (10.24)$$

We know that if $|m\rangle$ and $|n\rangle$ are eigenstates of \hat{A} , then the diagonal matrix elements $\langle m|\hat{A}|m\rangle = a_m$ and $\langle n|\hat{A}|n\rangle = a_n$, and the off-diagonal elements reduce to 0. Multiplying the matrices and taking the trace then gives (see Eq. (A7.7)):

$$\begin{aligned} \langle\hat{A}_p\rangle &= \text{Tr}(\hat{\rho}\hat{A}_p) = \text{Tr}\left[\begin{pmatrix} |c_m|^2 & c_m^*c_n \\ c_n^*c_m & |c_n|^2 \end{pmatrix} \begin{pmatrix} a_m & 0 \\ 0 & a_n \end{pmatrix}\right] \\ &= \text{Tr}\begin{pmatrix} |c_m|^2 a_m & c_n^*c_m a_n \\ c_m^*c_n a_m & |c_n|^2 a_n \end{pmatrix} \\ &= |c_m|^2 a_m + |c_n|^2 a_n \end{aligned} \quad (10.25)$$

as before.

Step (5): Compare the State Vectors with Classical Unit Vectors

We can now get some sense for what the state vectors represent by comparing their properties with those of classical unit vectors. Suppose the classical unit vector \mathbf{p} is confined to the xy -plane. This means we can express it in terms of unit vectors \mathbf{i} and \mathbf{j} defined along the x - and y -coordinates,

$$\mathbf{p} = p_x \mathbf{i} + p_y \mathbf{j}, \quad (10.26)$$

where p_x and p_y are the projections of \mathbf{p} along the x - and y -coordinates. The values of these coefficients can be determined from the inner products,

$$\begin{aligned} p_x &= (\mathbf{i} \cdot \mathbf{p}) = |\mathbf{i}||\mathbf{p}| \cos \theta = \cos \theta \\ p_y &= (\mathbf{j} \cdot \mathbf{p}) = |\mathbf{j}||\mathbf{p}| \cos(90^\circ - \theta) = |\mathbf{j}||\mathbf{p}| \sin \theta = \sin \theta, \end{aligned} \quad (10.27)$$

where θ is the angle between \mathbf{p} and the x -coordinate, and $|\mathbf{p}| = |\mathbf{i}| = |\mathbf{j}| = 1$, as these are defined as unit vectors. Let's now make a few comparisons.

The unit vectors are orthogonal:

$$\begin{aligned} \text{Classical : } (\mathbf{i} \cdot \mathbf{j}) &= \cos 90^\circ = 0 \\ \text{Quantum : } \langle m|n \rangle &= 0. \end{aligned} \quad (10.28)$$

The unit vectors are normalized:

$$\begin{aligned} \text{Classical : } (\mathbf{i} \cdot \mathbf{i}) &= \cos 0^\circ = 1 \\ \text{Quantum : } \langle m|m \rangle &= 1. \end{aligned} \quad (10.29)$$

The unit vectors provide a basis for representing other vectors:

$$\begin{aligned} \text{Classical : } \mathbf{p} &= p_x \mathbf{i} + p_y \mathbf{j} = \mathbf{i}(\mathbf{i} \cdot \mathbf{p}) + \mathbf{j}(\mathbf{j} \cdot \mathbf{p}) \\ \text{Quantum : } |\mathbf{p}\rangle &= c_m |m\rangle + c_n |n\rangle = |m\rangle \langle m|\mathbf{p}\rangle + |n\rangle \langle n|\mathbf{p}\rangle. \end{aligned} \quad (10.30)$$

The unit vectors provide a complete representation:

$$\begin{aligned} \text{Classical : } |\mathbf{i} \cdot \mathbf{p}|^2 + |\mathbf{j} \cdot \mathbf{p}|^2 &= \cos^2 \theta + \sin^2 \theta = 1 \\ \text{Quantum : } |\langle m|\mathbf{p}\rangle|^2 + |\langle n|\mathbf{p}\rangle|^2 &= 1. \end{aligned} \quad (10.31)$$

We can see immediately from this that the *quantum-mechanical state vectors are the unit vectors of the abstract Hilbert space.*

Step (6): Introduce the Time Evolution Operator

In Axiom 5 we *assume* the validity of the time-dependent Schrödinger equation without concerning ourselves with questions about its derivation. So, we adapt Eq. (5.28) and rewrite it by replacing the time-dependent wavefunction $\Psi(r, t)$ with a time-varying state vector, which we denote as $|n_t\rangle$:

$$\hat{H} |n_t\rangle = i\hbar \frac{\partial}{\partial t} |n_t\rangle. \quad (10.32)$$

If we further assume that the Hamiltonian operator for the quantum system under study is time-independent, then (10.32) can be integrated to give

$$|n_t\rangle = e^{-i\hat{H}t/\hbar} |n_0\rangle, \quad (10.33)$$

where $|n_0\rangle$ is the state vector at time $t = 0$. We can quickly confirm this by differentiating (10.33) with respect to time:

$$\frac{\partial}{\partial t} |n_t\rangle = \frac{\partial}{\partial t} (e^{-i\hat{H}t/\hbar} |n_0\rangle) = -\frac{i}{\hbar} \hat{H} e^{-i\hat{H}t/\hbar} |n_0\rangle = -\frac{i}{\hbar} \hat{H} |n_t\rangle. \quad (10.34)$$

Remember, $|n_0\rangle$ represents the state of the system at $t = 0$ and is therefore time-independent. Multiplying (10.34) through by $i\hbar$ and rearranging returns us to (10.32).

The appearance of the Hamiltonian operator in the exponent might cause some concern about how the operator should be applied, but don't forget that we can expand $e^{-i\hat{H}t/\hbar}$ as a Taylor series:

$$e^{-i\hat{H}t/\hbar} = 1 - \frac{i}{\hbar} \hat{H}t - \frac{1}{2!\hbar^2} \hat{H}^2 t^2 + \dots \quad (10.35)$$

If we now define the *time evolution operator* \hat{U} as $\hat{U} = e^{-i\hat{H}t/\hbar}$, this allows us to write (10.33) succinctly as

$$|n_t\rangle = \hat{U} |n_0\rangle. \quad (10.36)$$

Note that \hat{U} is not an 'operator for time', as we already established in our discussion in Chapter 7 (see Step (5) of Chapter 7's recipe) that no such operator exists. It is rather an operator that governs the time dependence of the state vector $|n_t\rangle$.

There are a couple of useful things to note about \hat{U} before we move on. It is said to be a *unitary operator*, meaning that it preserves any inner product that it operates on, in the same way that a unit matrix will not affect whatever it is multiplied with. For example, suppose we apply \hat{U} to a couple of arbitrary vectors $|a_t\rangle$ and $|b_t\rangle$, which form an inner product $\langle a_t | b_t \rangle$. We start by forming the inner product of $\hat{U}|a_t\rangle$ and $\hat{U}|b_t\rangle$ (remember,

the bra vector equivalent to $\hat{U}|a_t\rangle$ is $\langle a_t|\hat{U}^*$, where \hat{U}^* is the complex conjugate of \hat{U}):

$$(\langle a_t|\hat{U}^*) \cdot (\hat{U}|b_t\rangle) = \langle a_t|\hat{U}^*\hat{U}|b_t\rangle. \tag{10.37}$$

But we know that

$$\hat{U}^*\hat{U} = e^{+i\hat{H}t/\hbar}e^{-i\hat{H}t/\hbar} = 1. \tag{10.38}$$

So

$$\langle a_t|\hat{U}^*\hat{U}|b_t\rangle = \langle a_t|b_t\rangle \tag{10.39}$$

and the inner product is indeed preserved unchanged. In other words, whatever it is, the *relationship* between the vectors $|a_t\rangle$ and $|b_t\rangle$ is conserved through time.

More generally, $\hat{U}^\dagger\hat{U} = 1$, where \hat{U}^\dagger is the complex conjugate transpose of \hat{U} . This applies to situations where we might express the Hamiltonian operator in a matrix form, demanding a corresponding matrix structure for \hat{U} . In fact, for any Hermitian operator \hat{A} , $e^{i\hat{A}}$ is unitary (though it's worth noting that not all unitary operators are Hermitian).

The unitary nature of the time evolution operator is profoundly important. It means that in the absence of any external influence on the state vector $|n_t\rangle$, then we expect it to evolve smoothly and continuously in time in a manner entirely determined by the Hamiltonian of the system and the initial state, $|n_0\rangle$, according to the time-dependent Schrödinger equation. Nowhere in this prescription is there anything to suggest the kind of instantaneous, discontinuous, and indeterministic transition from one state to another that we associate with a quantum ‘jump’.

This is going to cause some considerable problems, as we will soon see.

.....

NOTES

1. P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th edn, Clarendon Press, Oxford, 1958, p. 16.
2. John von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton, NJ, 1955, p. 28.
3. Ibid. 31.
4. David Hilbert, ‘Mathematical Problems’, *Bulletin of the American Mathematical Society*, 8 (1902), 437–79 (English translation by Mary Winton Newson).
5. Roland Omnès, *Quantum Philosophy: Understanding and Interpreting Contemporary Science*, Princeton University Press, Princeton, NJ, 1999, pp. 81–3.
6. P. A. M. Dirac, ‘A New Notation for Quantum Mechanics’, *Proceedings of the Cambridge Philosophical Society*, 35 (1939), 416. Quoted in Helge S. Kragh, *Dirac: A Scientific Biography*, Cambridge University Press, Cambridge, UK, 1990, p. 177.
7. Dirac, ‘A New Notation for Quantum Mechanics’, 418.
8. Chris J. Isham, *Lectures on Quantum Theory: Mathematical and Structural Foundations*, Imperial College Press, London, p. 35.

FURTHER READING

- Atkins, Peter, and Friedman, Ronald, *Molecular Quantum Mechanics*, 5th edn, Oxford University Press, Oxford, 2011, Chapter 1.
- D’Espagnat, Bernard, *Conceptual Foundations of Quantum Mechanics*, 2nd edn, Addison-Wesley, New York, 1989, Chapters 1–3.
- Dirac, P. A. M., *The Principles of Quantum Mechanics*, 4th edn, Clarendon Press, Oxford, 1958, Chapters 1–3.
- Griffiths, David J., *Introduction to Quantum Mechanics*, 2nd edn, Cambridge University Press, Cambridge, UK, 2017, Chapter 3.
- Isham, Chris J., *Lectures on Quantum Theory: Mathematical and Structural Foundations*, Imperial College Press, London, Chapters 2 and 3.
- Kragh, Helge S., *Dirac: A Scientific Biography*. Cambridge University Press, Cambridge, UK, 1990, pp. 177–8.
- Susskind, Leonard, and Friedman, Art, *Quantum Mechanics: The Theoretical Minimum*, Penguin, London, 2015, Lecture 1, pp. 24–34.
- Von Neumann, John, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton, NJ, 1955, Chapters 1 and 2.

Von Neumann and the Problem of Quantum Measurement

The ‘Collapse of the Wavefunction’

Bohr’s debate with Heisenberg on the interpretation of the uncertainty principle (Chapter 7) prompted a period of deep introspection. Schrödinger had argued for a realistic interpretation of the wavefunction.* For him, the wavefunction was something physically meaningful and tangible; something that could be easily visualized. Heisenberg favoured a much more ‘anti-realist’ interpretation of quantum mechanics. This *doesn’t* mean that he rejected the idea of an objective external reality, or that invisible entities such as electrons really exist. It means that he wasn’t prepared to take the way that quantum mechanics *represents* this reality too literally. In particular he rejected any notion of the visualizability of some kind of underlying wave nature of matter, preferring to focus instead only on what can be empirically *observed*, such as the lines in an atomic spectrum, and the inherent discontinuity and uncertainty that such measurements implied. Bohr hovered between these extremes, perceiving the validity of both descriptions yet puzzled by the fact that he could find no words of his own.

His solution was the *principle of complementarity*. Whatever the true nature of the electron, the behaviour it exhibits is conditioned by the kinds of experiments we choose to perform. These, by definition, are experiments requiring apparatus of classical dimensions, resulting in effects sufficiently substantial to be observed and recorded in the laboratory, perhaps in the form of tracks in a cloud chamber, or the series of spots on an exposed photographic plate which build to an interference pattern, as we saw in Fig. 4.3.

So, a certain kind of experiment will yield effects that we interpret, using the language of classical physics, as electron diffraction and interference. We conclude that in this experiment the electron is a wave. Another kind of experiment will yield effects which we interpret in terms of the trajectories and collisions of localized electrons. We conclude that in this experiment the electron is a particle. These experiments are mutually exclusive.

* I’ll continue to use the term ‘wavefunction’ here, but will switch back to the formalism based on state vectors when we start to consider some more mathematics.

The Copenhagen Interpretation

Although Bohr was infamously obscure in many of his writings on the subject, and he was much less staunchly empiricist than Heisenberg, on balance I believe it is fair to conclude that Bohr adopted a generally anti-realist interpretation of the wavefunction. Although it's a bit of a stretch to provide only one Bohr quote in support of this conclusion (especially as this is not even a direct quote), I've nevertheless always found this rather telling. He is quoted as saying:¹

There is no quantum world. There is only an abstract quantum physical description. It is wrong to think that the task of physics is to find out how nature is. Physics concerns what we can say about nature.

This kind of anti-realism is nevertheless quite subtle. What Bohr is actually saying is that we're fundamentally limited by the classical nature of our apparatus and our measurements, and the language of classical waves and particles we use to describe what we see. It's therefore pretty pointless to speculate about the reality or otherwise of elements of the 'abstract quantum physical description', including the wavefunction, as we have absolutely no way of discovering anything about them.²

Heisenberg was initially resistant to Bohr's notion of complementarity, as it gave equal validity to the wave description associated with Heisenberg's rival, Schrödinger. As their debate became more bitter and personalized, Wolfgang Pauli was called to Bohr's institute in Copenhagen in early June 1927 to act as mediator. With Pauli's help, Bohr and Heisenberg forged an uneasy consensus.

The resulting interpretation of quantum mechanics is based at root on Bohr's anti-realist notion of wave–particle complementarity, and on the acceptance that the theory is complete—there are no missing ingredients. Physical observables are determined as the expectation values of their corresponding Hermitian operators, and non-commutation of these operators gives rise to the uncertainty principle. The theory deals only in quantum probabilities, according to the Born rule. And no, it's no coincidence that these foundations correspond closely with the axioms of the quantum formalism described in Chapter 10.

What was remarkable was the zeal with which the disciples of this new quantum orthodoxy embraced and preached the new gospel. Heisenberg spoke and wrote of the 'Kopenhagener Geist der Quantentheorie'; the 'Copenhagen spirit' of quantum theory.³ It became known as the *Copenhagen interpretation*.

Einstein didn't like it at all.

The Bohr–Einstein Debate

The stage was set for a great debate about the quantum representation of reality, and how this is meant to be interpreted. This commenced at the fifth Solvay congress in

Brussels in 1927, the first time that the protagonists were gathered together in one place to reflect on these new developments. Born and Heisenberg delivered a joint lecture, declaring that quantum mechanics is a complete theory, ‘whose fundamental physical and mathematical assumptions are no longer susceptible of any modification’.⁴ Schrödinger then delivered a lecture on wave mechanics. And, following an interruption to allow participants the opportunity to attend a competing conference that had been organized in Paris, Bohr presented a lecture on complementarity.

Then Einstein stood to raise an objection. It will be easier for us to understand Einstein’s argument in the context of two-slit interference involving single electrons (cf. Fig. 4.3). Before any kind of measurement, we suppose the electron wavefunction reflects the expected interference pattern—an alternating sequence of bright and dark fringes distributed across the screen. According to the Born rule, the probability of finding the electron is determined by the square of the amplitude of the wavefunction, so in principle it will be found in any location where this is non-zero—Fig. 11.1(a).

After measurement, we learn that the electron is ‘here’, in a single location (such as point A), forming a single bright spot on the screen—Fig. 11.1(b). However, Einstein now pointed out, we also learn *simultaneously* that the electron is definitely *not* ‘there’, where ‘there’ can be any location on the screen where we might have expected to find it, such as point B.

Einstein argued that this ‘assumes an entirely peculiar mechanism of action at a distance, which prevents the wave continually distributed in space from producing an action in two places on the screen’.⁵ This would later become widely known as ‘spooky action at a distance’. The particle, which according to the wavefunction is somehow distributed over a large region of space, becomes localized instantaneously, the act of measurement appearing to change the physical state of the system far from the point

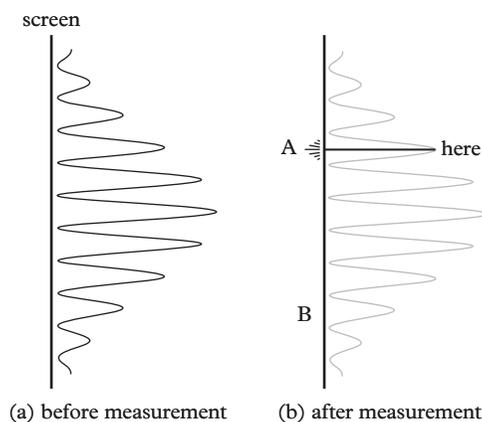


Figure 11.1 (a) Before measurement, the modulus-square of the electron wavefunction predicts a distribution of probabilities for where the electron might be found, spread across the screen. (b) After measurement, the electron is recorded to be found in one, and only one, location on the screen.

where the measurement is actually recorded. Einstein felt that this kind of action at a distance violates one of the key postulates of his special theory of relativity: no physical action, or information resulting in physical action, can be communicated at a speed faster than light. Any physical process that happens instantaneously over substantial distances violates this postulate.

We should note right away that all this talk about physical action betrays the fact that these concerns are based on a *realistic* interpretation of the wavefunction. This is not to say that Einstein wanted to ascribe reality to the wavefunction in the same way that Schrödinger did (we will see shortly that their views were quite different). But it's important to realize that, from the outset, the Bohr–Einstein debate involved a clash of *philosophical positions*. At great risk of oversimplifying, it was a confrontation between realism and anti-realism, at the level of the quantum theoretical representation.

As far as Bohr was concerned, the Copenhagen interpretation obliges us to resist the temptation to ask: *But how does nature actually do that?* Like emergency services personnel at the scene of a tragic accident, Bohr advises us to move along, as there's nothing to see here. And there lies the rub: for what is the purpose of a scientific theory if not to aid our *understanding* of the physical world? We *want* to rubberneck at reality. Arguably the only way to do this in quantum mechanics is to take the wavefunction more literally and realistically.

The discussion continued in the dining room of the Hôtel Britannique, where the conference participants were staying. Otto Stern described what happened:⁶

Einstein came down to breakfast and expressed his misgivings about the new quantum theory, every time [he] had invented some beautiful experiment from which one saw that [the theory] did not work. . . . Pauli and Heisenberg, who were there, did not pay much attention, 'ach was, das stimmt schon, das stimmt schon' [ah well, it will be all right, it will be all right]. Bohr, on the other hand, reflected on it with care and in the evening, at dinner, we were all together and he cleared up the matter in detail.

Einstein now sought to highlight what he perceived to be logical fallacies in the Copenhagen interpretation by developing a series of hypothetical tests, or *gedankenexperiments* (thought experiments). These were about matters of principle; they were not meant to be taken too literally as practical experiments that could be carried out in the laboratory.

He began by attempting to show up inconsistencies in the interpretation of the uncertainty principle, but each challenge was deftly rebutted by Bohr. However, under pressure from Einstein's insistent probing, the basis of Bohr's counter-arguments underwent a subtle shift. Bohr was obliged to fall back on the notion that measurements using classical apparatus are just too 'clumsy', implying limits on what can be *measured*, rather than limits on what we can *know*. This was precisely the interpretation which had prompted his bitter debate with Heisenberg.

In the eyes of the majority of physicists gathered in Brussels, Bohr won the day. But Einstein remained stubbornly unconvinced, and the seeds of a much more substantial challenge were sown.

The debate continued at the next Solvay Conference, which was held in 1930. Einstein's introduced a wickedly ingenious thought experiment designed to highlight

inconsistencies in the application of the energy–time uncertainty relation (this is Einstein’s ‘photon box’ experiment).⁷ Bohr was quite shocked, and he didn’t see the solution right away. Bohr’s eventual response used Einstein’s own general theory of relativity against him. Einstein conceded that Bohr’s response appeared to be ‘free of contradictions’, but in his view it still contained ‘a certain unreasonableness’.⁸ At the time this was hailed as a triumph for Bohr and for the Copenhagen interpretation.

But, once again, Bohr had been obliged to defend the integrity of the uncertainty principle using arguments based on an inevitable and sizeable *disturbance* of the observed system. At first sight, there seems no way around this. Surely, measurement of any kind will always involve interactions that are at least as large as the quantum system being measured. How can such a clumsy disturbance possibly be avoided?

Einstein chose to shift the focus of his challenge. Instead of arguing that quantum mechanics—and particularly the uncertainty principle—is *inconsistent*, he now sought to derive a logical paradox arising from what he saw to be the theory’s *incompleteness*.

Although another five years would elapse, Bohr was quite unprepared for Einstein’s next move.

Photon Spin and Circular Polarization

To prepare the mathematical ground for these closing chapters, I propose to interrupt the historical narrative here for a short discussion of the spin properties of photons and their relation with the phenomenon of light polarization. Actually, I have a further motive. I want you to watch very closely as we go through the logic of translating our *experience* of light polarization into a set of mathematical relationships constructed using the quantum formalism.

In our discussion of the Pauli principle in Chapter 8, I mentioned that photons are bosons, with a spin quantum number $s = 1$. We might be tempted to suppose that photons must therefore possess $(2s + 1) = 3$ different spin ‘orientations’, corresponding to $m_s = -1, 0$, and $+1$.

But we’d be wrong. Relativistic quantum mechanics forbids the orientation corresponding to $m_s = 0$ for particles moving at the speed of light. There are two ways you can try to reconcile this, both fairly inadequate. The $m_s = 0$ orientation corresponds to a component of the spin angular momentum *in the direction of motion*, and if we imagine the photon as a spherical particle of some kind (which it clearly is not) this implies that the photon spins around its axis with a velocity *faster* than the speed of light, which is impossible.

Another (and in my view slightly better) way of thinking about this is that according to special relativity, the length of any object travelling at the speed of light contracts to 0 according to $l = l_0/\gamma$, where l_0 is the ‘proper’ length and γ is the Lorentz factor (see Chapter 2). When $v \rightarrow c$, $\gamma \rightarrow \infty$ and $l \rightarrow 0$. A photon moving at the speed of light is therefore ‘flat’: *it has no dimension* in the direction in which it is travelling. And there can be no component of spin angular momentum in a dimension that doesn’t exist.

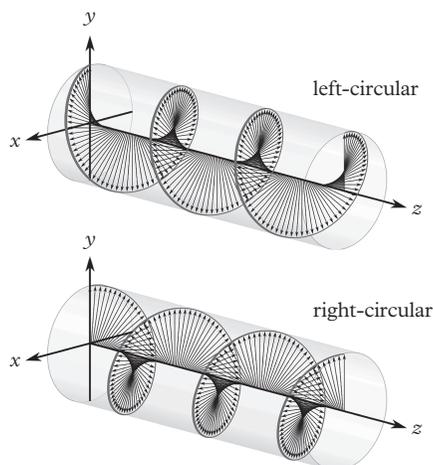


Figure 11.2 *Convention for left- and right-circular polarization, as judged from the perspective of the source of the radiation. Only the electric vector is shown.*

Either way, we must acknowledge that there are just two spin orientations for the photon, corresponding to $m_s = \pm 1$. By convention, we associate the $m_s = +1$ orientation with left-circularly polarized light and the $m_s = -1$ orientation with right-circularly polarized light, as judged from the perspective of the *source* of the light. These different polarizations are most readily visualized in terms of the old classical picture in which the electric and magnetic vectors are imagined to rotate clockwise or anti-clockwise around the direction of travel (see Fig. 11.2). This assignment is based on a convention typically adopted in quantum physics (but take care—the opposite convention is used in physical optics).

Although the spin property of a quantum particle should never be interpreted as if the particle were literally spinning on its axis, it is nevertheless manifested as an intrinsic angular momentum. Thus, a beam containing a large number of circularly polarized photons (such as a laser beam) will impart a measurable torque to a target. This angular momentum is not only a collective phenomenon: in the absorption of an individual photon resulting in electron excitation in an atom or molecule, as we've seen the angular momentum intrinsic to the photon is transferred to the excited electron and total angular momentum is conserved. That transfer has important, measurable effects on the absorption spectrum.

Projection Probabilities for Photon Spin States

It's perhaps possible that your most familiar experience with polarized light outside of the laboratory comes from wearing Polaroid sunglasses, which reduce glare by filtering out horizontally polarized light. This is an example of a *linear* polarizing filter, in this case

orientated so as to pass vertically polarized light and block horizontally polarized light. But how does this vertical/horizontal polarization relate to left/right circular polarization?

We can get some sense for this relationship by passing a source of circularly polarized light through a *calcite crystal*. There are a number of ways of obtaining a source of left-circularly polarized light (for example). These vary from a standard (i.e. unpolarized) light source passed through an optical device known as a quarter-wave plate to an atomic source that relies on the quantum mechanics of photon emission. An example of the latter is a beam of atoms that are excited to some electronically excited state from which emission occurs. As we've already seen, if angular momentum is to be conserved in the process, the emitted photon must carry away any excess angular momentum lost by the excited electron as it returns to a more stable quantum state. An appropriate choice of states between which the transition occurs can give rise to the emission of photons exclusively with $m_s = +1$. We will meet this kind of source again in Chapter 12.

Calcite is a naturally *birefringent* form of calcium carbonate. It has a crystal structure which has different refractive indices along two distinct crystal planes. One offers an axis of maximum transmission for vertically polarized light and the other offers an axis of maximum transmission for horizontally polarized light. The vertical and horizontal components of left-circularly polarized light are therefore physically separated by passage through the crystal, and their intensities can be measured separately. With careful machining, a calcite crystal can transmit virtually all the light incident on it.

We discover that a beam of left-circularly polarized light entering a calcite crystal will split into two beams, one of vertical and one of horizontal polarization. If the initial beam intensity is I_L , we can use detectors placed in the paths of the emergent beams (such as photomultipliers or photodiodes, connected to amplifiers and recording equipment) to confirm that the intensity of vertically polarized light $I_v = \frac{1}{2}I_L$ and likewise $I_h = \frac{1}{2}I_L$.

So how do we translate this experimental observation into a quantum description? With sufficiently sensitive equipment we discover that passing left-circularly polarized photons through the crystal one at a time we observe photons emerging from *either* the vertical *or* horizontal channels. If we repeat these measurements, we discover that the *probability* for each photon to emerge from the vertical channel is equal to $\frac{1}{2}$. Similarly, the probability for each photon to emerge from the horizontal channel is equal to $\frac{1}{2}$. We have no way of knowing beforehand which photons will emerge from which channels.

As we're starting with left-circularly polarized light we translate these results into a set of *projection probabilities*,

$$|\langle v|L\rangle|^2 = |\langle h|L\rangle|^2 = \frac{1}{2}, \tag{11.1}$$

where $|L\rangle$ represents the quantum state corresponding to a left-circularly polarized photon (with $m_s = +1$), and $|v\rangle$ and $|h\rangle$ represent the quantum states for vertical and horizontal polarization, respectively.

If we repeat this experiment with right-circularly polarized light, we get precisely the same results,

$$|\langle v|R\rangle|^2 = |\langle h|R\rangle|^2 = \frac{1}{2} \tag{11.2}$$

where $|R\rangle$ represents the quantum state corresponding to a right-circularly polarized photon (with $m_s = -1$).

What happens if we now rotate the calcite crystal by an angle φ clockwise as measured from the vertical axis (see Fig. 11.3)? This serves to define a new axis, which we denote v'/h' , rotated relative to v/h . We find that we get exactly the same results, implying that

$$\begin{aligned} |\langle v'|L\rangle|^2 &= |\langle h'|L\rangle|^2 = \frac{1}{2} \\ |\langle v'|R\rangle|^2 &= |\langle h'|R\rangle|^2 = \frac{1}{2}. \end{aligned} \tag{11.3}$$

Now it starts to get interesting. We set aside the calcite crystal and pick up a piece of Polaroid film. This is likely to be much less efficient, perhaps transmitting as little as 70% of vertically polarized light even when the film is orientated vertically to allow maximum transmission. But this is okay so long as we deal only with the *ratios* of transmitted light intensities, and we assume that the transmission efficiency doesn't change as we rotate the film (a big assumption).

We start with vertically polarized light and measure the transmitted intensity, I_v . If we now rotate the film clockwise through an angle φ to the vertical, once again we define a new axis, which we denote v'/h' . We discover that the ratio of intensities $I_{v'}/I_v = \cos^2\varphi$. This is *Malus's law*, named for Étienne-Louis Malus. Similarly, we find that $I_{h'}/I_h = \cos^2\varphi$. A further series of manipulations allows us to determine that $I_{v'}/I_h = \sin^2\varphi$ and $I_{h'}/I_v = \sin^2\varphi$. From these results we can deduce that

$$\begin{aligned} |\langle v'|v\rangle|^2 &= |\langle h'|h\rangle|^2 = \cos^2\varphi \\ |\langle v'|h\rangle|^2 &= |\langle h'|v\rangle|^2 = \sin^2\varphi. \end{aligned} \tag{11.4}$$

Our final step is to acknowledge that the states $|L\rangle/|R\rangle$, $|v\rangle/|h\rangle$, and $|v'\rangle/|h'\rangle$ all constitute different orthonormal *basis sets* which can be used to describe the spin states of photons, just as we use $|\uparrow\rangle/|\downarrow\rangle$ to describe the spin states of electrons. Assuming these pairs of

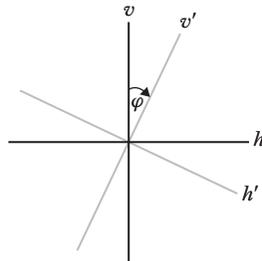


Figure 11.3 *Axis convention for vertical/horizontal polarization.*

Table 11.1 Projection Probabilities $|\langle f|i\rangle|^2$ for Photon Polarization States

Final state $\langle f $	Initial state $ i\rangle$		$ v'\rangle$	$ h'\rangle$	$ L\rangle$	$ R\rangle$
	$ v\rangle$	$ h\rangle$				
$\langle v $	1	0	$\cos^2\varphi$	$\sin^2\varphi$	$\frac{1}{2}$	$\frac{1}{2}$
$\langle h $	0	1	$\sin^2\varphi$	$\cos^2\varphi$	$\frac{1}{2}$	$\frac{1}{2}$
$\langle v' $	$\cos^2\varphi$	$\sin^2\varphi$	1	0	$\frac{1}{2}$	$\frac{1}{2}$
$\langle h' $	$\sin^2\varphi$	$\cos^2\varphi$	0	1	$\frac{1}{2}$	$\frac{1}{2}$
$\langle L $	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	0
$\langle R $	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	1

photon states are orthonormal allows us to summarize the projection probabilities in Table 11.1. However, while that is all very fine, what we really need are the projection *amplitudes* that correspond to the probabilities in Table 11.1.

Projection Amplitudes for Photon Spin States

Look back at Fig. 11.3. If we assume that the state vectors $|v\rangle/|h\rangle$, and $|v'\rangle/|h'\rangle$ behave as unit vectors, then we can quickly deduce that

$$|v'\rangle = |v\rangle\cos\varphi + |h\rangle\cos(90^\circ - \varphi) = |v\rangle\cos\varphi + |h\rangle\sin\varphi. \tag{11.5}$$

Multiplying from the left by the bra vector $\langle v|$ then gives

$$\langle v|v'\rangle = \langle v|v\rangle\cos\varphi + \langle v|h\rangle\sin\varphi = \cos\varphi. \tag{11.6}$$

This follows because $\langle v|v\rangle = 1$ and $\langle v|h\rangle = 0$ (the basis states are orthonormal). Similarly,

$$\langle h|v'\rangle = \langle h|v\rangle\cos\varphi + \langle h|h\rangle\sin\varphi = \sin\varphi. \tag{11.7}$$

It also follows that

$$|h'\rangle = |v\rangle\cos(\varphi + 90^\circ) + |h\rangle\cos\varphi = -|v\rangle\sin\varphi + |h\rangle\cos\varphi, \tag{11.8}$$

and so

$$\begin{aligned}\langle v|h' \rangle &= -\langle v|v \rangle \sin \varphi + \langle v|h \rangle \cos \varphi = -\sin \varphi \\ \langle h|h' \rangle &= -\langle h|v \rangle \sin \varphi + \langle h|h \rangle \cos \varphi = \cos \varphi.\end{aligned}\quad (11.9)$$

We can obviously use the same logic to show that $\langle v'|v \rangle = \langle v|v' \rangle = \cos \varphi$, and so on for all the other projection amplitudes involving $|v\rangle/|h\rangle$, and $|v'\rangle/|h'\rangle$. This is perfectly logical, as the choice of ‘vertical axis’ is obviously quite arbitrary. Note that these projection amplitudes are also entirely consistent with the projection probabilities listed in Table 11.1.

Projection amplitudes for states involving $|L\rangle/|R\rangle$ are a little trickier to work out. If the states $|L\rangle/|R\rangle$ provide a suitable basis for all the other spin states, then we can use the expansion theorem to express $|v'\rangle$ as a superposition (cf. Eq. (10.20)):

$$|v'\rangle = |L\rangle\langle L|v'\rangle + |R\rangle\langle R|v'\rangle. \quad (11.10)$$

If we now multiply from the left by the bra vector $\langle v|$ we get (from (11.6))

$$\langle v|v'\rangle = \langle v|L\rangle\langle L|v'\rangle + \langle v|R\rangle\langle R|v'\rangle = \cos \varphi. \quad (11.11)$$

Or, alternatively,

$$\langle v|L\rangle\langle L|v'\rangle + \langle v|R\rangle\langle R|v'\rangle = \frac{1}{2}(e^{i\varphi} + e^{-i\varphi}), \quad (11.12)$$

where we have replaced $\cos \varphi$ with its complex exponential form.

From Table 11.1, we know that $|\langle v|L\rangle|^2 = |\langle v|R\rangle|^2 = \frac{1}{2}$, and it obviously makes sense to assign the phase factors $e^{i\varphi}$ and $e^{-i\varphi}$ to the amplitudes $\langle L|v'\rangle$ and $\langle R|v'\rangle$. By convention, we assign the phase factor $e^{i\varphi}$ with projection amplitudes involving $|L\rangle$, consistent with the quantum number $m_s = +1$. So, from (11.12), we can deduce that

$$\langle v|L\rangle = \langle v|R\rangle = \frac{1}{\sqrt{2}} \quad (11.13)$$

$$\begin{aligned}\langle L|v'\rangle &= \frac{1}{\sqrt{2}}e^{i\varphi}, \text{ and} \\ \langle R|v'\rangle &= \frac{1}{\sqrt{2}}e^{-i\varphi}.\end{aligned}\quad (11.14)$$

This assignment might seem rather arbitrary, but we must remember that we have no way of knowing the ‘actual’ signs of the complex phase factors because this information is not revealed in experiments—we can only discover the projection probabilities which are the modulus-squares of the projection amplitudes. However, we can adopt a *phase*

convention which, if we stick to it rigorously, will always give results that are both internally consistent and consistent with experiment.⁹

We can follow the same logic for $|h'\rangle$,

$$|h'\rangle = |L\rangle\langle L|h'\rangle + |R\rangle\langle R|h'\rangle, \quad (11.15)$$

from which we get, using Eq. (11.9),

$$\langle v|h'\rangle = \langle v|L\rangle\langle L|h'\rangle + \langle v|R\rangle\langle R|h'\rangle = -\sin\varphi = -\frac{1}{2i}(e^{i\varphi} - e^{-i\varphi}), \quad (11.16)$$

and from which it follows that

$$\begin{aligned} \langle L|h'\rangle &= -\frac{1}{\sqrt{2}i}e^{i\varphi} = \frac{i}{\sqrt{2}}e^{i\varphi} \\ \langle R|h'\rangle &= \frac{1}{\sqrt{2}i}e^{-i\varphi} = -\frac{i}{\sqrt{2}}e^{-i\varphi}. \end{aligned} \quad (11.17)$$

From Eqs. (11.7) and (11.10) we have

$$\langle h|v'\rangle = \langle h|L\rangle\langle L|v'\rangle + \langle h|R\rangle\langle R|v'\rangle = \sin\varphi = \frac{1}{2i}(e^{i\varphi} - e^{-i\varphi}), \quad (11.18)$$

from which we can deduce that

$$\begin{aligned} \langle h|L\rangle\langle L|v'\rangle &= \frac{1}{2i}e^{i\varphi} = -\frac{i}{2}e^{i\varphi} \quad \text{hence } \langle h|L\rangle = -\frac{i}{\sqrt{2}}; \\ \langle h|R\rangle\langle R|v'\rangle &= -\frac{1}{2i}e^{-i\varphi} = \frac{i}{2}e^{-i\varphi} \quad \text{hence } \langle h|R\rangle = \frac{i}{\sqrt{2}}, \end{aligned} \quad (11.19)$$

where we have made use of the fact that $1/i = -i$.

To deduce the projection amplitudes $\langle v'|L\rangle$ and $\langle h'|L\rangle$, consider the superposition

$$|L\rangle = |v'\rangle\langle v'|L\rangle + |h'\rangle\langle h'|L\rangle \quad (11.20)$$

and multiply from the left by the bra vector $\langle L|$

$$\begin{aligned} \langle L|L\rangle &= \langle L|v'\rangle\langle v'|L\rangle + \langle L|h'\rangle\langle h'|L\rangle \\ &= \frac{1}{\sqrt{2}}e^{i\varphi}\langle v'|L\rangle + \frac{i}{\sqrt{2}}e^{i\varphi}\langle h'|L\rangle = 1. \end{aligned} \quad (11.21)$$

Clearly, we can only reconcile Eq. (11.21) if the amplitudes $\langle v'|L\rangle$ and $\langle h'|L\rangle$ are the *complex conjugates* of $\langle L|v'\rangle$ and $\langle L|h'\rangle$, respectively, which is precisely what we would expect.

Table 11.2 Projection Amplitudes $\langle f|i \rangle$ for Photon Polarization States

Final state $\langle f $	Initial state $ i\rangle$		$ v'\rangle$	$ h'\rangle$	$ L\rangle$	$ R\rangle$
	$ v\rangle$	$ h\rangle$				
$\langle v $	1	0	$\cos \varphi$	$-\sin \varphi$	$1/\sqrt{2}$	$1/\sqrt{2}$
$\langle h $	0	1	$\sin \varphi$	$\cos \varphi$	$-i/\sqrt{2}$	$i/\sqrt{2}$
$\langle v' $	$\cos \varphi$	$\sin \varphi$	1	0	$e^{-i\varphi}/\sqrt{2}$	$e^{i\varphi}/\sqrt{2}$
$\langle h' $	$-\sin \varphi$	$\cos \varphi$	0	1	$-ie^{-i\varphi}/\sqrt{2}$	$ie^{i\varphi}/\sqrt{2}$
$\langle L $	$1/\sqrt{2}$	$i/\sqrt{2}$	$e^{i\varphi}/\sqrt{2}$	$ie^{i\varphi}/\sqrt{2}$	1	0
$\langle R $	$1/\sqrt{2}$	$-i/\sqrt{2}$	$e^{-i\varphi}/\sqrt{2}$	$-ie^{-i\varphi}/\sqrt{2}$	0	1

We can now summarize the projection amplitudes in Table 11.2.

We now need to ask: *what just happened here?* There are at least two ways of interpreting what we’ve just done. We could say that we have used the quantum formalism simply as a way to *encode* our empirical observations of the behaviour and properties of circularly and linearly polarized light. The formalism provides a powerful tool for making predictions for any set of experimental manipulations we might make in the future. This is a firmly *anti-realist* interpretation: the state vectors used in the formalism have no real significance beyond their ability to carry encoded information about the physics.

For example, we can write the state vector $|L\rangle$ as a superposition involving $|v\rangle/|h\rangle$, or $|v'\rangle/|h'\rangle$, as follows:

$$|L\rangle = |v\rangle\langle v|L\rangle + |h\rangle\langle h|L\rangle = \frac{1}{\sqrt{2}}|v\rangle - \frac{i}{\sqrt{2}}|h\rangle = \frac{1}{\sqrt{2}}(|v\rangle - i|h\rangle), \tag{11.22}$$

or

$$\begin{aligned} |L\rangle &= |v'\rangle\langle v'|L\rangle + |h'\rangle\langle h'|L\rangle = \frac{1}{\sqrt{2}}e^{-i\varphi}|v'\rangle - \frac{i}{\sqrt{2}}e^{-i\varphi}|h'\rangle \\ &= \frac{e^{-i\varphi}}{\sqrt{2}}(|v'\rangle - i|h'\rangle). \end{aligned} \tag{11.23}$$

As we did in Chapter 7, we might ask once again: Which is the ‘correct’ form for the state vector $|L\rangle$? And, once again, the simple answer is that there really isn’t one. So long as we make use of the basis states and observe the rules of the quantum formalism, then we’re free to choose whatever basis is most appropriate to the problem we’re addressing or the measurement we’re about to perform.*

Alternatively, we could say that the formalism itself is founded on the physics of many different kinds of quantum systems, and all we’ve done here is use our observations of

* We should note in passing that if we define $|L'\rangle$ as the left-circularly polarized state in the v'/h' frame, then from Eq. (11.22) we have $|L'\rangle = (|v'\rangle - i|h'\rangle)/\sqrt{2}$. In other words, $|L'\rangle = e^{-i\varphi}|L\rangle$, or $|L\rangle = e^{i\varphi}|L'\rangle$. The two left-circular polarization states are not quite the same—one is *phase-shifted* compared with the other.

circularly and linearly polarized light to deduce the real properties of the quantum states of the photons that underlie this empirical behaviour. This is a *realist* interpretation: the state vectors represent the real quantum spin states of photons which, unlike the physical states of classical objects, can form superpositions.

Hold these thoughts, because our understanding of quantum mechanics is about to take a spectacularly bizarre turn.

Von Neumann's Theory of Quantum Measurement

We're now ready to contemplate this chapter's derivation. In fact, this is not so much a derivation based on mathematics as a derivation based on logical arguments. Einstein's *gedankenexperiments* were focused on the process of measurement in quantum mechanics and the relationships between quantum systems and the classical measuring devices (such as calcite crystals, Polaroid film, photomultipliers or photodiodes, amplifiers, recording equipment, and—let's not forget—human experimenters) that we use to observe and record their behaviour.

Bohr did not believe that a specific theory of quantum measurement was required—quantum mechanics was already complete without it. It was enough for Bohr to distinguish between the 'quantum world' and the 'classical world' and acknowledge that we are prevented from acquiring knowledge of the former through our experiences in the latter. But this seemingly arbitrary split (where does the 'quantum world' end and the 'classical world' begin?) left some nagging doubts. Many years later the Irish physicist John Bell would call it the 'shifty split'.¹⁰

Although von Neumann's axiomatic approach to the quantum formalism conformed quite closely to the building blocks of the Copenhagen interpretation, unlike both Bohr and Heisenberg he felt it was necessary in *Mathematical Foundations* to address the problem of quantum measurement directly. Significantly, he saw no need to introduce an arbitrary split between worlds. As far as he was concerned, the new quantum mechanics is the principal theory of matter and radiation and applies not just to quantum objects such as electrons and photons, but to classical measuring devices, too.*

Although what follows is a much-simplified account, it more or less summarizes what von Neumann proposed.

The Ingredients

1. A definition of the measurement 'process' in terms of three distinct components, I, II, and III.
2. The expansion theorem applied to the quantum state for left-circularly polarized photons, Eq. (11.21).

* Some commentators have incorporated von Neumann's approach to the measurement problem into their understanding of the 'Copenhagen interpretation'. Strictly speaking, this is incorrect. The Copenhagen interpretation is based on an arbitrary split between quantum and classical worlds, so 'measurement' is not actually a problem.

3. The notion of a ‘measurement operator’, \hat{M} , whose eigenstates are the states detected and whose eigenvalues are the measurement outcomes.
4. Von Neumann’s *projection postulate*, which is more popularly known as the ‘collapse of the wavefunction’.
5. A speculative proposal for component III.

The Recipe

We begin in Step (1) by defining what we mean by ‘measurement’. As usual, it is much more straightforward to see how the measurement process is supposed to work when we apply it to a real system, and in Step (2) we use the expansion theorem to express the quantum state of left-circularly polarized photons in terms of a basis of vertical and horizontal polarization states. This defines the quantum system in component I of von Neumann’s measurement process. If this is a process through which a quantum system produces specific measurement outcomes, then in principle we can define a ‘measurement operator’, \hat{M} , which holds all the necessary physical manipulations, with eigenstates which are composites of the quantum states of component I *and* the states of the classical measuring instrument, and with eigenvalues that are simply the outcomes we observe. We do this in Step (3).

We will discover that this doesn’t give us quite what we’re looking for, and von Neumann had no alternative but to propose a further postulate in component III to explain what we see. He also had some suggestions for the mechanism that might be responsible. We will consider both the postulate and the mechanism in Step (4).

Step (1): Von Neumann’s Measurement Process

In *Mechanical Foundations*, von Neumann defined the components of the measurement process as follows.

- I is the quantum system under investigation;
- II is the physical measuring instrument; and
- III is the ‘observer’.

Step (2): Define the Quantum System under Investigation: State Preparation

It will help in what follows to define the quantum system on which we’re going to make our measurements—component I. As we’ve now done all the ground work, I propose to select a system in which we start with a collection—or an *ensemble*—of photons all prepared in a quantum state of left-circular polarization, $|L\rangle$.

State preparation must be distinguished from measurement itself. Measurement is defined as an operation performed on a quantum system which probes the state immediately *before* the measurement and yields an outcome which can be interpreted from some physically registered event or series of events (such as the click of a Geiger counter, firing of a photomultiplier tube, or, most simply, the precipitation of silver atoms in a photographic emulsion). The measurement process changes the state of the system (and often destroys the system altogether). State preparation, on the other hand, is an operation performed on a system consisting of many particles which forces it into an ensemble of identical states or collections of states.

The states referred to are those of the system *after* the operation has taken place. In doing this, much the same processes take place, and not all particles in the original system necessarily survive. For example, we can prepare an ensemble of left-circularly polarized photons by creating specific excited states of atoms (for example, in an ‘atomic beam’), and gathering the light emitted as these atoms return to the lowest-energy ground state. I’ll provide a few more details about this kind of source in Chapter 12.

But how do we know that such photons are all prepared in the ‘same’ quantum state? In general, we know that systems exposed to certain preparation operations will behave in predictable (but probabilistic) ways when subjected to certain measurement operations. The formalism requires an initial state vector of the prepared system and information regarding the possible measurement outcomes. Provided we stick to the rules of the quantum formalism, we can expect to derive accurate predictions.

For the purposes of this derivation, we suppose that our physical measurement (component II) involves the use of a calcite crystal, sensitive photomultipliers to detect photons emerging from the vertical and horizontal channels, amplifiers, and a recording device which will *count* the numbers of photons detected from each channel. In these circumstances, we connect the quantum state of the *input* prepared system, $|L\rangle$, with the *output* quantum states of the calcite crystal, $|v\rangle/|h\rangle$, using the superposition given in Eq. (11.21). We then detect, amplify, and count the photons emerging separately from both channels.

Step (3): The Measurement Operator, Eigenstates, and Eigenvalues

There is clearly more to this measurement than the calcite crystal itself. If—as von Neumann proposed—there is to be no artificial split between the quantum system and our classical measuring apparatus, then we are perfectly at liberty to apply the logic of the quantum formalism also to component II.

The measurement process goes something like this. A left-circularly polarized photon enters the calcite crystal and emerges from either the vertical or horizontal channel. Suppose it emerges from the vertical channel. The photon is detected, and the electrical signal from the detector is amplified. The amplified signal passes to an electronic device which cumulatively counts the number of times this sequence occurs. The count for the vertical channel increases by 1. We apply the same logic for photons emerging from the

horizontal channel. Following von Neumann, we apply quantum mechanics throughout. For the sake of simplicity, we will assume that both sets of apparatus have identical properties and efficiencies (although, in practice, we can usually *calibrate* these and take any differences into account). This means we can consider them as a *single* instrument.

We know that the initial quantum state of the system under investigation is $|L\rangle$. But what are the quantum states of the measuring instrument? We define a basis consisting of just three states:

- An initial neutral or ‘standby’ state corresponding to the instrument prior to the detection of a photon from either channel of the calcite crystal: let’s call this $|I_0\rangle$;
- A state corresponding to the detection, amplification, and counting of a photon emerging from the vertical channel, $|I_v\rangle$; and
- A state corresponding to the detection, amplification, and counting of a photon emerging from the horizontal channel, $|I_h\rangle$.

It’s worth noting that if we assume 100% efficiency—every photon entering the instrument results in detection—then the states $|I_0\rangle$ and $|I_v\rangle/|I_h\rangle$ represent the instrument *at different times* in the measurement process.

Before any measurement takes place, the total quantum system and measuring instrument can be considered to be in a *composite state*, formed from $|L\rangle$ and $|I_0\rangle$, which we will write as $|LI_0\rangle$. Strictly speaking, this is the *tensor product* of $|L\rangle$ and $|I_0\rangle$, $|LI_0\rangle = |L\rangle \otimes |I_0\rangle$, but for our purposes here we can safely assume $|LI_0\rangle = |L\rangle|I_0\rangle$.

Given what we know the calcite crystal is set up to do, it makes sense to expand $|L\rangle$ in a basis of $|v\rangle/|h\rangle$ states,

$$\begin{aligned} |LI_0\rangle &= |L\rangle|I_0\rangle = (|v\rangle\langle v|L\rangle + |h\rangle\langle h|L\rangle)|I_0\rangle \\ &= |v\rangle|I_0\rangle\langle v|L\rangle + |h\rangle|I_0\rangle\langle h|L\rangle \\ &= |vI_0\rangle\langle v|L\rangle + |hI_0\rangle\langle h|L\rangle, \end{aligned} \tag{11.24}$$

where $|vI_0\rangle$ and $|hI_0\rangle$ are composite states corresponding to photons emerging from the vertical and horizontal channels, respectively, but which have yet to be detected.

Clearly, the process of detection, amplification, and counting converts the states $|vI_0\rangle$ and $|hI_0\rangle$ into *measurement eigenstates*:

$$\begin{aligned} |vI_0\rangle &\rightarrow |vI_v\rangle \quad \text{and} \\ |hI_0\rangle &\rightarrow |hI_h\rangle. \end{aligned} \tag{11.25}$$

We anticipate that this process will take a finite time, t . If we now define a measurement operator \hat{M} to encompass every step in the measurement chain from detection to counting, then (from Chapter 10) the corresponding time evolution operator is given by $\hat{U} = e^{-i\hat{M}t/\hbar}$ and we can write (11.25) a little more explicitly as

$$\begin{aligned} |vI_v\rangle &= \hat{U}|vI_0\rangle \quad \text{and} \\ |hI_h\rangle &= \hat{U}|hI_0\rangle. \end{aligned} \tag{11.26}$$

This allows us to deduce how the initial composite state $|LI_0\rangle$ in Eq. (11.24) will evolve in time:

$$\begin{aligned} \hat{U}|LI_0\rangle &= \hat{U}|vI_0\rangle\langle v|L\rangle + \hat{U}|hI_0\rangle\langle h|L\rangle \\ &= |vI_v\rangle\langle v|L\rangle + |hI_h\rangle\langle h|L\rangle. \end{aligned} \tag{11.27}$$

What we observe with each measurement event is that either the vertical or horizontal counter increases by 1. We define an increase in the vertical count as 1_v , and an increase in the horizontal count as 1_h . These are the ‘observables’, so we’re free to define these as the *measurement eigenvalues*, such that

$$\begin{aligned} \hat{M}|vI_v\rangle &= 1_v|vI_v\rangle \quad \text{and} \\ \hat{M}|hI_h\rangle &= 1_h|hI_h\rangle. \end{aligned} \tag{11.28}$$

This is the approach sanctioned by Dirac, who in *The Principles* wrote:¹¹

if the system is in a state such that a measurement of a real dynamical variable ... is certain to give one particular result ... then the state is an eigenstate ... and the result of the measurement is the eigenvalue ... to which this eigenstate refers.

If we now apply the measurement operator to the evolved state given by $\hat{U}|LI_0\rangle$, we get, from (11.27),

$$\begin{aligned} \hat{M}\hat{U}|LI_0\rangle &= \hat{M}|vI_v\rangle\langle v|L\rangle + \hat{M}|hI_h\rangle\langle h|L\rangle \\ &= 1_v|vI_v\rangle\langle v|L\rangle + 1_h|hI_h\rangle\langle h|L\rangle \end{aligned} \tag{11.29}$$

and we’re now in a position to deduce the expectation value for the measurement,

$$\begin{aligned} \langle \hat{M} \rangle &= \langle I_0L|\hat{U}^* \hat{M}\hat{U}|LI_0\rangle \\ &= (\langle L|v\rangle\langle I_vv| + \langle L|h\rangle\langle I_hh|) (1_v|vI_v\rangle\langle v|L\rangle + 1_h|hI_h\rangle\langle h|L\rangle) \\ &= 1_v|\langle v|L\rangle|^2 \langle I_vv|vI_v\rangle + 1_h\langle L|v\rangle\langle h|L\rangle\langle I_vv|hI_h\rangle \\ &\quad + 1_v\langle L|h\rangle\langle v|L\rangle\langle I_hh|vI_v\rangle + 1_h|\langle h|L\rangle|^2 \langle I_hh|hI_h\rangle \\ &= 1_v|\langle v|L\rangle|^2 + 1_h|\langle h|L\rangle|^2, \end{aligned} \tag{11.30}$$

where we have assumed that the composite states $|vI_v\rangle/|hI_h\rangle$ are orthonormal.

We now define the probability of detecting a photon from the vertical channel $P_v = |\langle v|L\rangle|^2$ and the probability of detecting a photon from the horizontal channel as $P_h = |\langle h|L\rangle|^2$, giving

$$\langle \hat{M} \rangle = P_v 1_v + P_h 1_h = \frac{1}{2} (1_v + 1_h), \quad (11.31)$$

where we have applied the projection probabilities given in Table 11.1.

Obviously, the individual channel counts can't be half-integral. We interpret Eq. (11.31) to mean that each photon has a probability of $\frac{1}{2}$ of being found emerging from the vertical channel (adding 1 to the vertical count) and a probability of $\frac{1}{2}$ of being found emerging from the horizontal channel (adding 1 to the horizontal count).

This is all very logical. But we know that when we perform each measurement, *we get only one outcome*. Equation (11.31) implies an expectation value for each photon detected that is 50% vertical *and* 50% horizontal. But what we actually see for each photon is 100% vertical *or* 100% horizontal.

How does *and* become *or*?

Step (4): Apply von Neumann's Projection Postulate

Let's recap. Setting aside for a moment all the business about forming composite states of the quantum system and the measuring instrument, we can see from Eq. (11.31) that the factors critically determining the measurement outcomes are the projection probabilities, and hence the projection amplitudes $\langle v|L \rangle$ and $\langle h|L \rangle$.^{*} These carry through the entire measurement chain.[†]

This happens because, once we have set up the superposition given by (11.24), there is nothing in the logic of the formalism that 'breaks' this up, allowing either $|vI_0 \rangle$ and $|hI_0 \rangle$ or $|vI_v \rangle$ and $|hI_v \rangle$ to 'separate'. I've already said in Chapter 10 that the time-evolution operator is unitary and, as such, it describes a smooth and continuous transformation from $|vI_0 \rangle$ to $|vI_v \rangle$ (and $|hI_0 \rangle$ to $|hI_v \rangle$) according to the time-dependent Schrödinger equation. What it *can't* do is transform $|vI_0 \rangle$ to $|vI_v \rangle$ (or $P_v \rightarrow 1$) whilst simultaneously forcing $|hI_0 \rangle$ to 0 (or $P_h \rightarrow 0$), and vice versa. In fact, there is *nothing* in the quantum formalism that can account for this.

Von Neumann understood that this is a problem, and felt he therefore had no alternative but to introduce a further postulate. He distinguished between two fundamentally different types of quantum process. The second, which he called process 2, is the continuous, deterministic, and completely reversible evolution of the composite quantum and instrument states, governed by the time-dependent Schrödinger equation. The first, which he referred to as process 1, is the discontinuous, irreversible transformation of a mixture of states (say $|vI_v \rangle$ and $|hI_v \rangle$), into just one measurement outcome.

In other words, von Neumann postulated the projection of *and* into *or*. This is his 'projection postulate', or what we now call the 'reduction of the state vector' or (much

^{*} And so, in what follows, I will refer to states such as $|v \rangle$ and $|h \rangle$ as the measurement eigenstates, on the understanding that these go on to form composite states with the instrument used to detect them.

[†] Although, on a practical note, we should admit that detection, amplification, and counting efficiencies will typically be below 100% at each stage and so will affect the experimental determination of P_v and P_h , such that these are slightly reduced compared with the quantum-mechanical predictions. More on this in Chapter 12.

more familiarly) the ‘collapse of the wavefunction’. In the context of our example, applying this postulate suggests something like

$$\langle \hat{M} \rangle = \frac{1}{2} (1_v + 1_h) \rightarrow 1_v \text{ or } 1_h. \quad (11.32)$$

How is this supposed to work? As we will see, the answer depends on whether you want to interpret the state vector or wavefunction realistically, or not.

Von Neumann defined component III of the measurement process to be the ‘observer’. But if the quantum mechanics described by process 2 applies equally well to classical measuring devices, then there is again no good reason to suppose that it ceases to apply when considering the function of human sense organs, their connections to the brain, and the brain itself. If we simply extend the measurement chain—component II—to include the observer’s sensory apparatus, then the expectation value \hat{M} becomes a mix of equally probable ‘observer’s brain registers 1_v ’ and ‘observer’s brain registers 1_h ’.

But this is just a physical registration and it is at this point—and no earlier, von Neumann argued—that we need to break the infinite regress and invoke the projection postulate. Based on conversations he had had with his Hungarian compatriot Leo Szilard, he suggested that component III represents the observer’s ‘abstract ego’. In other words, process 1—the collapse of the wavefunction—only occurs when the observer becomes *conscious* of the measurement outcome.

This can still be interpreted in two fundamentally different ways. An anti-realist would agree with the description of components I and II and interpret III not as a physical collapse, but as the registering of the measurement outcome and the updating of the observer’s *state of knowledge* about it. This most definitely involves the observer’s conscious mind, but only in a passive way.

But it seems that von Neumann himself held to a different view. Component III is intended as the place where process 1 occurs, considered as a *real physical collapse*. His long conversations with Szilard concerned the latter’s work on entropy reduction in thermodynamic systems through interference by intelligent beings, a variation on Maxwell’s Demon.¹² The philosopher Max Jammer notes that this kind of paper ‘marked the beginning of certain thought-provoking speculations about the effect of a *physical intervention* of mind on matter’.¹³

Wigner’s Friend

But now we need to ask ourselves: Just *who* is the observer? Suppose Alice makes a measurement in the laboratory but renowned Hungarian theorist Eugene Wigner, with whom Alice is a close friend, is delayed in the corridor, having an animated conversation with von Neumann. Alice has modified the measuring instrument we have considered thus far. She has replaced the photon counter with a simple light bulb. If a photon is detected in the vertical channel, the light flashes. If it emerges from the horizontal channel, the light doesn’t flash. She runs the experiment just once, and observes the light flash.

Wigner is still in the corridor. As far as he is concerned, the expectation value of the measurement is still given by Eq. (11.31), but now with 1_v and 1_h replaced by Alice's conscious experience of the light flashing, or not. Wigner now enters the laboratory, and the following conversation ensues.

'Did you see the light flash?' asks Wigner.

'Yes,' replies Alice.

As far as Wigner himself is concerned, the measurement outcome has just registered in his conscious mind and the wavefunction collapses. But, after some reflection, he decides to probe his friend a little further.

'What did you feel about the flash before I asked you?'

Understandably, Alice is starting to get a little irritated. 'I told you already, I *did* see a flash,' she replies, testily.

Not wishing to put any further strain on his relationship with Alice, he decides to accept what she's telling him. He concludes that the wavefunction must have already collapsed *before* he entered the laboratory and asked the question, and the superposition (11.31) that he took to be the mathematically correct description is, in fact, wrong. This superposition 'appears absurd because it implies that my friend was in a state of suspended animation before [she] answered my question'.¹⁴ He wrote:

It follows that the being with a consciousness must have a different role in quantum mechanics than the inanimate measuring device.... It is not necessary to see a contradiction here from the point of view of orthodox quantum mechanics, and there is none if we believe that the alternative is meaningless, whether my friend's consciousness contains either the impression of having seen a flash or of not having seen a flash. However, to deny the existence of the consciousness of a friend to this extent is surely an unnatural attitude, approaching solipsism, and few people, in their hearts, will go along with it.

This is the *paradox of Wigner's friend*. To resolve it we must presume that the irreversible collapse of the wavefunction is triggered by the *first* conscious mind it encounters.

Schrödinger's Cat

As I've already mentioned, Einstein didn't share Schrödinger's realistic vision of the wavefunction as some kind of 'matter wave'. In an exchange of letters during 1935, Einstein sought to persuade Schrödinger to think of the wavefunction in terms of *statistics*. For example, we describe the properties of an atomic gas in terms of physical quantities such as temperature and pressure. But if we consider the gas as a collection of atoms, we can use the classical theories developed by Ludwig Boltzmann and James Clerk Maxwell to deduce expressions for temperature and pressure as the result of *statistical averaging* over a range of atomic motions. In this case, we deal with statistics and probabilities only because we have no way of following the motions of each individual atom in the gas. Of course, we might not be able to account for such motions except in terms of statistics, but this doesn't mean that atoms (and their motions) aren't real.

If quantum probability is, after all, a statistical probability born of ignorance, then there must exist a further underlying reality that we are ignorant of, just as 'hidden' atomic

motions underlie the temperature and pressure of a gas. This was Einstein's point: as this underlying reality makes no appearance in quantum mechanics, then the theory cannot be considered to be complete.

Schrödinger's interpretation couldn't possibly be right, and Einstein sought to dissuade him in a letter dated 8 August 1935. In this letter Einstein asked him to imagine a charge of gunpowder that, at any time over a year, may spontaneously explode. At the beginning of the year, the gunpowder is described by a wavefunction (or a state vector). But how should we describe the situation through the course of the year? Until we look to see what's happened, we would have to regard the wavefunction as a superposition of the wavefunctions corresponding to an explosion, and to a non-explosion. He wrote:¹⁵

Through no art of interpretation can this [wavefunction] be turned into an adequate description of a real state of affairs; [for] in reality there is just no intermediary between exploded and not-exploded.

Schrödinger was eventually persuaded, and came to share Einstein's views. But the gunpowder experiment had set him thinking. As there is nothing in the mathematical formulation of quantum mechanics that accounts for the collapse of the wavefunction, then—just as von Neumann had done—why not imagine that a superposition reaches all the way up the measurement chain? On the 19 August he wrote back:¹⁶

Contained in a steel chamber is a Geigercounter prepared with a tiny amount of uranium, so small that in the next hour it is just as probable to expect one atomic decay as none. An amplified relay provides that the first atomic decay shatters a small bottle of prussic acid. This and—cruelly—a cat is also trapped in the steel chamber. According to the [wavefunction] for the total system, after an hour, *sit venia verbo* [pardon the phrase], the living and dead cat are smeared out in equal measure.

This is the famous, and eternally enduring, paradox of Schrödinger's cat. Einstein was in complete agreement. A total wavefunction consisting of contributions from the wavefunctions of a live and dead cat is surely a fiction. Better to try to interpret the wavefunction in terms of statistics. If the experiment is duplicated, the laboratory filled with hundreds of chambers each containing a cat, then after an hour we predict that in a certain number of these the cat will be dead. The Geiger counter in each box clicks or doesn't click. If it clicks, the relay is activated, the prussic acid is released and the cat is killed. If it doesn't click, the cat survives. Nowhere in this experiment is a cat ever in some kind of peculiar purgatory.

Schrödinger intended the cat paradox as a rather tongue-in-cheek dig at the apparent incompleteness of quantum mechanics, rather than a direct challenge to the Copenhagen interpretation. It does not seem to have elicited any kind of formal response from Bohr.

The community of physicists had in any case moved on by this time, and probably had little appetite for an endless philosophical debate that, in the view of the majority, had already been satisfactorily addressed by Bohr. In the meantime, the 'Copenhagener Geist' had become formalized and enshrined in the quantum formalism.

NOTES

1. Niels Bohr, quoted by Aage Petersen, 'The Philosophy of Niels Bohr', *Bulletin of the Atomic Scientists*, 19 (1963), 12.
2. A careful analysis of Bohr's philosophical influences and writings suggest that he was closer to the tradition known as *pragmatism* than to positivism. Pragmatism, founded by American philosopher Charles Sanders Pierce, has many of the characteristics of positivism in that they both roundly reject metaphysics. There are differences, however. We can think of the positivist doctrine as one of 'seeing is believing': what we can know is limited by what we can observe empirically. The pragmatist doctrine admits a more practical (or, indeed, pragmatic) approach: what we can know is limited not by what we can see, but by what we can *do*. See, for example, Dugald Murdoch, *Niels Bohr's Philosophy of Physics*, Cambridge University Press, Cambridge, UK, 1987.
3. Werner Heisenberg, *The Physical Principles of the Quantum Theory*, University of Chicago Press, Chicago, 1930. Republished in 1949 by Dover, New York. This quote appears in the preface.
4. Max Born and Werner Heisenberg, 'Quantum Mechanics', *Proceedings of the Fifth Solvay Congress, 1928*. English translation from Guido Bacciagaluppi and Antony Valentini, *Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference*, Cambridge University Press, Cambridge, UK, 2009, p. 437.
5. Albert Einstein 'General Discussion', *Proceedings of the Fifth Solvay Congress, 1928*. English translation Bacciagaluppi and Valentini, *Quantum Theory at the Crossroads*, p. 488.
6. Otto Stern, interview with Res Jost, 2 December 1961. Quoted in Abraham Pais, *Subtle is the Lord: The Science and the Life of Albert Einstein*, Oxford University Press, Oxford, 1982, p. 445.
7. See, for example, Jim Baggott, *The Quantum Story: A History in 40 Moments*, Oxford University Press, Oxford, 2011, Chapter 15.
8. Albert Einstein, quoted by Hendrik Casimir in a letter to Abraham Pais, 31 December 1977. Quoted in Pais, *Subtle is the Lord*, p. 449.
9. For a similar (though slightly different) convention, see Richard P. Feynman, Robert B. Leighton, and Matthew Sands, *The Feynman Lectures on Physics*, Vol III, Addison-Wesley, Reading, MA, 1965, pp. 11.9–11.12.
10. John Bell, quoted by Andrew Whitaker, *John Stewart Bell and Twentieth-Century Physics: Vision and Integrity*, Oxford University Press, Oxford, 2016, p. 57.
11. P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th edn, Clarendon Press, Oxford, 1958, p. 35.
12. L. Szilard, 'On Entropy Reduction in a Thermodynamic System by Interference by Intelligent Beings', *Zeitschrift für Physik*, 53 (1929), 840–56. NASA Technical Translation F-16723.
13. Max Jammer, *The Philosophy of Quantum Mechanics*, Wiley, New York, 1974, p. 480. The italics are mine.
14. Eugene Wigner, in I. J. Good (ed.), 'Remarks on the Mind–Body Question', *The Scientist Speculates: An Anthology of Partly-Baked Ideas*, Heinemann, London, 1961, pp. 284–302. This is reproduced in John Archibald Wheeler and Wojciech Hubert Zurek (eds), *Quantum Theory and Measurement*, Princeton University Press, Princeton, NJ, 1983, pp. 168–81. These quotes appear on pp. 176–8.

15. Albert Einstein, letter to Erwin Schrödinger, 8 August 1935. Quoted in Arthur Fine, *The Shaky Game: Einstein, Realism and the Quantum Theory*, 2nd edn, University of Chicago Press, Chicago, 1996, p. 78.
16. Erwin Schrödinger, letter to Albert Einstein, 19 August 1935. Quoted in Fine, *The Shaky Game*, pp. 82–3.

.....

FURTHER READING

- D’Espagnat, Bernard, *Conceptual Foundations of Quantum Mechanics*, 2nd edn, Addison-Wesley, New York, 1989, Part 4.
- Dirac, P. A. M., *The Principles of Quantum Mechanics*, 4th edn, Clarendon Press, Oxford, 1958, pp. 34–41.
- Isham, Chris J., *Lectures on Quantum Theory: Mathematical and Structural Foundations*, Imperial College Press, London, 1995, pp. 175–88.
- Jammer, Max, *The Philosophy of Quantum mechanics*, Wiley, New York, 1974, Chapter 11.
- Susskind, Leonard, & Friedman, Art, *Quantum Mechanics: The Theoretical Minimum*, Penguin, London, 2015, Lectures 6 and 7.
- Von Neumann, John, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton, NJ, 1955, Chapter 6.
- Wheeler, John Archibald, and Zurek, Wojciech Hubert (eds), *Quantum Theory and Measurement*, Princeton University Press, Princeton, NJ, 1983. This is an anthology of papers on the problem of quantum measurement. See especially pp. 152–67 and 168–81.

Einstein, Bohm, Bell, and the Derivation of Bell's Inequality

Entanglement and Quantum Non-locality

By 1935, the Copenhagen interpretation had become the orthodoxy. It was already the default way in which physicists were meant to think about quantum mechanics. Einstein referred to this as 'Talmudic'; a 'religious' philosophy that is to be interpreted only through its qualified priests, who insist on its essential truth, and who will countenance no rivals.¹

The philosopher Karl Popper called it a schism:²

One remarkable aspect of these discussions was the development of a split in physics. Something emerged which may be fairly described as a quantum orthodoxy: a kind of party, or school, or group, led by Niels Bohr, with the very active support of Heisenberg and Pauli; less active sympathizers were Max Born and P[ascual] Jordan and perhaps even Dirac. In other words, all the greatest names in atomic theory belonged to it, except two great men who strongly and consistently dissented: Albert Einstein and Erwin Schrödinger.

If he was to continue his challenge, Einstein needed to find a way to render Bohr's disturbance (or clumsiness) defence either irrelevant or inadmissible. Despite its seeming impossibility, this meant imagining a physical situation in which it is indeed possible, in principle, to acquire knowledge of the physical state of a quantum system without disturbing it in any way. Working with two young theorists, Boris Podolsky and Nathan Rosen, Einstein devised a new challenge that was extraordinarily cunning. It seemed that they had found a way to do the impossible.

Entangled States

Imagine a situation in which two quantum particles interact or are formed together in some physical process, and then move apart. These particles may be photons, for example, emitted in rapid succession from an atom, or they could be electrons or atoms.

For convenience, we'll label these particles as 1 and 2. For our purposes we just need to suppose that, as a result of the operation of some law of conservation, the two particles are each produced in quantum states that are obliged to be *orthogonal*.

At this stage it really doesn't matter what these states are, so let's just imagine that the particles are photons in states of left- and right-circular polarization. Suppose the law of conservation says that if photon 1 is found to be in a state of left-circular polarization, $|L_1\rangle$, then photon 2 must be in a state of right-circular polarization, $|R_2\rangle$, as judged from the perspective of the source (rather than a detector). Similarly, if particle 1 is found to be in the state $|R_1\rangle$, then particle 2 must be in the state $|L_2\rangle$. The reason for this particular choice of combination will become apparent later in this chapter.

The two photons form composite states, $|L_1R_2\rangle = |L_1\rangle|R_2\rangle$ and $|R_1L_2\rangle = |R_1\rangle|L_2\rangle$, which we can presume are both equally probable. We don't know which composite state we're going to get from any specific individual physical event that creates it, but we know from our discussion of the Pauli principle in Chapter 8 that the correct way to proceed in these circumstances is to form these into a normalized superposition (cf. Eq. (8.16)),

$$|\zeta_{12}\rangle = \frac{1}{\sqrt{2}} (|L_1\rangle|R_2\rangle + |R_1\rangle|L_2\rangle), \quad (12.1)$$

where I've used the Greek letter ζ (zeta) to indicate the composite pair state (from the Greek word $\zeta\epsilon\acute{\upsilon}\gamma\omicron\varsigma$, meaning 'pair').

The photons move a long distance apart, each eventually passing through a *linear* polarization analyser (such as a calcite crystal) before being detected, amplified, and counted. Both polarization analysers are aligned along common vertical/horizontal axes, so the possible measurement eigenstates are

$$\begin{aligned} |v_1v_2\rangle &= |v_1\rangle|v_2\rangle && \text{photon 1 vertical/photon 2 vertical} \\ |v_1h_2\rangle &= |v_1\rangle|h_2\rangle && \text{photon 1 vertical/photon 2 horizontal} \\ |h_1v_2\rangle &= |h_1\rangle|v_2\rangle && \text{photon 1 horizontal/photon 2 vertical} \\ |h_1h_2\rangle &= |h_1\rangle|h_2\rangle && \text{photon 1 horizontal/photon 2 horizontal.} \end{aligned} \quad (12.2)$$

We know what to do next. To analyse this situation we must expand the initial composite state $|\zeta_{12}\rangle$ in the basis of the measurement eigenstates:

$$|\zeta_{12}\rangle = |v_1v_2\rangle\langle v_1v_2|\zeta_{12}\rangle + |v_1h_2\rangle\langle v_1h_2|\zeta_{12}\rangle + |h_1v_2\rangle\langle h_1v_2|\zeta_{12}\rangle + |h_1h_2\rangle\langle h_1h_2|\zeta_{12}\rangle. \quad (12.3)$$

I think we know by now where this is heading. So, let's assume a joint measurement operator \hat{M}_{12} with eigenvalues 1_v1_v (the detectors register vertical polarization for photons 1 and 2, respectively), 1_v1_h , 1_h1_v , and 1_h1_h . From Eqs. (11.30) and (11.31), we know we can write the expectation value as

$$\langle\hat{M}_{12}\rangle = P_{vv}1_v1_v + P_{vh}1_v1_h + P_{hv}1_h1_v + P_{hh}1_h1_h, \quad (12.4)$$

where $P_{vv} = \langle v_1 v_2 | \zeta_{12} \rangle|^2$ is the probability of observing the combination vertical/vertical, and so on for P_{vh} , P_{hv} , and P_{hh} .

It will prove useful in what follows to define a generalized *correlation function*, C_{12} , which summarizes the extent of the correlation between the two photons. This is based on the expectation value given by Eq. (12.4), in which we (arbitrarily) assign $1_v = +1$ and $1_h = -1$:

$$C_{12} = P_{vv} - P_{vh} - P_{hv} + P_{hh}. \quad (12.5)$$

Having set up these general expressions—which will prove useful very soon—we can now go on to deduce the individual projection amplitudes using the entries in Table 11.2:

$$\begin{aligned} \langle v_1 v_2 | \zeta_{12} \rangle &= \langle v_1 | \langle v_2 | \frac{1}{\sqrt{2}} (|L_1\rangle |R_2\rangle + |R_1\rangle |L_2\rangle) \\ &= \frac{1}{\sqrt{2}} (\langle v_1 | L_1 \rangle \langle v_2 | R_2 \rangle + \langle v_1 | R_1 \rangle \langle v_2 | L_2 \rangle) \\ &= \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \right) = \frac{1}{\sqrt{2}}. \end{aligned} \quad (12.6)$$

Similarly,

$$\begin{aligned} \langle v_1 h_2 | \zeta_{12} \rangle &= \frac{1}{\sqrt{2}} (\langle v_1 | L_1 \rangle \langle h_2 | R_2 \rangle + \langle v_1 | R_1 \rangle \langle h_2 | L_2 \rangle) \\ &= \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}} \cdot \frac{i}{\sqrt{2}} - \frac{1}{\sqrt{2}} \cdot \frac{i}{\sqrt{2}} \right) = 0, \end{aligned} \quad (12.7)$$

and

$$\begin{aligned} \langle h_1 v_2 | \zeta_{12} \rangle &= 0 \\ \langle h_1 h_2 | \zeta_{12} \rangle &= \frac{1}{\sqrt{2}}. \end{aligned} \quad (12.8)$$

So from (12.3) we have

$$|\zeta_{12}\rangle = \frac{1}{\sqrt{2}} (|v_1\rangle |v_2\rangle + |h_1\rangle |h_2\rangle). \quad (12.9)$$

From Eq. (12.4), the expectation value reduces to

$$\langle \hat{M}_{12} \rangle = P_{vv} 1_v 1_v + P_{hh} 1_h 1_h. \quad (12.10)$$

And from Eq. (12.5) we see that $C_{12} = 1$, which simply means that the outcomes are perfectly correlated.

By the time we make measurements on them, we presume that photons 1 and 2 have moved a long distance apart. We can, in principle, make a measurement on either photon to discover its state. Of course, for each measurement we will only ever see *one* outcome. We must therefore presume that the composite state $|\zeta_{12}\rangle$ collapses to deliver only one outcome: *either* $|v_1\rangle |v_2\rangle$ *or* $|h_1\rangle |h_2\rangle$, such that in a series of repeated measurements on

identically prepared systems we will get either vertical/vertical or horizontal/horizontal 50 percent of the time.

Now suppose we make a measurement on photon 1 and discover that it is vertically polarized. Following von Neumann's logic, this *must* mean that $|\zeta_{12}\rangle$ collapsed to leave photon 2 in a vertically polarized state. Likewise, if we discover that photon 1 is horizontally polarized, this *must* mean that $|\zeta_{12}\rangle$ collapsed to leave photon 2 in a horizontally polarized state. Based on (12.10), there are no other possible outcomes.

In the review article in which Schrödinger introduced his cat paradox, he also stated:³

Any 'entanglement of predictions' that takes place can obviously only go back to the fact that the two [particles] at some earlier time formed in a true sense *one* system, that is [they] were interacting, and have left behind *traces* on each other.

Such a pair of particles are now said to be *entangled*.

We have no way of knowing in advance if photon 1 will be measured to be vertically or horizontally polarized. But this really doesn't matter, for once we know the state of photon 1, we also know the state of photon 2 with certainty, even though we may not have measured it. In other words, *we can discover the state of photon 2 with certainty without disturbing it in any way*. All we have to assume is that any measurement we make on photon 1 in no way affects or disturbs photon 2, which could be an arbitrarily long distance away (say halfway across the universe). We conclude that the state of photon 2 (and by inference, the state of photon 1) *must surely have been determined all along*.

A Reasonable Definition of Reality

In their 1935 paper, which was titled 'Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?', Einstein, Podolsky, and Rosen (EPR) offered a philosophically loaded 'definition' of physical reality:⁴

If, without in any way disturbing a system, we can predict with certainty (i.e. with a probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.

We can see what the authors were trying to do. If the wavefunction is interpreted realistically, then it ought to account for the reality of the physical quantities—such as the states of photons 1 and 2—that it purports to describe. It clearly doesn't. There is nothing in the formulation that describes what these states are *before* we make a measurement on photon 1, so quantum mechanics cannot be complete.

The alternative is to accept that the reality of the state of photon 2 is determined by the nature of a measurement we *choose* to make on a completely different particle an arbitrarily long distance away. Whatever we think might be going on, a realistic interpretation of the wavefunction implies some kind of 'spooky action at a distance'

at odds with the special theory of relativity. EPR argued that: 'No reasonable definition of reality could be expected to permit this.'⁵

Details of this latest challenge were reported in *The New York Times* before the EPR paper was published, in a news article headlined 'Einstein Attacks Quantum Theory'. This provided a non-technical summary of the main arguments, with extensive quotations from Podolsky who, it seems, had been the principal author of the paper.

There is much in the language and nature of the arguments employed in the paper that Einstein appears later to have regretted, especially the reality criterion. He deplored *The New York Times* article and the publicity surrounding it. All the more disappointing, perhaps, as the main challenge presented by EPR does not require this (or any) criterion, though it does rest on the presumption that, however reality is defined, it is presumed to be *local*, meaning that—no matter how they might have been formed—as photons 1 and 2 move apart, they are assumed to exist completely independently of each other. This is sometimes referred to as 'Einstein separability'.

This new challenge sent shockwaves through the small community of quantum physicists. It hit Bohr like a 'bolt from the blue'.⁶ Pauli was furious. Dirac exclaimed: 'Now we have to start all over again, because Einstein proved that it does not work.'⁷

Bohr's response, when it came a short time later, inevitably targeted the reality criterion as the principal weakness. He argued that the stipulation 'without in any way disturbing a system' is essentially ambiguous, since the quantum system is influenced by the very conditions which define its future behaviour. In other words, the composite state $|\zeta_{12}\rangle$ is deliberately set up with coded information based on what we already know from previous experience. This allows us to predict P_{vv} and P_{hh} . The measurements then simply update our knowledge. All is well, provided we don't ask how nature manages this particular conjuring trick.

Einstein was, at least, successful in pushing Bohr to give up his clumsiness defense, and to adopt a more firmly anti-realist position. Those in the physics community who cared about these things seemed to accept that Bohr's response had put the record straight. But not everybody was satisfied.

Hidden Variables

In his debate with Bohr and Schrödinger, Einstein had hinted at a *statistical* interpretation. In his opinion, quantum probabilities, derived as the modulus-squares of the projection amplitudes, actually represent statistical probabilities, averaged over large numbers of physically real particles. We resort to probabilities because we're ignorant of the states of the physically real quantum things.

Einstein toyed with just such an approach in May 1927. This was a modification of quantum mechanics that combined classical wave and particle descriptions, with the wavefunction taking the role of a 'guiding field' (in German, a *Führungsfeld*), guiding or 'piloting' the physically real particles. In this kind of scheme, the wavefunction is responsible for all the wave-like effects, such as diffraction and interference, but the

particles maintain their integrity as localized, physically real entities. Instead of waves *or* particles, as the Copenhagen interpretation demands, Einstein's adaptation of quantum mechanics was constructed from waves *and* particles.

But Einstein lost his enthusiasm for this approach within a matter of weeks of formulating it. It hadn't come out as he'd hoped. The wavefunction had taken on a significance much greater than merely statistical. Einstein thought the problem was that distant particles were exerting some kind of strange force on one another, which he really didn't like. But the real problem was that the guiding field is capable of exerting spooky non-local influences. He withdrew a paper he had written on the approach before it could be published. It survives in the Einstein Archives as a handwritten manuscript.⁸

This experience probably led Einstein to conclude that his initial belief—that quantum mechanics could be completed through a more direct fusion of classical wave and particle concepts—was misguided. He subsequently expressed the opinion that a complete theory could only emerge from a much more radical revision of the entire theoretical structure. He felt that quantum mechanics would eventually be replaced by an elusive grand unified field theory, the search for which took up most of his intellectual energy in the last decades of his life.

This early attempt by Einstein at completing quantum mechanics is known generally as a *hidden variables formulation*, or just a 'hidden variables theory'. It is based on the idea that there is some aspect of the physics that governs what we see in an experiment, but which makes no appearance in the representation. There are, of course, many precedents for this kind of approach in the history of science. As I've already explained, Boltzmann formulated a statistical theory of thermodynamics based on the 'hidden' motions of real atoms and molecules. Likewise, in Einstein's abortive attempt to rethink quantum mechanics, it is the positions and motions of real particles, guided by the wavefunction, that are hidden.

However, in *Mathematical Foundations*, von Neumann presented a proof which appeared to demonstrate that all hidden variable extensions of quantum mechanics are impossible.⁹ This seemed to be the end of the matter. If hidden variables are impossible, why bother even to speculate about them?

And, indeed, silence prevailed for nearly twenty years. The dogmatic Copenhagen view prevailed, seeping into the mathematical formalism and becoming the quantum physicists' conscious or unconscious default interpretation. The physics community moved on and just got on with it, content to 'shut up and calculate'.¹⁰

Then David Bohm broke the silence.

Enter Bohm

In February 1951, Bohm published a textbook, simply called *Quantum Theory*, in which he followed the party line and dismissed the challenge posed by EPR's 'bolt from the blue', much as Bohr had done. But even as he was writing the book he was already having misgivings. He felt that something had gone seriously wrong.

Einstein welcomed the book, and invited Bohm to meet with him in Princeton sometime in the spring of 1951. The doubts over the interpretation of quantum theory that had begun to creep into Bohm's mind now crystallized into a sharply defined problem. 'This encounter had a strong effect on the direction of my research,' Bohm later wrote, 'Because I then became seriously interested in whether a deterministic extension of quantum theory could be found'.¹¹ The Copenhagen interpretation had transformed what was really just a method of calculation into an explanation of reality, and Bohm was more committed to the preconceptions of causality and determinism than perhaps he had realized.

In *Quantum Theory*, Bohm asserted that 'no theory of mechanically determined hidden variables can lead to *all* of the results of the quantum theory'.¹² Bohm went on to develop a derivative of the EPR thought experiment which he published in a couple of papers in 1952 and which he elaborated in 1957 with physicist Yakir Aharonov.¹³ This is based on the idea of fragmenting a diatomic molecule (such as hydrogen, H₂) into two spin-aligned atoms.

Through their efforts, Bohm and Aharonov brought the EPR experiment down from the lofty heights of pure thought and into the practical world of the physics laboratory. In fact, the purpose of their 1957 paper was to claim that experiments capable of measuring correlations between distant entangled particles had already been carried out. For those few physicists paying attention, Bohm's assertion and the notion of a practical test suggested some mind-blowing possibilities.

Enter Bell

John Bell was paying attention. In 1964, he had an insight that was completely to transform questions about the representation of reality at the quantum level. After reviewing and dismissing von Neumann's 'impossibility proof' as flawed and irrelevant, he derived what was to become known as *Bell's inequality*. 'Probably I got that equation into my head and out on to paper within about one weekend,' he later explained. 'But in the previous weeks I had been thinking intensely all around these questions. And in the previous years it had been at the back of my head continually.'¹⁴

The Ingredients

1. Projection amplitudes for photon polarization states, Table 11.2.
2. The complex exponential forms for $\cos A$ and $\sin A$: $\cos A = \frac{1}{2}(e^{iA} + e^{-iA})$ and $\sin A = \frac{1}{2i}(e^{iA} - e^{-iA})$.
3. The trigonometric identity $\cos^2 A - \sin^2 A = \cos 2A$.
4. Bertlmann's socks.

The Recipe

So far in our discussion of the system of entangled photons, we've assumed that the two polarizing analysers used to discover the linear polarization states of photons 1 and 2 are orientated so that they are aligned on common vertical/horizontal axes. But what if we now *rotate* one (or both) of these analysers? In Step (1), we elaborate the projection amplitudes, projection probabilities, expectation value, and correlation function for this situation.

We take a curious diversion in Step (2). Bell was constantly on the lookout for 'everyday' examples of pairs of objects that are spatially separated but whose properties are correlated, as these provide accessible analogues for the EPR experiment. He found a perfect example in the dress sense of one of his colleagues at CERN, Reinhold Bertlmann. So, in this step I will introduce you to Bertlmann's socks. In a paper published in 1981, Bell made use of a series of hypothetical experiments involving prolonged washing of these socks at different temperatures to develop some numerical relationships involving the outcomes, and we consider these in Step (3). These relationships are generalized—in Step (4)—to experiments on pairs of socks, from which Bell's inequality can be derived.

We return to quantum mechanics in Step (5), which details the derivation of Bell's theorem.

Step (1): Measurements with Different Analyser Orientations

We'll start by considering the situation in which the two polarization analysers are both orientated at different angles relative to the (arbitrary) laboratory vertical axis. We suppose that the analyser for photon 1 is orientated at an angle α measured clockwise from the vertical axis. This defines the v'/h' axes (as before, see Fig. 11.3). We suppose that the analyser for photon 2 is orientated at an angle β , defining another set of axes which we denote as v''/h'' . As before, the possible measurement eigenstates are

$$\begin{aligned}
 |v'_1 v''_2\rangle &= |v'_1\rangle |v''_2\rangle \text{ photon 1 vertical}'/\text{photon 2 vertical}'' \\
 |v'_1 h''_2\rangle &= |v'_1\rangle |h''_2\rangle \text{ photon 1 vertical}'/\text{photon 2 horizontal}'' \\
 |h'_1 v''_2\rangle &= |h'_1\rangle |v''_2\rangle \text{ photon 1 horizontal}'/\text{photon 2 vertical}'' \\
 |h'_1 h''_2\rangle &= |h'_1\rangle |h''_2\rangle \text{ photon 1 horizontal}'/\text{photon 2 horizontal}'',
 \end{aligned}
 \tag{12.11}$$

where vertical' indicates that photon 1 emerges from the vertical channel of analyser 1, and vertical'' indicates that photon 2 emerges from the vertical channel of analyser 2, and so on. We can use the entries in Table 11.2 to deduce the projection amplitude $\langle v'_1 v''_2 | \zeta_{12} \rangle$ as follows:

$$\begin{aligned}
\langle v'_1 v'_2 | \zeta_{12} \rangle &= \langle v'_1 | \langle v'_2 | \frac{1}{\sqrt{2}} (|L_1\rangle |R_2\rangle + |R_1\rangle |L_2\rangle) \\
&= \frac{1}{\sqrt{2}} (\langle v'_1 | L_1 \rangle \langle v'_2 | R_2 \rangle + \langle v'_1 | R_1 \rangle \langle v'_2 | L_2 \rangle) \\
&= \frac{1}{\sqrt{2}} \left(\frac{e^{-i\alpha}}{\sqrt{2}} \frac{e^{i\beta}}{\sqrt{2}} + \frac{e^{i\alpha}}{\sqrt{2}} \frac{e^{-i\beta}}{\sqrt{2}} \right) \\
&= \frac{1}{\sqrt{2}} \left(\frac{1}{2} e^{i(\beta-\alpha)} + \frac{1}{2} e^{-i(\beta-\alpha)} \right) \\
&= \frac{1}{\sqrt{2}} \cos \varphi,
\end{aligned} \tag{12.12}$$

where $\varphi = \beta - \alpha$, and we have made use of the complex exponential form for $\cos \varphi$. We note in passing that when $\varphi = 0^\circ$, $\langle v'_1 v'_2 | \zeta_{12} \rangle = \langle v_1 v_2 | \zeta_{12} \rangle$ and from Eq. (12.12) we recover (12.6).

It follows that

$$\begin{aligned}
\langle v'_1 h'_2 | \zeta_{12} \rangle &= -\frac{1}{\sqrt{2}} \sin \varphi \\
\langle h'_1 v'_2 | \zeta_{12} \rangle &= \frac{1}{\sqrt{2}} \sin \varphi \\
\langle h'_1 h'_2 | \zeta_{12} \rangle &= \frac{1}{\sqrt{2}} \cos \varphi.
\end{aligned} \tag{12.13}$$

We learn that for this particular entangled system the projection amplitudes depend (rather neatly!) only on the *differences* between the orientation angles of the polarization analysers.

The corresponding projection probabilities are then

$$\begin{aligned}
P_{vv}(\varphi) &= |\langle v'_1 v'_2 | \zeta_{12} \rangle|^2 = \frac{1}{2} \cos^2 \varphi \\
P_{vh}(\varphi) &= |\langle v'_1 h'_2 | \zeta_{12} \rangle|^2 = \frac{1}{2} \sin^2 \varphi \\
P_{hv}(\varphi) &= |\langle h'_1 v'_2 | \zeta_{12} \rangle|^2 = \frac{1}{2} \sin^2 \varphi \\
P_{hh}(\varphi) &= |\langle h'_1 h'_2 | \zeta_{12} \rangle|^2 = \frac{1}{2} \cos^2 \varphi.
\end{aligned} \tag{12.14}$$

And the expectation value for the joint measurements is, therefore,

$$\langle \hat{M}_{12} \rangle = P_{vv}(\varphi) 1_v 1_v + P_{vh}(\varphi) 1_v 1_h + P_{hv}(\varphi) 1_h 1_v + P_{hh}(\varphi) 1_h 1_h. \tag{12.15}$$

The correlation function for the joint measurement is

$$\begin{aligned}
C_{12}(\varphi) &= P_{vv}(\varphi) - P_{vh}(\varphi) - P_{hv}(\varphi) + P_{hh}(\varphi) \\
&= \cos^2 \varphi - \sin^2 \varphi \\
&= \cos 2\varphi,
\end{aligned} \tag{12.16}$$

where we have made use of the trigonometric identity $\cos^2 A - \sin^2 A = \cos 2A$.

Notice how the correlation now varies as we change the difference angle between the analyser orientations, from $C_{12}(\varphi) = +1$ (perfect correlation) when $\varphi = 0^\circ$, to $C_{12}(\varphi) = 0$ (no correlation) when $\varphi = 45^\circ$, to $C_{12}(\varphi) = -1$ (perfect anti-correlation) when $\varphi = 90^\circ$.

Step (2): Bertlmann's Socks

The photons are correlated, but is this really so mysterious? In a paper titled 'Bertlmann's Socks and the Nature of Reality', published in 1981, Bell wrote:¹⁵

The philosopher in the street, who has not suffered a course in quantum mechanics, is quite unimpressed by Einstein–Podolsky–Rosen correlations. He can point to many examples of similar correlations in everyday life. The case of Bertlmann's socks is often cited. Dr Bertlmann likes to wear two socks of different colours. Which colour he will have on a given foot on a given day is quite unpredictable. But when you see that the first sock is pink you can be already sure that the second sock will not be pink. Observation of the first, and experience of Bertlmann, gives immediate information about the second. There is no accounting for tastes, but apart from that there is no mystery here. And is not this EPR business just the same?

This situation is illustrated in Fig. 12.1.*

Let's suppose that Dr Bertlmann is a physicist who is very interested in the physical properties of his socks. Imagine that he has secured a contract from a consumer research organization to study how his socks stand up to the rigours of prolonged washing at different temperatures. Being a theoretical physicist, he knows that he can discover some simple relationships between the *numbers* of socks that pass (+ result) or fail (– result) such tests without actually having to perform them using real socks and real washing machines. This makes his study inexpensive and therefore attractive to his sponsors.

Step (3): Washing Socks and Outcome Spaces

Bertlmann decides to subject his left socks (which we label collectively as socks 1) to three different tests:

- Test α , washing for 1000 cycles at 0°C ;
- Test β , washing for 1000 cycles at $22\frac{1}{2}^\circ\text{C}$; and
- Test γ , washing for 1000 cycles at 45°C .

* Reinhard Bertlmann was a colleague of Bell's at CERN in the early 1980s. Bertlmann had decided some time before that it was 'crazy' to wear matching socks and that the correct thing to do is wear socks of different colours on each foot. See Andrew Whitaker, *John Stewart Bell and Twentieth-Century Physics: Vision and Integrity*, Oxford University Press, 2016, p. 350.

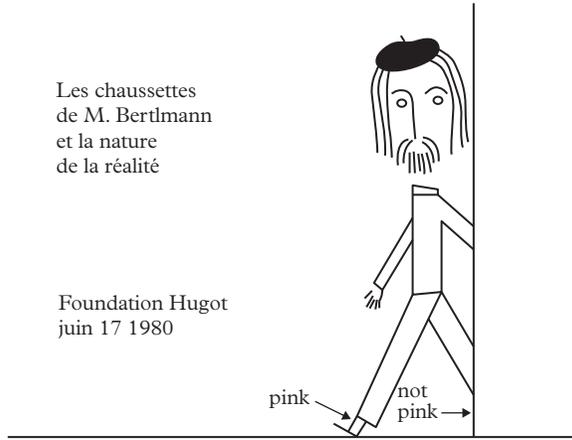


Figure 12.1 Bertlmann's socks and the nature of reality.

Source: Reproduced with permission from Bell, J. S. (1981). Bertlmann's socks and the nature of reality. *J. Phys. Symposiums*, 42(C2):41–62.

Of course, we are assuming that the properties of the socks are uniform—there is no physical difference (perhaps other than colour) between socks in a large collection. We can define the result or outcome 'spaces' for these three tests as follows:

$$\begin{array}{ccc}
 \begin{array}{|c|} \hline \alpha_+ \\ \hline \alpha_- \\ \hline \end{array} & \begin{array}{|c|c|} \hline \beta_+ & \beta_- \\ \hline \end{array} & \begin{array}{|c|} \hline \gamma_+ \\ \hline \gamma_- \\ \hline \end{array} \\
 \alpha & \beta & \gamma
 \end{array} \tag{12.17}$$

Think of it this way. When Bertlmann discovers that a sock successfully passes the α test, he places a tick in the α_+ space; if it fails he places a tick in the α_- space. Likewise for all the other spaces.

Bertlmann now defines three experiments. In experiment 1, he determines how many socks pass test α and fail test β . This number is denoted as $n(\alpha_+, \beta_-)$. You can think of this as the count of the number of ticks that lie in the overlapping α_+ and β_- outcome spaces. In experiment 2 he determines the number of socks passing β and failing γ , $n(\beta_+, \gamma_-)$. And in experiment 3 he determines the number passing α and failing γ , $n(\alpha_+, \gamma_-)$. These three experiments map to the outcome spaces as follows:

$$\begin{array}{ccc}
 \begin{array}{|c|} \hline \begin{array}{|c|} \hline \alpha_+ \\ \hline \alpha_- \\ \hline \end{array} \\ \hline \end{array} & \begin{array}{|c|} \hline \begin{array}{|c|c|} \hline \beta_+ & \beta_- \\ \hline \end{array} \\ \hline \end{array} & \begin{array}{|c|} \hline \begin{array}{|c|} \hline \gamma_+ \\ \hline \gamma_- \\ \hline \end{array} \\ \hline \end{array} \\
 n(\alpha_+, \beta_-) & n(\beta_+, \gamma_-) & n(\alpha_+, \gamma_-)
 \end{array} \tag{12.18}$$

where the shaded areas indicate the overlapping outcome spaces in each experiment.

We can see from this that $n(\alpha_+, \beta_-)$ is the sum of two subspaces:

$$\begin{array}{c}
 \begin{array}{|c|c|} \hline \text{○} & \text{■} \\ \hline \text{○} & \text{○} \\ \hline \end{array} & = & \begin{array}{|c|c|} \hline \text{○} & \text{■} \\ \hline \text{○} & \text{○} \\ \hline \end{array} & + & \begin{array}{|c|c|} \hline \text{○} & \text{○} \\ \hline \text{○} & \text{■} \\ \hline \end{array} \\
 n(\alpha_+, \beta_-) & & n(\alpha_+, \beta_-, \gamma_+) & & n(\alpha_+, \beta_-, \gamma_-)
 \end{array} \tag{12.19}$$

Similarly,

$$\begin{array}{c}
 \begin{array}{|c|c|} \hline \text{■} & \text{○} \\ \hline \text{○} & \text{○} \\ \hline \end{array} & = & \begin{array}{|c|c|} \hline \text{■} & \text{○} \\ \hline \text{○} & \text{○} \\ \hline \end{array} & + & \begin{array}{|c|c|} \hline \text{○} & \text{○} \\ \hline \text{■} & \text{○} \\ \hline \end{array} \\
 n(\beta_+, \gamma_-) & & n(\alpha_+, \beta_+, \gamma_-) & & n(\alpha_-, \beta_+, \gamma_-)
 \end{array} \tag{12.20}$$

So, if we now add $n(\alpha_+, \beta_-)$ and $n(\beta_+, \gamma_-)$, from (12.19) and (12.20), we get

$$\begin{array}{c}
 \begin{array}{|c|c|} \hline \text{■} & \text{○} \\ \hline \text{○} & \text{○} \\ \hline \end{array} & = & \begin{array}{|c|c|} \hline \text{■} & \text{○} \\ \hline \text{○} & \text{○} \\ \hline \end{array} & + & \begin{array}{|c|c|} \hline \text{○} & \text{○} \\ \hline \text{○} & \text{■} \\ \hline \end{array} \\
 n(\alpha_+, \beta_-) + n(\beta_+, \gamma_-) & & n(\alpha_+, \gamma_-) & & n(\alpha_+, \beta_-, \gamma_-) + n(\alpha_-, \beta_+, \gamma_-)
 \end{array} \tag{12.21}$$

Or

$$\begin{array}{c}
 \begin{array}{|c|c|} \hline \text{■} & \text{○} \\ \hline \text{○} & \text{○} \\ \hline \end{array} & \geq & \begin{array}{|c|c|} \hline \text{■} & \text{○} \\ \hline \text{○} & \text{○} \\ \hline \end{array} \\
 n(\alpha_+, \beta_-) + n(\beta_+, \gamma_-) & & n(\alpha_+, \gamma_-)
 \end{array} \tag{12.22}$$

Step (4): Generalize for Experiments on Pairs of Socks

Astute readers will have already spotted the flaw in Bertlmann’s reasoning. We really have no way of knowing if any given sock will pass one test and at the same time fail another. If we try to perform a sequence of tests on any individual sock, then we can’t be sure that surviving the rigours of 1000 washing cycles will leave the sock in its pristine state, ready for a subsequent test. And if failing a test means that the sock is destroyed, then it is obviously unavailable for any further test.

But then Bertlmann remembers that his socks always come in *pairs*. Aside from differences in colour, if the socks in each pair are assumed to have otherwise identical physical properties, then we can safely assume that the result of a test performed on sock 2 *implies* that the same result would have been obtained for sock 1, even though we haven’t performed this test on it directly. He must further assume that *whatever test he chooses to perform on sock 2 in no way affects the outcome of any other test he might perform on sock 1*, but this seems so obviously valid that he doesn’t give it a second thought.

Now the three different sets of experiments are carried out on three samples containing the same total number of pairs of socks, N . In experiment 1, for each pair, sock 1 is subjected to test α and sock 2 is subjected to test β . If sock 2 fails test β , this implies that sock 1 would also have failed test β had it been performed on 1. The number of pairs of socks for which sock 1 passes test α and sock 2 fails test β , which Bertlmann denotes as $N_{+-}(\alpha, \beta)$, must be equal to the (theoretical) number of socks 1 which pass test α and fail test β , i.e. $N_{+-}(\alpha, \beta) = n(\alpha_+, \beta_-)$. The same logic follows for experiments 2 and 3.

Bertlmann can now generalize this result for any batch of pairs of socks. By dividing each of these numbers by the total number of pairs of socks N he arrives at the relative frequencies with which each joint result is obtained. He identifies these relative frequencies as probabilities for obtaining the results for experiments yet to be performed on any batch of pairs of socks that, statistically, have the same properties, i.e. $P_{+-}(\alpha, \beta) = N_{+-}(\alpha, \beta) / N$, and so on. From Eq. (12.22) he is led to the inescapable conclusion

$$P_{+-}(\alpha, \beta) + P_{+-}(\beta, \gamma) \geq P_{+-}(\alpha, \gamma). \quad (12.23)$$

This is *Bell's inequality*.

As we can see, this has nothing whatsoever to do with quantum mechanics or hidden variables. It is simply a logical conclusion derived from the relationships between independent sets of numbers and their related probabilities. Any pair of socks which pass α and fail γ will contribute to the probability $P_{+-}(\alpha, \gamma)$. But such a pair will also either pass or fail β , contributing either to $P_{+-}(\beta, \gamma)$ or $P_{+-}(\alpha, \beta)$. Thus, there is simply no way that $P_{+-}(\alpha, \gamma)$ can exceed the sum $P_{+-}(\alpha, \beta) + P_{+-}(\beta, \gamma)$.

Step (5): Bell's Theorem: Quantum Non-locality

This has, no doubt, been an illuminating diversion, but we need to get back to quantum mechanics. Actually, this is really quite straightforward. Instead of experimenting with pairs of socks, we experiment with pairs of photons. Instead of washing at different temperatures, we perform experiments with different orientations of the polarization analysers.

Our three experiments are now:

	Orientation of analyser 1	Orientation of analyser 2	Difference in orientations (φ)
Experiment 1	α	β	$\varphi_1 = \beta - \alpha$
Experiment 2	β	γ	$\varphi_2 = \gamma - \beta$
Experiment 3	α	γ	$\varphi_3 = \gamma - \alpha$

Analysers 1 and 2 refer to the polarization analysers through which photons 1 and 2, respectively, will pass. Again, the orientation angles α, β, γ are measured relative to the (arbitrary) laboratory vertical axis.

We can follow precisely the same logic, replacing pass/fail (+/-) results with detection from the vertical/horizontal channels of the analysers (*v/h*). We retain the assumption that *the relative orientation of analyser 2 can in no way affect the outcome of the measurement performed on photon 1*. In essence, this means that the photons are assumed to be ‘Einstein separable’, just like Bertlmann’s socks. We invoke the existence of some kind of hidden variable which governs the linear polarization properties of the two photons, such that their properties are predetermined *before* they pass through the analysers and are detected. The photons are said to be *locally real*, meaning that their properties are determined all along and photon 1 cannot be influenced by whatever we choose to do to photon 2, and vice versa.

If we accept this, then from Eq. (12.23), we have

$$P_{vh}(\varphi_1) + P_{vh}(\varphi_2) \geq P_{vh}(\varphi_3). \quad (12.24)$$

Combining Eqs. (12.14) and (12.24) gives

$$\frac{1}{2} \sin^2 \varphi_1 + \frac{1}{2} \sin^2 \varphi_2 \geq \frac{1}{2} \sin^2 \varphi_3. \quad (12.25)$$

In these experiments we’re free to set whatever orientation angles we like for the analysers. So, let’s set $\alpha = 0^\circ$, $\beta = 22\frac{1}{2}^\circ$, and $\gamma = 45^\circ$, such that $\varphi_1 = 22\frac{1}{2}^\circ$, $\varphi_2 = 22\frac{1}{2}^\circ$, and $\varphi_3 = 45^\circ$. From (12.25) we get

$$\frac{1}{2} \sin^2 22\frac{1}{2}^\circ + \frac{1}{2} \sin^2 22\frac{1}{2}^\circ \geq \frac{1}{2} \sin^2 45^\circ \text{ or } 0.146 \geq 0.250. \quad (12.26)$$

The conclusion is inescapable. For this particular configuration of the analysers, *quantum mechanics predicts that Bell’s inequality should be violated*.

Bell’s inequality is quite general. It does not depend on what kind of hidden variable theory we might devise, so long as it is locally real. This generality allowed Bell to formulate a ‘no-go’ theorem:¹⁶

If the [hidden variable] extension is local it will not agree with quantum mechanics, and if it agrees with quantum mechanics it will not be local.

A complementary no-go theorem was devised in 1967 by Simon Kochen and Ernst Specker.¹⁷

The Aspect Experiments

The real repercussions of Bell’s 1966 papers were felt through the work of a small group of theoreticians and experimentalists who had read the papers and had become obsessed

with the problem that they posed, and the first direct tests of Bell's inequality were performed in 1972, by Stuart Freedman and John Clauser. These experiments produced the violations of Bell's inequality predicted by quantum mechanics but, because of some further assumptions that were necessary in order to extrapolate the data, only a weaker form of the inequality was tested.

Other results followed, but the first comprehensive experiments designed specifically to test the general form of Bell's inequality were those performed by Alain Aspect and his colleagues Philippe Grangier, Gérard Roger, and Jean Dalibard, at the Institut d'Optique Théoretique et Appliquée, Université Paris-Sud in Orsay, in 1981 and 1982.^{18,19}

Aspect and his colleagues settled on excited calcium atoms as the source of entangled photons. In the lowest energy 'ground' electronic state of the calcium atom, the outermost 4s orbital is filled with two spin-paired electrons ($4s^2$). If one of these electrons absorbs a photon of the right wavelength, then the electron is excited to a higher-energy 4p orbital. In this process, the photon that is absorbed imparts a quantum of angular momentum, and this appears as orbital angular momentum of the excited electron, the value of L increasing by 1. If there is no change in the spin orientations of the two electrons, the excited state is still a singlet state, S is equal to 0, and, since $L = 1$, there is only one possible value for \mathcal{J} : $\mathcal{J} = 1$. This excited state is labeled $4s^1 4p^1(^1P_1)$ (cf. the discussion of the electronic states of helium in Chapter 8).

Now suppose that it is possible to excite a second electron (the one 'left behind' in the 4s orbital) also into this same excited 4p orbital, but in a way that maintains the alignment of the electron spins. In other words, we create a doubly excited state in which the electron spins remain paired ($4p^2$). This gives rise to three different electronic states corresponding to the three different ways of combining the angular momentum vectors. In one of these the orbital angular momentum vectors of the individual electrons cancel, $L = 0$ and, since $S = 0$, we have $\mathcal{J} = 0$. This particular doubly excited state is labeled $4p^2(^1S_0)$.

This doubly excited state undergoes a rapid cascade emission through the intermediate $4s^1 4p^1(^1P_1)$ state to return to the ground state (see Fig. 12.2). Two photons are emitted. Because the quantum number \mathcal{J} changes from $0 \rightarrow 1 \rightarrow 0$ in the cascade, the net angular momentum of the photon pair must be zero: they are emitted in opposite states of circular polarization, either $|L_1 R_2\rangle$ or $|R_1 L_2\rangle$ (as judged from the perspective of the source), described by the superposition in Eq. (12.1). The photons are entangled.

In fact, the photons have wavelengths in the visible region. Photon 1, from the $4p^2(^1S_0) \rightarrow 4s^1 4p^1(^1P_1)$ transition, has a wavelength of 551.3 nm (green) and photon 2, from the $4s^1 4p^1(^1P_1) \rightarrow 4s^2(^1S_0)$ transition, has a wavelength of 422.7 nm (blue).

Aspect and his colleagues used two high-power lasers to produce the excited calcium atoms, which were formed in an atomic 'beam', produced by passing gaseous calcium from a high-temperature oven through a tiny hole into a vacuum chamber. Subsequent collimation of the atoms entering the sample chamber provided a well-defined beam of atoms. The low density of atoms at the point of intersection with the laser beams ensured that the calcium atoms did not collide with each other or with the walls of the chamber before absorbing and subsequently emitting photons.

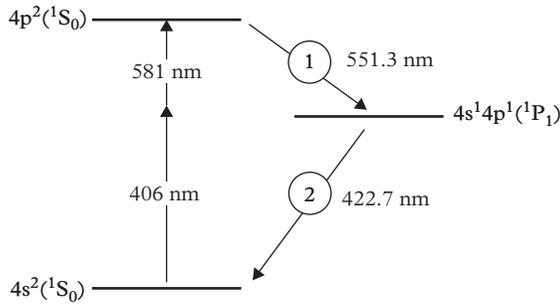


Figure 12.2 *Electronic states of atomic calcium used to generate pairs of entangled photons in tests of Bell’s inequality.*

The physicists monitored the light emitted in opposite directions from the atomic beam source, using coloured filters to isolate the green photons (photons 1) on the left and the blue photons (photons 2) on the right. The photons were then passed into an arrangement consisting of two polarization analysers, four photomultipliers to amplify the signals from the detected photons, and electronic devices designed to detect and record *coincident* signals from the photomultipliers.

Each polarization analyser was mounted on a platform which allowed it to be rotated about its optical axis. Experiments could therefore be performed for different relative orientations of the two analysers, placed about 13 metres apart. The electronics were set to look for coincidences in the arrival and detection of photons 1 and 2 within a time window of just 20 ns. Any kind of ‘spooky’ signal passed between the photons, ‘informing’ photon 2 of the polarization state of photon 1, for example, would therefore need to travel the 13 metres between the detectors within this time window. In fact, it takes about twice this amount of time for a signal moving at the speed of light to cover this distance. The measurements were therefore ‘space-like’ separated.

Aspect and his colleagues measured the correlation $C_{12}(\varphi)$ for seven different sets of analyser orientations. Their results are shown in Fig. 12.3. We saw from Eq. (12.16) that the quantum-mechanical prediction for $C_{12}(\varphi)$ is simply $\cos 2\varphi$, and this is the curve plotted through the data points, after corrections for experimental imperfections.* As anticipated, the predictions demonstrate that perfect correlation and perfect anti-correlation were not quite realized in these experiments. However, it is quite clear that the measured values of $C_{12}(\varphi)$ agree well with the predictions.

* Not all the photons could be physically ‘gathered’ in the detection system, the polarization analysers didn’t transmit all the photons incident on them, and some photons ‘leaked’ through the wrong analyser channels. All these inefficiencies were established in a series of calibrations.

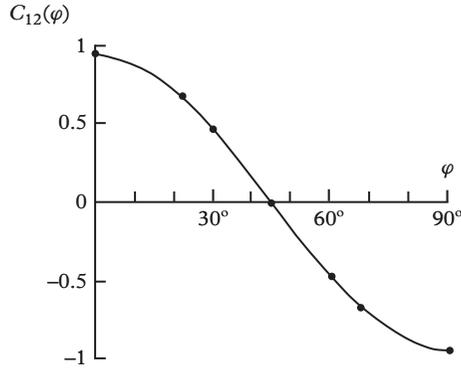


Figure 12.3 Results of measurements of the correlation between entangled photons for different orientations of the polarization analysers. The data points include error bars. The curve is the quantum-mechanical prediction—Eq. (12.16)—modified to take account of experimental inefficiencies.

Source: Reprinted with permission from Aspect, A. et al. Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities. *Physical Review Letters*, 49(2):91-93, © 1982. 1982 by the American Physical Society.

The physicists also tested a generalized version of Bell's inequality based on four different analyser configurations (with angles $\alpha, \beta, \gamma, \delta$), applicable to non-ideal experiments in which perfect correlation or anti-correlation can't be achieved. I won't derive this here, but it is

$$|C_{12}(\varphi_1) - C_{12}(\varphi_4)| + |C_{12}(\varphi_2) + C_{12}(\varphi_3)| \leq 2. \tag{12.27}$$

The orientations were $\varphi_1 = \beta - \alpha = 22 \frac{1}{2}^\circ$, $\varphi_2 = \beta - \gamma = -22 \frac{1}{2}^\circ$, $\varphi_3 = \delta - \gamma = 22 \frac{1}{2}^\circ$, and $\varphi_4 = \delta - \alpha = 67 \frac{1}{2}^\circ$. We can use the general form for $C_{12}(\varphi)$ given in Eq. (12.16) to predict that

$$|\cos 2\varphi_1 - \cos 2\varphi_4| + |\cos 2\varphi_2 + \cos 2\varphi_3| \leq 2 \tag{12.28}$$

$$\left| \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \right| + \left| \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \right| = \frac{4}{\sqrt{2}} = 2\sqrt{2} \leq 2.$$

And we see that, once again, quantum mechanics predicts that the inequality is violated. Aspect and his colleagues obtained the result 2.697 ± 0.015 , a violation of the inequality by 83% of the theoretical maximum ($2\sqrt{2} = 2.828$).

You will note that nowhere in the above has it been necessary to introduce a specific local hidden variable theory to compare and contrast with the predictions of quantum mechanics. This might make you somewhat suspicious. If so, you can allay your suspicions by taking a quick look at Appendix 8, which summarizes a very simple

(but intuitive) local hidden variables theory. This predicts results which do not violate Bell's inequality.

Closing the Loopholes

Of course, this was not the end of the matter. For those physicists with deeply held realist convictions, there just had to be something else going on. More questions were asked: 'What if the hidden variables are somehow influenced by the way the experiment is set up?' This was just the first in a series of 'loopholes', invoked in attempts to argue that these results didn't necessarily rule out all possible local hidden variable theories.

Aspect himself had anticipated this first loophole, and he and his colleagues performed further experiments to close it off. The experimental arrangement was modified to include devices which could randomly switch the paths of the photons, directing each of them towards analysers orientated at different angles. This prevented the photons from 'knowing' in advance along which path they would be traveling, and hence through which analyser they would eventually pass. This is equivalent to changing the relative orientations of the two analysers *while the photons are in flight*. It made no difference. Bell's inequality was still violated.²⁰

The problem can't be made to go away simply by increasing the distance between the source of the entangled particles and the detectors. Experiments have been performed with detectors located in Bellevue and Bernex, two small Swiss villages outside Geneva almost 11 kilometers apart.²¹ Subsequent experiments placed detectors in La Palma and Tenerife in the Canary Islands, 144 kilometers apart. Bell's inequality was still violated.²²

Okay, but what if the hidden variables are still somehow sensitive even to random choices in the experimental setup, simply because these choices are made on the same timescale? In experiments reported in 2018, the experimental settings were determined by the colours of photons detected from distant quasars, the active nuclei of distant galaxies. The random choice of settings was therefore already made nearly 8 billion years *before* the experiment was performed, as this is how long it took for the trigger photons to reach the Earth. Bell's inequality was still violated.²³

There are other loopholes, and these too have been closed off in experiments involving both entangled photons and ions. Experiments involving entangled *triplets* of photons performed in 2000 ruled out all manner of locally realistic hidden variable theories without recourse to Bell's inequality.²⁴

If we want to adopt a realistic interpretation, then it seems we must accept that this reality is determinedly *non-local*.

Leggett's Inequality: Crypto Non-local Hidden Variables

But can we still meet reality halfway? In these experiments, we assume that the properties of the entangled particles are governed by some, possibly very complex, set of hidden variables. These possess unique values that determine the quantum states of the particles

and their subsequent interactions with the measuring devices. We further assume that the particles are formed with a statistical distribution of these variables determined only by the physics and not by the way the experiment is set up.

Local hidden variable theories are characterized by two further assumptions. In the first, we assume (as did EPR) that the *outcome* of the measurement on particle 1 can in no way affect the outcome of the measurement on 2, and vice versa. In the second, we assume that the *setting* of the device we use to make the measurement on 1 can in no way affect the outcome of the measurement on 2, and vice versa.

The experimental violations of Bell's inequality show that one or other (or both) of these assumptions is invalid. But they don't tell us which.

In a paper published in 2003, Anthony Leggett chose to drop the setting assumption. This means that the behaviour of the particles and the outcomes of subsequent measurements *is* assumed to be influenced by the way the measuring devices are set up. This is still all very spooky and highly counter-intuitive:²⁵

nothing in our experience of physics indicates that the orientation of distant [measuring devices] is either more or less likely to affect the outcome of an experiment than, say, the position of the keys in the experimenter's pocket or the time shown by the clock on the wall.

By keeping the outcome assumption, we define a class of non-local hidden variable theories in which the individual particles possess defined properties before the act of measurement. What is actually measured will of course depend on the settings, and changing these settings will somehow affect the behaviour of distant particles (hence, 'non-local'). Leggett referred to this broad class of theories as 'crypto' non-local hidden variable theories.

He went on to show that dropping the setting assumption is in itself still insufficient to reproduce all the results of quantum mechanics. Just as Bell had done in 1964, he derived an inequality that is valid for all classes of crypto non-local hidden variable theories but which is predicted to be violated by quantum mechanics. At stake then was the rather simple question of whether quantum particles have the properties we assign to them *before the act of measurement*. Put another way, here was an opportunity to test whether quantum particles have what we might want to consider as 'real' properties *before* they are measured.

The results of experiments designed to test Leggett's inequality were reported in 2007 and, once again, the answer is pretty unequivocal. For a specific arrangement of the settings in these experiments, Leggett's inequality demands a result which is less than or equal to 3.779. Quantum mechanics predicts 3.879, a violation of less than 3%. The experimental result was 3.8521, with an error of ± 0.0227 . Leggett's inequality was violated.²⁶ Several variations of experiments to test Leggett's inequality have been performed more recently. All confirm this result.

The debate continues. But we must acknowledge that in any realistic interpretation in which the wavefunction is assumed to represent the real physical state of a quantum system, *the wavefunction must be non-local*.

NOTES

1. Albert Einstein, letter to Erwin Schrödinger, 19 June 1935. Quoted in Arthur Fine, *The Shaky Game: Einstein, Realism and the Quantum Theory*, 2nd edn, University of Chicago Press, Chicago, 1996, p. 69.
2. Karl R. Popper, *Quantum Theory and the Schism in Physics*, Unwin Hyman, London, 1982, pp. 99–100.
3. Erwin Schrödinger, ‘The Present Situation in Quantum Mechanics: A Translation of Schrödinger’s “Cat Paradox” Paper’ (translated by John D. Trimmer), *Proceedings of the American Philosophical Society*, 124 (1980), 323–8. This paper is reproduced in John Archibald Wheeler and Wojciech Hubert Zurek (eds), *Quantum Theory and Measurement*, Princeton University Press, Princeton, NJ, 1983, pp. 152–67. This quote appears on p. 161.
4. Albert Einstein, Boris Podolsky, and Nathan Rosen, ‘Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?’, *Physical Review*, 47 (1935), 777–80. This paper is reproduced in Wheeler and Zurek, *Quantum Theory and Measurement*, pp. 138–41. This quote appears on p. 138.
5. Einstein, Podolsky, and Rosen, in Wheeler and Zurek, *Quantum Theory and Measurement*, pp. 138–41. This quote appears on p. 141.
6. Léon Rosenfeld, in Stefan Rozenthal (ed.), *Niels Bohr: His Life and Work as Seen by His Friends and Colleagues*, North-Holland, Amsterdam, 1967, pp. 114–36. Extract reproduced in Wheeler and Zurek, *Quantum Theory and Measurement*, pp. 137 and 142–3. This quote appears on p. 142.
7. Paul Dirac, interview with Niels Bohr, 17 November 1962, *Archive for the History of Quantum Physics*. Quoted in Mara Beller, *Quantum Dialogue*, University of Chicago Press, Chicago, 1999, p. 145.
8. Darrin W. Belousek, ‘Einstein’s 1927 Unpublished Hidden-Variable Theory: Its Background, Context and Significance’, *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 27 (1996), 437–61. Peter Holland takes a closer look at Einstein’s reasons for rejecting this approach in ‘What’s Wrong with Einstein’s 1927 Hidden-Variable Interpretation of Quantum Mechanics’, *Foundations of Physics*, 35 (2005), 177–96; arXiv:quant-ph/0401017v1, 5 January 2004.
9. ‘It should be noted that we need not go any further into the mechanism of the “hidden parameters”, since we now know that the established results of quantum mechanics can never be re-derived with their help.’ John von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton, NJ, 1955, p. 324.
10. The phrase ‘shut up and calculate’ is frequently attributed to Richard Feynman, but it appears to have been coined by N. David Mermin. As a research student studying quantum mechanics in the 1950s, Mermin’s questions about meaning and interpretation were rebuffed by his professors: “‘You’ll never get a PhD if you allow yourself to be distracted by such frivolities,” they kept advising me, “so get back to serious business and produce some results.” “Shut up”, in other words, “and calculate.” And so I did, and probably turned out much the better for it. At Harvard, they knew how to administer tough love in those olden days.’ N. David Mermin, ‘Could Feynman Have Said This?’, *Physics Today*, May 2004, 10–11.
11. According to Basil Hiley, one of Bohm’s long-term collaborators, Bohm said of his meeting with Einstein: ‘After I finished [*Quantum Theory*] I felt strongly that there was something seriously wrong. Quantum theory had no place in it for an adequate notion of an individual

- actuality. My discussions with Einstein clarified and reinforced my opinion and encouraged me to look again.' Quoted by Basil Hiley, personal communication to the author, 1 June 2009.
12. David Bohm, *Quantum Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1951, p. 623.
 13. D. Bohm and Y. Aharonov, 'Discussion of Experimental Proof for the Paradox of Einstein, Rosen, and Podolsky', *Physical Review*, **108** (1957), 1070.
 14. John Bell, in P. C. W. Davies and J. R. Brown (eds), *The Ghost in the Atom*, Cambridge University Press, Cambridge, UK, 1986, p. 57.
 15. John Bell, 'Bertlmann's Socks and the Nature of Reality', *Journal de Physique Colloque C2*, Supplement 3, **42** (1981), 41–61. Reproduced in J. S. Bell, *Speakable and Unsayable in Quantum Mechanics*, Cambridge University Press, Cambridge, UK, 1987, pp. 139–58. This quote appears on p. 139.
 16. John Bell, 'Locality in Quantum Mechanics: Reply to Critics', *Epistemological Letters*, November 1975, 2–6. This paper is reproduced in Bell, *Speakable and Unsayable in Quantum Mechanics*, pp. 63–6. This quote appears on p. 65.
 17. Simon Kochen and E. P. Specker, 'The Problem of Hidden Variables in Quantum Mechanics', *Journal of Mathematics and Mechanics*, **17** (1967), 59–87.
 18. Alain Aspect, Philippe Grangier, and Gérard Roger, 'Experimental Tests of Realistic Local Theories via Bell's Theorem', *Physical Review Letters*, **47** (1981), 460–3.
 19. Alain Aspect, Philippe Grangier, and Gérard Roger, 'Experimental Realization of Einstein–Podolsky–Rosen–Bohm Gedankenexperiment: A New Violation of Bell's Inequalities', *Physical Review Letters*, **49** (1982), 91–4.
 20. Alain Aspect, Jean Dalibard, and Gérard Roger, 'Experimental Test of Bell's Inequalities Using Time-Varying Analyzers', *Physical Review Letters*, **49** (1982), 1804–7.
 21. W. Tittel, J. Brendel, N. Gisin, and H. Zbinden, 'Long-Distance Bell-Type Tests Using Energy-Time Entangled Photons', *Physical Review A*, **59** (1999), 4150–63.
 22. Thomas Scheidl, Rupert Ursin, Johannes Kofler, Sven Ramelow, Xiao-Song Ma, Thomas Herbst, et al., 'Violation of Local Realism with Freedom of Choice', *Proceedings of the National Academy of Sciences*, **107** (2010), 19708–13.
 23. Dominik Rauch, Johannes Handsteiner, Armin Hochrainer, Jason Gallicchio, Andrew S. Friedman, Calvin Leung, et al., 'Cosmic Bell Test Using Random Measurement Settings from High-Redshift Quasars', *Physical Review Letters*, **121** (2018), 080403.
 24. Jian-Wei Pan, Dik Bouwmeester, Matthew Daniell, Harald Weinfurter, and Anton Zeilinger, 'Experimental Test of Quantum Nonlocality in Three-Photon Greenberger–Horne–Zeilinger Entanglement', *Nature*, **403** (2000), 515–19.
 25. A.J. Leggett, 'Nonlocal Hidden-variable Theories and Quantum Mechanics: An Incompatibility Theorem', *Foundations of Physics*, **33** (2003), 1469–93. This quote appears on pp. 1474–5.
 26. Simon Gröblacher, Tomasz Paterek, Rainer Kaltenbaek, Caslav Brukner, Marek Zukowski, Markus Aspelmeyer, et al., 'An Experimental Test of Non-local Realism', *Nature*, **446** (2007), 871–5. In case you were wondering, Bell's inequality is violated in these experiments, too.

.....

FURTHER READING

- Aczel, Amir D., *Entanglement: The Greatest Mystery in Physics*, Wiley, London, 2003.
- Bell, J. S., *Speakable and Unsayable in Quantum Mechanics*, Cambridge University Press, Cambridge, UK, 1987.

- Davies, P. C. W., and Brown, J. R. (eds), *The Ghost in the Atom: A Discussion of the Mysteries of Quantum Physics*, Cambridge University Press, Cambridge, UK, 1986.
- D'Espagnat, Bernard, *Conceptual Foundations of Quantum Mechanics*, 2nd edn, Addison-Wesley, New York, 1989, Part 3.
- Popper, Karl R., *Quantum Theory and the Schism in Physics*, Unwin Hyman, London, 1982.
- Raymer, Michael G., *Quantum Physics: What Everyone Needs to Know*, Oxford University Press, Oxford, 2017.
- Schilpp, P. A. (ed.), *Albert Einstein: Philosopher-Scientist*, The Library of Living Philosophers, Open Court, La Salle, IL, 1949, Chapter II-7.
- Susskind, Leonard, and Friedman, Art, *Quantum Mechanics: The Theoretical Minimum*, Penguin, London, 2015, Lectures 6 and 7.
- Wheeler, John Archibald, and Zurek, Wojciech Hubert (eds), *Quantum Theory and Measurement*, Princeton University Press, Princeton, NJ, 1983.
- Whitaker, Andrew, *John Stewart Bell and Twentieth-Century Physics: Vision and Integrity*, Oxford University Press, Oxford, 2016.

Epilogue

A Game of Theories

The Quantum Representation of Reality

I hope that, having made it to the end, you have enjoyed the journey. I set out to do three things. I wanted to show you that whilst quantum mechanics is mathematically challenging, some basic knowledge and a bit of effort will carry you a long way. I also wanted to show you how quantum mechanics was derived from careful studies of the *physics*. The rather abstract quantum formalism based on state vectors in Hilbert space was introduced only when it was deemed desirable to lend the theory greater mathematical consistency, and to reject some of its historical metaphysical baggage. I firmly believe that the best way to come to terms with this formalism is to first understand how and why it came about. As we've seen, it was only partially successful. Mathematical consistency might have been won, but philosophical clarity was certainly lost.

Debates about the interpretation of the quantum representation continue to this day. Many practicing scientists see such debates as rather pointless or futile. The theory works very well, so why trouble yourself about what it means? Why not just 'shut up and calculate'? This brings me to my third ambition. I had hoped that by presenting the historical development of the theory in this way, you might get the impression that any lack of comprehension of its meaning on your part is absolutely not your fault. The quantum-mechanical representation challenges our comprehension of what any (and all) scientific theories are meant to be telling us about the nature of physical reality. Honestly, it's okay to have doubts.

The question of interpretation remains so stubborn because *the choice we face is a philosophical one*. There is absolutely nothing scientifically wrong with an anti-realist interpretation in which there is no mystery and the theory is complete. We make use of what we know about a quantum system based on previous experiments and experience and we *code* this in the wavefunction, or state vector. All we do is use the theory to manipulate *information*, as a way of connecting past events to predictions of the future. If the wavefunction is just coded information, then it isn't physically real and there's no need to postulate a collapse. All that changes with the measurement is our knowledge, and 'This change is unproblematic,' writes Italian theorist Carlo Rovelli, 'for the same reason for which my information about China changes discontinuously any time I read an article about China in the newspaper.'¹

If, instead, we choose to pull on the loose thread, we find that we are obliged to take the representation at face value, and interpret its concepts more realistically. The theory then gives us all those things about the quantum world that we find utterly baffling and, just as Einstein himself concluded, we're obliged to accept that the theory can't be complete.

I thought I'd conclude by summarizing the efforts of the past 90 years, framed in the context of the Bohr–Einstein debate.²

Bohr was right: Quantum mechanics is complete and there are no problems. If you're discomfited by Copenhagen, take heart. There are alternatives, such as Rovelli's relational interpretation and so-called information-theoretic interpretations. Though unpalatable to many with different metaphysical prejudices, there's absolutely nothing wrong with this conclusion.

Bohr was (mostly) right: Quantum mechanics is complete but, to make better sense, the conventional representation needs to be interpreted in a different way. In one approach the axioms of the theory are reworked to make quantum mechanics an essentially probabilistic theory. In another, quantum probabilities are computed through the use of 'consistent histories'. In another, quantum probabilities are interpreted subjectively, using something called Bayesian decision theory.

Einstein was right (I): Quantum mechanics can be completed by including hidden variables, which sit at a deeper level 'beneath' the wavefunction and govern the physics. As we learned in Chapter 12, certain classes of local and crypto non-local hidden variables have now been largely ruled out through experimental tests. Other types of non-local hidden variable theories (including so-called 'pilot wave' interpretations) are hard if not impossible to distinguish experimentally from conventional quantum mechanics. But adopting a structure based on waves *and* particles means accepting the reality of 'spooky' action at a distance. The choice then becomes a matter of taste.

Einstein was right (II): Quantum mechanics can be completed by introducing a physical mechanism to account for the collapse of the wavefunction. Such mechanisms may invoke some kind of spontaneous collapse, or they may involve gravity (curved spacetime), or (following von Neumann and Wigner) the influence of a conscious mind.

Einstein was right (III): But in a way that I think would have appalled him. In a quantum measurement, *all* the possibilities are realized at once, but each in *different universes*. This is the 'many worlds' interpretation, the ultimate consequence of assuming the wavefunction is real.

The debate continues. After a thirty-year personal journey in search of the meaning of quantum mechanics, I can happily attest to the fact I still don't understand it. But I think I now understand why. And, having long favoured Einstein's philosophical position, I now confess to some doubts.

Like the great philosopher Han Solo, I've got a very bad feeling about this.

NOTES

1. Matteo Smerlak and Carlo Rovelli, 'Relational EPR', *Foundations of Physics*, 37 (2007), 427–45.
2. I've written about the various interpretations of quantum mechanics in a popular book titled *Quantum Reality: The Quest for the Real Meaning of Quantum Mechanics - A Game of Theories*. I regard this as a 'companion' to the present volume. Both are published by Oxford University Press.

Appendix 1

Cavity Modes

We assume that the radiation trapped inside the cavity consists of electromagnetic waves with frequency ν and wavelength λ . Many different frequencies will persist inside the cavity at equilibrium, provided they meet the condition for *standing waves*. Such standing waves are called ‘cavity modes’. The function $\rho(\nu, T)$ then represents the energy-density of these modes per unit volume per unit frequency interval.

The cavity modes derive their energy from the internal energy of the oscillators, $U(\nu, T)$, so the volume density of this is simply $U(\nu, T)/V$, where V is the volume of the cavity. The density per unit frequency interval is then simply the variation of the *number* of cavity modes, N_m , with frequency, or $dN_m/d\nu$. Hence,

$$\rho(\nu, T) = \frac{U(\nu, T)}{V} \left(\frac{dN_m}{d\nu} \right). \tag{A1.1}$$

We can sense-check this with a little dimensional analysis. We can deduce from Planck’s radiation law, Eq. (1.3), that the density function $\rho(\nu, T)$ has the units Jsm^{-3} . The energy (in J) derives from $U(\nu, T)$. Dividing this by the volume of the cavity gives us units of m^{-3} . N_m is simply a number, and $dN_m/d\nu$ is the rate of change of this number with frequency, which has units $(\text{s}^{-1})^{-1}$. So the dimensions of $\rho(\nu, T)$ in Eq. (A1.1) are $\text{J}(\text{s}^{-1})^{-1}\text{m}^{-3}$, or joules per unit frequency interval per cubic metre.

We begin by evaluating N_m . In one dimension (in the x -direction, say), if the length of the cavity is l then the standing wave condition means that only wavelengths which ‘fit’ between the end walls of the cavity will interfere constructively—see Fig. A1.1. In other words,

$$n_x \lambda = 2l \quad \text{or} \quad n_x \left(\frac{c}{\nu} \right) = 2l, \quad \text{so that} \quad n_x = \frac{2\nu l}{c}, \tag{A1.2}$$

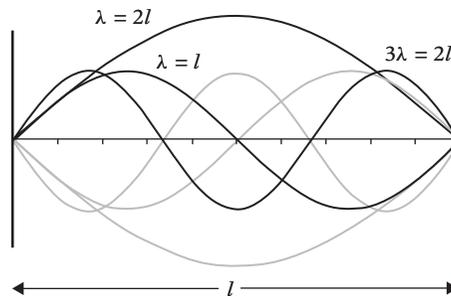


Figure A1.1 The standing wave condition.

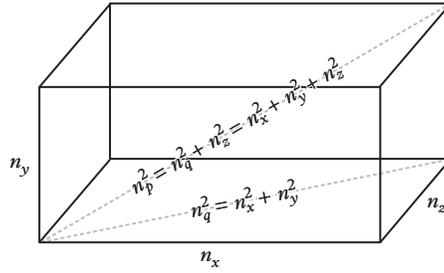


Figure A1.2 Pythagoras’ theorem can be used to generalize the expression for the number of modes to three dimensions.

where n_x is a positive integer. In practical terms, the values for n_x in Eq. (A1.2) will be very large. For example, for a cavity length of 20 centimetres, a standing light wave with a wavelength of 580 nm (yellow) will have (roughly) $n_x = 690,000$. The modes will also be very closely spaced, such that we can assume that they form a continuum along the ‘ n_x -coordinate’. If we assume for the moment that the cavity forms a cube, this allows us to make use of Pythagoras’ theorem to generalize (A1.2) to three dimensions (see Fig. A1.2),

$$n_p = \frac{2vl}{c}, \tag{A1.3}$$

where $n_p^2 = n_x^2 + n_y^2 + n_z^2$. We will soon discover that the result for $\rho(v, T)$ is actually independent of the cavity size and shape.

If the modes can be assumed to form a continuum, we can approximate N_m as the volume of a sphere of radius n_p . There are two caveats. We know that light has two forms of polarization—right circular and left circular (or, perhaps more familiarly, vertical and horizontal plane polarization). This means there are twice the number of modes. Secondly, if n_p is evaluated from the centre of the sphere, then we must exclude from consideration all negative values of n_x , n_y , and n_z , as these make no physical sense. This means considering only the positive octant of the sphere, or one-eighth of the volume (see Fig. A1.3). Hence,

$$N_m \cong 2 \left(\frac{1}{8} \right) \left(\frac{4}{3} \pi n_p^3 \right) \cong \frac{1}{3} \pi n_p^3. \tag{A1.4}$$

But we know from Eq. (A1.3) that $n_p = 2vl/c$, so

$$N_m \cong \frac{8\pi v^3 l^3}{3c^3}. \tag{A1.5}$$

Equation (A1.5) suggests that the number of modes inside the cavity will increase with increasing frequency. This makes perfect sense—the higher the frequency, the greater the number of nodes in the wave and so the greater the likelihood of meeting the standing wave condition. At high frequencies the modes will get increasingly closer together.

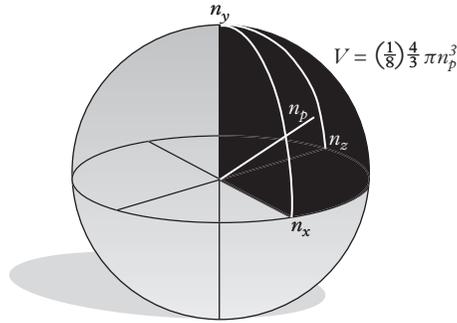


Figure A1.3 If we approximate the modes as a continuum, we can determine the total number as the positive octant of a sphere of volume with radius n_p .

The derivative of N_m with respect to frequency is then given by

$$\frac{dN_m}{dv} = \frac{8\pi v^2 l^3}{c^3}. \quad (\text{A1.6})$$

We can now insert this into Eq. (A1.1), remembering that for the moment we have assumed the cavity is a cube, so $V = l^3$,

$$\rho(v, T) = \frac{U(v, T)}{l^3} \frac{8\pi v^2 l^3}{c^3} = \frac{8\pi v^2}{c^3} U(v, T), \quad (\text{A1.7})$$

and we see that this result is indeed independent of cavity size and shape. This is the relation between radiation density and oscillator energy that Planck deduced in May 1899, Eq. (1.5).

Appendix 2

Lorentz Transformation for Energy and Linear Momentum

Consider three reference frames with the following relationships:

S_0 , the rest frame;

S , a moving frame with velocity u relative to S_0 , with coordinates t, x, y, z and velocity components u_x, u_y, u_z ; and

S' , a moving frame with velocity u' relative to S_0 , with coordinates t', x', y', z' and velocity components u'_x, u'_y, u'_z . S' is also moving with velocity v relative to S .

We begin by establishing relationships between the velocity components of u and u' and the velocity v . From the Lorentz transformation from S to S' , Table 2.1, we have

$$t' = \gamma_v \left(t - \frac{v}{c^2} x \right), \quad \text{where} \quad \gamma_v = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (\text{A2.1})$$

$$x' = \gamma_v (x - vt) \quad (\text{A2.2})$$

$$y' = y \quad (\text{A2.3})$$

$$z' = z. \quad (\text{A2.4})$$

For a small incremental change dx' we can use Eq. (A2.2) to deduce

$$dx' = \gamma_v (dx - v dt) \quad \text{or} \quad dx' = \gamma_v (u_x - v) dt, \quad (\text{A2.5})$$

where we have made use of the fact that $u_x = dx/dt$ and $dx = u_x dt$. Likewise,

$$dt' = \gamma_v \left(dt - \frac{v}{c^2} dx \right) \quad \text{or} \quad dt' = \gamma_v \left(1 - \frac{v}{c^2} u_x \right) dt. \quad (\text{A2.6})$$

Dividing (A2.5) by (A2.6) gives

$$\frac{dx'}{dt'} = u'_x = \frac{u_x - v}{1 - vu_x/c^2}. \quad (\text{A2.7})$$

Similarly,

$$dy' = dy = u_y dt \quad \text{and so} \quad \frac{dy'}{dt'} = u'_y = \frac{u_y}{\gamma_v(1 - vu_x/c^2)} \quad (\text{A2.8})$$

$$dz' = dz = u_z dt \quad \text{and so} \quad \frac{dz'}{dt'} = u'_z = \frac{u_z}{\gamma_v(1 - vu_x/c^2)}. \quad (\text{A2.9})$$

In what follows we will need to consider the Lorentz transformations from S_0 to S and from S_0 to S' , and it will therefore be helpful to deduce some relationships between the corresponding Lorentz factors. Let's start by considering $\gamma_{u'}$:

$$\frac{1}{\gamma_{u'}^2} = 1 - \frac{u'^2}{c^2} = 1 - \left(\frac{u_x'^2}{c^2} + \frac{u_y'^2}{c^2} + \frac{u_z'^2}{c^2} \right). \quad (\text{A2.10})$$

This can get quite complicated quite quickly, so let's consider each of the velocity components in turn. From (A2.7),

$$1 - \frac{u_x'^2}{c^2} = 1 - \frac{(u_x - v)^2/c^2}{(1 - vu_x/c^2)^2} = \frac{(1 - vu_x/c^2)^2 - (u_x - v)^2/c^2}{(1 - vu_x/c^2)^2}. \quad (\text{A2.11})$$

On expanding the brackets in the numerator we get, after some rearrangement,

$$1 - \frac{u_x'^2}{c^2} = \frac{(1 - v^2/c^2)(1 - u_x^2/c^2)}{(1 - vu_x/c^2)^2} = \frac{1 - u_x^2/c^2}{\gamma_v^2(1 - vu_x/c^2)^2}. \quad (\text{A2.12})$$

Similarly,

$$\frac{u_y'^2}{c^2} = \frac{u_y/c^2}{\gamma_v^2(1 - vu_x/c^2)^2} \quad (\text{A2.13})$$

$$\frac{u_z'^2}{c^2} = \frac{u_z/c^2}{\gamma_v^2(1 - vu_x/c^2)^2}. \quad (\text{A2.14})$$

Subtracting (A2.13) and (A2.14) from (A2.12) then gives

$$1 - \frac{u'^2}{c^2} = \frac{(1 - u^2/c^2)}{\gamma_v^2(1 - vu_x/c^2)^2} = \frac{1}{\gamma_v^2 \gamma_u^2 (1 - vu_x/c^2)^2}. \quad (\text{A2.15})$$

We can now invert this last expression and take the square root of both sides:

$$\gamma_{u'} = \gamma_v \gamma_u (1 - vu_x/c^2) \quad (\text{A2.16})$$

We now have everything we need. Recall from Eqs. (4.6) and (4.9) that transforming from S_0 to S and from S_0 to S' gives

$$E = \gamma_u m_0 c^2 \quad \text{and} \quad E' = \gamma_{u'} m_0 c^2 \quad (\text{A2.17})$$

$$p = \gamma_u m_0 u \quad \text{and} \quad p' = \gamma_{u'} m_0 u' \quad (\text{A2.18})$$

We can resolve (A2.17) into components as follows:

$$p_x = \gamma_u m_0 u_x \quad \text{and} \quad p'_x = \gamma_{u'} m_0 u'_x \quad (\text{A2.19})$$

$$p_y = \gamma_u m_0 u_y \quad \text{and} \quad p'_y = \gamma_{u'} m_0 u'_y \quad (\text{A2.20})$$

$$p_z = \gamma_u m_0 u_z \quad \text{and} \quad p'_z = \gamma_{u'} m_0 u'_z. \quad (\text{A2.21})$$

From (A2.17) and (A2.16) we have

$$E' = \gamma_v \gamma_u (1 - v u_x / c^2) m_0 c^2 = \gamma_v (\gamma_u m_0 c^2 - v \gamma_u m_0 u_x). \quad (\text{A2.22})$$

From this last expression, (A2.17), and (A2.19), it immediately follows that

$$E' = \gamma_v (E - v p_x) \quad (\text{A2.23})$$

Also, from (A2.19), (A2.16) and (A2.7) we have for p'_x

$$p'_x = \frac{\gamma_v \gamma_u (1 - v u_x / c^2) m_0 (u_x - v)}{(1 - v u_x / c^2)} = \gamma_v (\gamma_u m_0 u_x - \gamma_u m_0 v). \quad (\text{A2.24})$$

But we know that from (A2.19) $p_x = \gamma_u m_0 u_x$ and from (A2.17) $\gamma_u m_0 = E/c^2$, so

$$p'_x = \gamma_v \left(p_x - \frac{v}{c^2} E \right). \quad (\text{A2.25})$$

Similarly, from (A2.20), (A2.21) and (A2.8), (A2.9) we have

$$p'_y = \frac{\gamma_v \gamma_u (1 - v u_x / c^2) m_0 u_y}{\gamma_v (1 - v u_x / c^2)} = \gamma_u m_0 u_y = p_y \quad (\text{A2.26})$$

$$p'_z = \frac{\gamma_v \gamma_u (1 - v u_x / c^2) m_0 u_z}{\gamma_v (1 - v u_x / c^2)} = \gamma_u m_0 u_z = p_z. \quad (\text{A2.27})$$

These results are summarized on the left-hand side of Table 4.1.

Appendix 3

Energy Levels of the Hydrogen Atom

The radial component of the Schrödinger wave equation for the hydrogen atom is given in Eq. (5.37):

$$-\frac{\hbar^2}{2m_e} \frac{1}{r} \frac{\partial^2}{\partial r^2} r R_{nl}(r) + \left(\frac{-e^2}{4\pi \epsilon_0 r} + \frac{l(l+1)\hbar^2}{2m_e r^2} \right) R_{nl}(r) = E_n R_{nl}(r). \quad (\text{A3.1})$$

And from (5.38), the radial functions are

$$R_{nl}(r) = \sqrt{\left(\frac{2}{na_0} \right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}} (\rho_n)^l e^{-\rho_n/2} L_{n-l-1}^{2l+1}(\rho_n). \quad (\text{A3.2})$$

It's quite instructive to insert this expression for the radial function into the wave equation and solve for the energy, E_n . To simplify matters, I'm going to drop the subscript notation and the bracketed reference to the dependent variable throughout (though we mustn't forget that they're there) and replace the complicated square-root normalization factor in (A3.2) with the term N . With this notation, we can write Eq. (A3.2) as $R = N \rho^l e^{-\rho/2} L$. We'll also focus only on the kinetic energy term in (A3.1):

$$-\frac{\hbar^2}{2m_e} \frac{1}{r} \frac{\partial^2}{\partial r^2} r R = -\frac{\hbar^2}{2m_e} \frac{1}{r} \frac{\partial^2}{\partial r^2} r N \rho^l e^{-\rho/2} L. \quad (\text{A3.3})$$

As it stands, this looks pretty complicated—we need to differentiate the product of three functions that depend on r : ρ^l , $e^{-\rho/2}$, and L . But we can do a couple of things to make this a little simpler. We recall that $\rho = 2r/na_0$, so we can use the chain rule to turn this into a differential equation in ρ :

$$\frac{\partial}{\partial r} r R = \frac{\partial}{\partial \rho} \frac{d\rho}{dr} r R = \frac{\partial}{\partial \rho} \frac{2}{na_0} r R = \frac{\partial}{\partial \rho} \rho R \quad (\text{A3.4})$$

and

$$\frac{\partial^2}{\partial r^2} r R = \frac{\partial}{\partial r} \left(\frac{\partial}{\partial r} r R \right) = \frac{\partial}{\partial \rho} \frac{d\rho}{dr} \left(\frac{\partial}{\partial \rho} \rho R \right) = \frac{2}{na_0} \frac{\partial^2}{\partial \rho^2} \rho R. \quad (\text{A3.5})$$

Equation (A3.3) now becomes

$$-\frac{\hbar^2}{2m_e} \frac{1}{r} \frac{\partial^2}{\partial r^2} r R = -\frac{\hbar^2}{2m_e} \frac{2}{na_0 r} \frac{\partial^2}{\partial \rho^2} \rho R. \quad (\text{A3.6})$$

It's best to evaluate this expression in stages, and we begin by taking a single derivative with respect to ρ , noting as we do that ρ times ρ^l is ρ^{l+1} . Applying the product rule, we get*

$$\frac{\partial}{\partial \rho} \rho R = \frac{\partial}{\partial \rho} (N \rho^{l+1} e^{-\rho/2}) L = N \left[(l+1) \rho^l e^{-\rho/2} - \frac{\rho}{2} \rho^l e^{-\rho/2} \right] L + (N \rho^{l+1} e^{-\rho/2}) \frac{\partial L}{\partial \rho}. \quad (\text{A3.7})$$

We can see a little better what's happening here by gathering terms that are multiplied by $N \rho^l e^{-\rho/2} L$ and replacing them with R , giving

$$\frac{\partial}{\partial \rho} \rho R = (l+1)R - \frac{\rho}{2}R + (N \rho^{l+1} e^{-\rho/2}) \frac{\partial L}{\partial \rho}. \quad (\text{A3.8})$$

You might be getting a little nervous about that last term in (A3.8) containing the differential of the associated Laguerre polynomial, but please be reassured—I have a magic trick that will make this complexity go away.

Unfortunately, we need to differentiate (A3.8) again, and there are now three terms to consider rather than just one. Best to take them one term at a time, starting with the first:

$$\frac{\partial}{\partial \rho} [(l+1)R] = \frac{\partial}{\partial \rho} \left[(l+1)(N \rho^l e^{-\rho/2}) \right] L = \frac{l(l+1)}{\rho} R - \frac{(l+1)}{2} R + (l+1)(N \rho^l e^{-\rho/2}) \frac{\partial L}{\partial \rho}. \quad (\text{A3.9})$$

Likewise for the second term,

$$\frac{\partial}{\partial \rho} \left(\frac{\rho}{2} R \right) = \frac{\partial}{\partial \rho} \left[\frac{\rho}{2} (N \rho^l e^{-\rho/2}) \right] L = \frac{(l+1)}{2} R - \frac{\rho}{4} R + \frac{\rho}{2} (N \rho^l e^{-\rho/2}) \frac{\partial L}{\partial \rho}. \quad (\text{A3.10})$$

And, finally,

$$\frac{\partial}{\partial \rho} \left[(N \rho^{l+1} e^{-\rho/2}) \frac{\partial L}{\partial \rho} \right] = \left\{ \left[(l+1) - \frac{\rho}{2} \right] (N \rho^l e^{-\rho/2}) \right\} \frac{\partial L}{\partial \rho} + \rho (N \rho^l e^{-\rho/2}) \frac{\partial^2 L}{\partial \rho^2}. \quad (\text{A3.11})$$

We now subtract (A3.10) from (A3.9), add (A3.11), and gather terms to give

$$\frac{\partial^2}{\partial \rho^2} \rho R = \left[\frac{l(l+1)}{\rho} + \frac{\rho}{4} - (l+1) \right] R + [2(l+1) - \rho] (N \rho^l e^{-\rho/2}) \frac{\partial L}{\partial \rho} + \rho (N \rho^l e^{-\rho/2}) \frac{\partial^2 L}{\partial \rho^2}. \quad (\text{A3.12})$$

Here's the magic trick. The associated Laguerre polynomials are solutions of the *associated Laguerre equation* which, in the notation used here, can be written

$$\rho \frac{\partial^2 L}{\partial \rho^2} + [2(l+1) - \rho] \frac{\partial L}{\partial \rho} + (n-l-1)L = 0. \quad (\text{A3.13})$$

Of course, this isn't magic. Eq. (A3.13) is an inherent property of the associated Laguerre polynomials (and can quickly be tested by reference to the specific examples given in Table 5.2).

* $d(uvw)/dx = w(v du/dx + u dv/dx) + uv dw/dx$, where $u = N \rho^{l+1}$, $v = e^{-\rho/2}$, and $w = L$.

What is magical is that we can now replace the complicated derivative terms in (A3.12):

$$\rho \frac{\partial^2 L}{\partial \rho^2} + [2(l+1) - \rho] \frac{\partial L}{\partial \rho} = -(n-l-1)L. \quad (\text{A3.14})$$

And, of course, L , when multiplied by $N\rho^l e^{-\rho/2}$, returns the radial function R once more. Equation (A3.12) then becomes

$$\frac{\partial^2}{\partial \rho^2} \rho R = \left[\frac{l(l+1)}{\rho} + \frac{\rho}{4} - (l+1) \right] R - (n-l-1)R. \quad (\text{A3.15})$$

Or

$$\frac{\partial^2}{\partial \rho^2} \rho R = \left[\frac{l(l+1)}{\rho} + \frac{\rho}{4} - n \right] R. \quad (\text{A3.16})$$

We can now put this result into (A3.6):

$$-\frac{\hbar^2}{2m_e} \frac{1}{r} \frac{\partial^2}{\partial r^2} rR = -\frac{\hbar^2}{2m_e} \frac{2}{na_0 r} \left[\frac{l(l+1)}{\rho} + \frac{\rho}{4} - n \right] R. \quad (\text{A3.17})$$

We now replace ρ with $2r/na_0$ and multiply the terms inside the bracket by the factor $2/na_0 r$, giving

$$-\frac{\hbar^2}{2m_e} \frac{1}{r} \frac{\partial^2}{\partial r^2} rR = -\frac{\hbar^2}{2m_e} \left[\frac{l(l+1)}{r^2} + \frac{1}{n^2 a_0^2} - \frac{2}{a_0 r} \right] R, \quad (\text{A3.18})$$

which is Eq. (5.45) in Chapter 5. As we saw there, when inserted into the wave Eq. (A3.1), we can cancel the radial function R . We then find that the term $\hbar^2/m_e a_0 r$ exactly cancels the contribution from the Coulomb potential, $-e^2/4\pi\epsilon_0 r$, and the term $-l(l+1)\hbar^2/2m_e r^2$ exactly cancels the positive contribution from the angular momentum in the effective potential. We're therefore just left with

$$-\frac{\hbar^2}{2m_e n^2 a_0^2} = -\frac{m_e e^4}{8\epsilon_0^2 \hbar^2} \cdot \frac{1}{n^2} = E_n, \quad (\text{A3.19})$$

in agreement with Bohr's original formula.

Appendix 4

Orthogonality of the Hydrogen Atom Wavefunctions

From Eqs. (5.41) and (5.42) we have

$$\psi_{100} = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} \quad (\text{A4.1})$$

$$\psi_{200} = \frac{1}{4\sqrt{2\pi a_0^3}} \left(2 - \frac{r}{a_0}\right) e^{-r/2a_0}, \quad (\text{A4.2})$$

where we have substituted for ρ_n . Both of these solutions are spherically symmetric, so once again we can write

$$\int \psi_{200} \psi_{100} d\tau = \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi \int_0^\infty \psi_{200} \psi_{100} dr. \quad (\text{A4.3})$$

The integrals over the angles evaluate to 4π , and so

$$\int \psi_{200} \psi_{100} d\tau = \frac{1}{\sqrt{2a_0^3}} \int_0^\infty r^2 \left(2 - \frac{r}{a_0}\right) e^{-3r/2a_0} dr. \quad (\text{A4.4})$$

This looks a little daunting, but we can rearrange it as follows:

$$\int \psi_{200} \psi_{100} d\tau = \frac{2}{\sqrt{2a_0^3}} \int_0^\infty \left(r^2 - \frac{r^3}{2a_0}\right) e^{-3r/2a_0} dr. \quad (\text{A4.5})$$

Or

$$\int \psi_{200} \psi_{100} d\tau = \frac{2}{\sqrt{2a_0^3}} \left(\int_0^\infty r^2 e^{-3r/2a_0} dr - \int_0^\infty \frac{r^3}{2a_0} e^{-3r/2a_0} dr \right). \quad (\text{A4.6})$$

The last integral in Eq. (A4.6) is of the general form of another standard integral:

$$\int x^n e^{bx} dx = \frac{x^n e^{bx}}{b} - \frac{n}{b} \int x^{n-1} e^{bx} dx. \quad (\text{A4.7})$$

Substituting $x = r$ and $b = -3/2a_0$ gives

$$\int_0^{\infty} \frac{r^3}{2a_0} e^{-3r/2a_0} dr = \left[-\frac{r^3 e^{-3r/2a_0}}{3} + \int_0^{\infty} r^2 e^{-3r/2a_0} dr \right]_0^{\infty}. \quad (\text{A4.8})$$

We can now see what happens. The integral in (A4.8) cancels the first integral in (A4.6), and we're left with

$$\int \psi_{200} \psi_{100} d\tau = \frac{2}{3\sqrt{2}a_0^3} \left[r^3 e^{-3r/2a_0} \right]_0^{\infty} = 0. \quad (\text{A4.9})$$

As before, when $r = \infty$ the exponential term declines to 0. And when $r = 0$, $r^3 = 0$, so the integral evaluates as 0.

Appendix 5

The Integral Cauchy-Schwarz Inequality

Consider an arbitrary complex function Φ and its complex conjugate Φ^* , for which

$$\int \Phi^* \Phi d\tau \geq 0. \quad (\text{A5.1})$$

We define Φ such that the equality holds only when $\Phi = 0$.

Suppose we express Φ as a linear combination of two functions f and g , such that

$$\Phi = f + \lambda g \quad \text{and} \quad \Phi^* = f^* + \lambda^* g^*, \quad (\text{A5.2})$$

where λ is a complex coefficient. If we substitute for Φ and Φ^* in (A5.1), we get

$$\int \Phi^* \Phi d\tau = \int (f^* + \lambda^* g^*)(f + \lambda g) d\tau \geq 0. \quad (\text{A5.3})$$

Or

$$\int (f^* f + \lambda f^* g + \lambda^* g^* f + \lambda^* \lambda g^* g) d\tau \geq 0 \quad (\text{A5.4})$$

$$\int |f|^2 d\tau + \lambda \int (f^* g) d\tau + \lambda^* \int (g^* f) d\tau + \lambda^* \lambda \int |g|^2 d\tau \geq 0.$$

We now multiply each term in (A5.4) by $\int |g|^2 d\tau$, giving

$$\int |f|^2 d\tau \cdot \int |g|^2 d\tau + \lambda \left(\int (f^* g) d\tau \cdot \int |g|^2 d\tau \right) + \lambda^* \left(\int (g^* f) d\tau \cdot \int |g|^2 d\tau \right) + \lambda^* \lambda \left(\int |g|^2 d\tau \right)^2 \geq 0. \quad (\text{A5.5})$$

Now the inequality (A5.5) must hold no matter how we might choose to define the complex coefficient, λ . So, we choose to define it as follows:

$$\lambda = -\frac{\int (g^* f) d\tau}{\int |g|^2 d\tau} \quad \text{and} \quad \lambda^* = -\frac{\int (f^* g) d\tau}{\int |g|^2 d\tau}. \quad (\text{A5.6})$$

Inserting these expressions for λ and λ^* into (A5.5) give:

$$\int |f|^2 d\tau \cdot \int |g|^2 d\tau - \int (f^* g) d\tau \cdot \int (g^* f) d\tau \geq 0. \quad (\text{A5.7})$$

Or

$$\int |f|^2 d\tau \cdot \int |g|^2 d\tau \geq \left| \int (f^*g) d\tau \right|^2, \quad (\text{A5.8})$$

which is the integral form of the Cauchy–Schwarz inequality.

This inequality appears in many forms, and one of the simplest derivations is available from simple vector algebra. If f and g are now considered as classical vectors with a common origin and with an angle, θ , between them, then the inner product of these (or ‘dot’ product or scalar product if you prefer) is given by

$$f \cdot g = |f| |g| \cos \theta. \quad (\text{A5.9})$$

It follows that

$$|f \cdot g|^2 = |f|^2 |g|^2 \cos^2 \theta. \quad (\text{A5.10})$$

Or

$$|f|^2 |g|^2 = \frac{|f \cdot g|^2}{\cos^2 \theta}. \quad (\text{A5.11})$$

But $\cos^2 \theta$ has a maximum value of 1 and so, in general,

$$|f|^2 |g|^2 \geq |f \cdot g|^2, \quad (\text{A5.12})$$

and the equality holds for values of θ equal to 0 and integral multiples of π .

Appendix 6

Orbital Angular Momentum in Quantum Mechanics

We start, as we do so often in quantum mechanics, with a reminder of the classical treatment. Picture an object of mass m rotating in a plane around a fixed point. We express the angular momentum L (or the ‘moment of momentum’) of the object as a vector quantity determined from the cross product of the radial vector and the instantaneous linear momentum vector (see Fig. A6.1):

$$L = r \times p. \quad (\text{A6.1})$$

We decompose both r and p into components as follows,

$$\begin{aligned} r &= xi + yj + zk \\ p &= p_x i + p_y j + p_z k, \end{aligned} \quad (\text{A6.2})$$

where i, j, k are unit vectors along the x, y, z axes.

Thus, we have

$$\begin{aligned} L &= (xi + yj + zk) \times (p_x i + p_y j + p_z k) \\ &= xp_x(i \times i) + xp_y(i \times j) + xp_z(i \times k) \\ &\quad + yp_x(j \times i) + yp_y(j \times j) + yp_z(j \times k) \\ &\quad + zp_x(k \times i) + zp_y(k \times j) + zp_z(k \times k). \end{aligned} \quad (\text{A6.3})$$

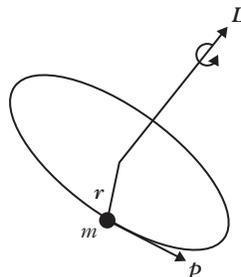


Figure A6.1 The angular momentum vector L of an object of mass m rotating around a fixed point is given by $L = r \times p$.

Now the scalar magnitude of the cross product of two vectors \mathbf{a} and \mathbf{b} is given by $|\mathbf{a}||\mathbf{b}| \sin \theta$, where θ is the angle between the vectors drawn from the same origin. Hence,

$$\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = 0, \quad (\text{A6.4})$$

since in these cases $\theta = 0$ and so $\sin \theta = 0$. It also follows that if we reverse the direction of vector \mathbf{b} (for example), then the magnitude of $\mathbf{a} \times \mathbf{b}$ is now given by $|\mathbf{a}||\mathbf{b}| \sin(\theta + 2\pi) = -|\mathbf{a}||\mathbf{b}| \sin \theta = -\mathbf{b} \times \mathbf{a}$. Finally, the cross product of any two unit vectors (for which $\theta = \pi$, $\sin \theta = 1$) must be equal to the third unit vector. Hence,

$$\begin{aligned} \mathbf{i} \times \mathbf{j} &= -\mathbf{j} \times \mathbf{i} = \mathbf{k} \\ \mathbf{j} \times \mathbf{k} &= -\mathbf{k} \times \mathbf{j} = \mathbf{i} \\ \mathbf{k} \times \mathbf{i} &= -\mathbf{i} \times \mathbf{k} = \mathbf{j}. \end{aligned} \quad (\text{A6.5})$$

This allows us to re-write Eq. (A6.3) as

$$\begin{aligned} \mathbf{L} &= xp_y \mathbf{k} - xp_z \mathbf{j} - yp_x \mathbf{k} + yp_z \mathbf{i} + zp_x \mathbf{j} - zp_y \mathbf{i} \\ &= (yp_z - zp_y) \mathbf{i} + (zp_x - xp_z) \mathbf{j} + (xp_y - yp_x) \mathbf{k}. \end{aligned} \quad (\text{A6.6})$$

In other words, the components L_x, L_y, L_z of the angular momentum vector are given by

$$L_x = yp_z - zp_y \quad L_y = zp_x - xp_z \quad L_z = xp_y - yp_x \quad (\text{A6.7})$$

and $L^2 = L_x^2 + L_y^2 + L_z^2$, where L is the magnitude of the angular momentum vector \mathbf{L} .

We can now turn these expressions for the components of the classical angular momentum into their quantum mechanical equivalents simply by replacing the classical linear momentum p_x with the operator $\hat{p}_x = -i\hbar \partial / \partial x$, and likewise for p_y and p_z :

$$\hat{L}_x = y\hat{p}_z - z\hat{p}_y \quad \hat{L}_y = z\hat{p}_x - x\hat{p}_z \quad \hat{L}_z = x\hat{p}_y - y\hat{p}_x. \quad (\text{A6.8})$$

From this we can swiftly deduce commutation relations as follows:

$$\begin{aligned} [\hat{L}_x, \hat{L}_y] &= \hat{L}_x \hat{L}_y - \hat{L}_y \hat{L}_x \\ &= (y\hat{p}_z - z\hat{p}_y)(z\hat{p}_x - x\hat{p}_z) - (z\hat{p}_x - x\hat{p}_z)(y\hat{p}_z - z\hat{p}_y) \\ &= [y\hat{p}_z, z\hat{p}_x] - [y\hat{p}_z, x\hat{p}_z] - [z\hat{p}_y, z\hat{p}_x] + [z\hat{p}_y, x\hat{p}_z]. \end{aligned} \quad (\text{A6.9})$$

We can simplify this last expression for $[\hat{L}_x, \hat{L}_y]$ by noting that

$$\begin{aligned} [y\hat{p}_z, z\hat{p}_x] &= y\hat{p}_z z\hat{p}_x - z\hat{p}_x y\hat{p}_z = y\hat{p}_x \hat{p}_z z - y\hat{p}_x z \hat{p}_z = y\hat{p}_x [\hat{p}_z, z] \\ [y\hat{p}_z, x\hat{p}_z] &= y\hat{p}_z x\hat{p}_z - x\hat{p}_z y\hat{p}_z = yx\hat{p}_z^2 - xy\hat{p}_z^2 = 0 \\ [z\hat{p}_y, z\hat{p}_x] &= z\hat{p}_y z\hat{p}_x - z\hat{p}_x z\hat{p}_y = z^2 \hat{p}_y \hat{p}_x - z^2 \hat{p}_x \hat{p}_y = 0 \\ [z\hat{p}_y, x\hat{p}_z] &= z\hat{p}_y x\hat{p}_z - x\hat{p}_z z\hat{p}_y = x\hat{p}_y z \hat{p}_z - x\hat{p}_y \hat{p}_z z = x\hat{p}_y [z, \hat{p}_z]. \end{aligned} \quad (\text{A6.10})$$

This gives

$$[\hat{L}_x, \hat{L}_y] = y\hat{p}_x [p_z, z] + p_y x [z, p_z]. \quad (\text{A6.11})$$

But we know that $[z, p_z] = i\hbar$, and $[p_z, z] = -i\hbar$, so

$$[\hat{L}_x, \hat{L}_y] = i\hbar (x\hat{p}_y - y\hat{p}_x) = i\hbar \hat{L}_z. \quad (\text{A6.12})$$

We can obviously repeat this process for $[\hat{L}_y, \hat{L}_z]$ and $[\hat{L}_z, \hat{L}_x]$, or we can acknowledge that all we need to do is cyclically permute the indices:

$$[\hat{L}_x, \hat{L}_y] = i\hbar \hat{L}_z \quad [\hat{L}_y, \hat{L}_z] = i\hbar \hat{L}_x \quad [\hat{L}_z, \hat{L}_x] = i\hbar \hat{L}_y. \quad (\text{A6.13})$$

We'll go on to consider the consequences of these commutation relations later in this Appendix.

Given the operator \hat{L} , what then are the eigenvalues of angular momentum? The standard textbook answer involves the introduction of so-called 'shift operators' and an examination of their properties, which can be done entirely without reference to the precise nature of the wavefunctions. However, it turns out that we actually already have everything we need for a 'quick-and-dirty' answer to this question.

Again, we reach back to the treatment of angular momentum in classical mechanics. From Eq. (A6.1), the scalar magnitude of the angular momentum L is given by $|r||p| \sin \theta = rp = rmv$ (the vectors \mathbf{r} and \mathbf{p} form a right angle, for which $\sin \theta = 1$, m is the mass of the rotating object, and v is the orbital or angular velocity, measured—for example—in radians per second).

For the particular case of an orbiting electron with charge $-e$ and mass m_e , an electric current I is generated which is determined by the rate at which the charge passes a given point in the orbit, according to $I = -e/\tau$, where $\tau = 2\pi r/v$ is the orbital *period*. This current gives rise to a *magnetic dipole moment*, of magnitude m , given by IA , where $A = \pi r^2$ is the area enclosed by the orbit. Hence,

$$m = IA = -\frac{ev}{2\pi r} \cdot \pi r^2 = -\frac{evr}{2}. \quad (\text{A6.14})$$

But for an orbiting electron, $L = rm_e v$, or $rv = L/m_e$ and so

$$m = -\frac{e}{2m_e} L = \gamma_e L, \quad (\text{A6.15})$$

where $\gamma_e = -e/2m_e$ is called the *gyromagnetic ratio*.

In a situation where the classical electron moves in a circular orbit, acted on by a conservative central force, the total energy E is conserved according to

$$\frac{1}{2} m_e v^2 + V(r) = E, \quad (\text{A6.16})$$

where $V(r)$ is the potential energy. But once again we know that $L = rm_e v$, so we can substitute for v and re-write Eq. (A6.16) in terms of L :

$$\frac{L^2}{2m_e r^2} + V(r) = E. \quad (\text{A6.17})$$

We've seen something like this before. The term $L^2/2m_e r^2$ is often referred to as a 'centrifugal potential energy'. An equivalent term appears in the Schrödinger equation for the electron in a hydrogen atom, Eq. (5.37), which we explained as a kind of centrifugal force which tends to push the electron outwards, resisting the attraction from the positively charged nucleus.

So, we can simply compare (A6.17) and (5.37),

$$\frac{L^2}{2m_e r^2} = \frac{l(l+1)\hbar^2}{2m_e r^2}, \quad (\text{A6.18})$$

in which l is the orbital angular momentum quantum number. This allows us to conclude that the orbital angular momentum is quantized with allowed values of L given by

$$L = \sqrt{l(l+1)}\hbar. \quad (\text{A6.19})$$

The corresponding eigenfunctions are of course the spherical harmonics, $Y_l^{m_l}(\theta, \phi)$, which we encountered in Chapter 5. From Eqs. (A6.18) and (5.33) we can deduce that

$$\hat{L}^2 = -\hbar^2 \Lambda^2 = -\hbar^2 \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right). \quad (\text{A6.20})$$

I don't propose to prove it here, but the component operators can be similarly expressed in spherical polar co-ordinates as follows:

$$\begin{aligned} \hat{L}_x &= i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right) \\ \hat{L}_y &= -i\hbar \left(\cos \phi \frac{\partial}{\partial \theta} - \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right) \\ \hat{L}_z &= -i\hbar \frac{\partial}{\partial \phi}. \end{aligned} \quad (\text{A6.21})$$

Let's take a quick look at the action of \hat{L}_z on the functions $Y_l^{m_l}(\theta, \phi)$. From Eq. (5.34) we see that we can factor these functions into separate θ - and ϕ -dependent parts,

$$Y_l^{m_l}(\theta, \phi) = N_l^{m_l} \Theta_l^{m_l}(\theta) \Phi_l^{m_l}(\phi), \quad (\text{A6.22})$$

where

$$\begin{aligned}
 N_l^{m_l} &= \sqrt{\frac{(2l+1)(l-|m_l|)!}{2(l+|m_l|)!}} \\
 \Theta_l^{m_l}(\theta) &= P_l^{|m_l|}(\cos\theta) \quad \text{and} \\
 \Phi^{m_l}(\phi) &= \frac{1}{\sqrt{2\pi}} e^{im_l\phi}.
 \end{aligned}
 \tag{A6.23}$$

So, we can deduce that

$$\begin{aligned}
 \hat{L}_z Y_l^{m_l}(\theta, \phi) &= -i\hbar N_l^{m_l} \Theta_l^{m_l}(\theta) \frac{\partial}{\partial \phi} \left(\frac{1}{\sqrt{2\pi}} e^{im_l\phi} \right) \\
 &= -i\hbar N_l^{m_l} \Theta_l^{m_l}(\theta) im_l \left(\frac{1}{\sqrt{2\pi}} e^{im_l\phi} \right) \\
 &= m_l \hbar Y_l^{m_l}(\theta, \phi).
 \end{aligned}
 \tag{A6.24}$$

And we see that the eigenvalues of the operator for the z -component of the angular momentum, L_z , are given by $m_l \hbar$.

There's one last thing to note. The commutator $[\hat{L}_x, \hat{L}_y] = i\hbar \hat{L}_z$ implies that whilst we can determine the magnitude of the z -component of the orbital angular momentum (A6.24) and the overall magnitude of the angular momentum L (A6.19), the x - and y -components are uncertain. In fact, from (A6.13) we see that we can specify *any* component, but only at the cost of uncertainty

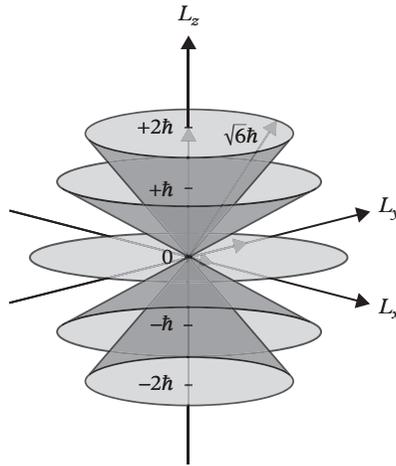


Figure A6.2 If we specify the z -component of the orbital angular momentum then the x - and y -components are uncertain. This means that the angular momentum vector may lie anywhere on the surface of a cone of length $\sqrt{l(l+1)}\hbar$. This is illustrated here for $l = 2$.

in the other two. From Eq. (7.43) we have

$$\sigma_{L_x}^2 \sigma_{L_y}^2 \geq \left| \frac{1}{2} [\hat{L}_x, \hat{L}_y] \right|^2 = \left| \frac{1}{2} i\hbar \langle \hat{L}_z \rangle \right|^2. \quad (\text{A6.25})$$

Or

$$\sigma_{L_x} \sigma_{L_y} \geq \frac{1}{2} \hbar |\langle \hat{L}_z \rangle|, \quad (\text{A6.26})$$

where $\langle \hat{L}_z \rangle$ is the expectation value of \hat{L}_z .

We conclude from this that the magnitude of the angular momentum of an electron orbiting a central nucleus is quantized according to $L = \sqrt{l(l+1)}\hbar$, and that the magnitude of the z -component is quantized according to $L_z = m_l\hbar$. Because the components L_x and L_y are uncertain, we represent the angular momentum vector as a *cone*, orientated along the z -axis, of length $\sqrt{l(l+1)}\hbar$. This is illustrated in Fig. A6.2 for the specific example $l = 2$, for which the length is $\sqrt{6}\hbar$ and there are five— $(2l+1)$ —possible values for m_l ranging from $+2$, $+1$, 0 , -1 , and -2 .

Note that, from (A6.15), the magnetic dipole moment in the z -direction, m_z , is quantized and given by

$$m_z = \gamma_e L_z = \gamma_e m_l \hbar. \quad (\text{A6.27})$$

Appendix 7

A Very Brief Introduction to Matrices

A matrix is a regular array of numbers, symbols, or expressions organized into rows and columns. It provides a convenient way of manipulating a large number of individual *matrix elements* simultaneously. Although Heisenberg developed a version of quantum mechanics based on matrices (of infinite dimensions)—called matrix mechanics (see Chapter 6)—matrices appear more broadly in quantum mechanics and are not confined to Heisenberg’s formulation.

A matrix with m rows and n columns may form a single column ($n = 1$), a square array ($m = n$), or a rectangular array ($m \neq n$). For example, the square 2×2 matrix M has two rows and two columns,

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}, \quad (\text{A7.1})$$

where the numbers (or symbols, or expressions) m_{11} , etc., are the matrix elements.

The matrix M is said to be *diagonal* if the only non-zero elements lie on the leading diagonal, from top left to bottom right, i.e. if $m_{12} = m_{21} = 0$ in Eq. (A7.1). If this is the case and if $m_{11} = m_{22} = 1$, the result is the *unit matrix* or *identity matrix*, often given the symbol 1.

The *trace* of a matrix (symbol Tr) is the sum of the leading diagonal elements, from top left to bottom right:

$$\text{Tr}(M) = \text{Tr} \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} = m_{11} + m_{22}. \quad (\text{A7.2})$$

The *transpose* of M , denoted M^T , is formed by swapping or transposing the off-diagonal elements:

$$M^T = \begin{pmatrix} m_{11} & m_{21} \\ m_{12} & m_{22} \end{pmatrix}. \quad (\text{A7.3})$$

In the event that one or more of the matrix elements of M are complex (i.e. they contain i), then there is some interest in the complex-conjugate transpose, denoted M^\dagger (and sometimes referred to as the ‘dagger’ of the matrix):

$$M^\dagger = \left(M^T\right)^* = \begin{pmatrix} m_{11}^* & m_{21}^* \\ m_{12}^* & m_{22}^* \end{pmatrix}. \quad (\text{A7.4})$$

The *determinant* of M is given by

$$\det M = |M| = \begin{vmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{vmatrix} = m_{11}m_{22} - m_{12}m_{21}. \quad (\text{A7.5})$$

The *sum* of M and another 2×2 matrix N is given by

$$\begin{aligned} M + N &= \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} + \begin{pmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{pmatrix} \\ &= \begin{pmatrix} m_{11} + n_{11} & m_{12} + n_{12} \\ m_{21} + n_{21} & m_{22} + n_{22} \end{pmatrix}. \end{aligned} \quad (\text{A7.6})$$

The *product* of M and N is given by

$$\begin{aligned} MN &= \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{pmatrix} \\ &= \begin{pmatrix} m_{11}n_{11} + m_{12}n_{21} & m_{11}n_{12} + m_{12}n_{22} \\ m_{21}n_{11} + m_{22}n_{21} & m_{21}n_{12} + m_{22}n_{22} \end{pmatrix}, \end{aligned} \quad (\text{A7.7})$$

and we can see immediately that M and N do not necessarily commute, since

$$\begin{aligned} NM &= \begin{pmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{pmatrix} \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \\ &= \begin{pmatrix} n_{11}m_{11} + n_{12}m_{21} & n_{11}m_{12} + n_{12}m_{22} \\ n_{21}m_{11} + n_{22}m_{21} & n_{21}m_{12} + n_{22}m_{22} \end{pmatrix}. \end{aligned} \quad (\text{A7.8})$$

For example, suppose

$$M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}. \quad (\text{A7.9})$$

Then

$$\begin{aligned} MN &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix} \quad \text{and} \\ NM &= \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}. \end{aligned} \quad (\text{A7.10})$$

Let's finish by applying the rules of matrix multiplication to the Pauli spin matrices given in Eq. (9.19):

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (\text{A7.11})$$

Thus,

$$\sigma_x \sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = i\sigma_z. \quad (\text{A7.12})$$

Similarly,

$$\sigma_y \sigma_z = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = i \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = i \sigma_x \quad (\text{A7.13})$$

and

$$\sigma_z \sigma_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = i \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = i \sigma_y. \quad (\text{A7.14})$$

Appendix 8

A Simple Local Hidden Variables Theory

I propose to continue with the example in which two entangled photons are produced in opposite states of circular polarization. In a local hidden variables extension of quantum mechanics, we assume the existence of a physical mechanism which predetermines the *linear* polarization states of photons 1 and 2. This mechanism relies on a certain variable or variables—to be defined—and, because these are not revealed in the experiments (nor required by the standard quantum-mechanical formalism), they are necessarily ‘hidden’. The presumption is that in its most widely accepted formalism, quantum mechanics is not complete.

So let’s suppose that hidden within each circularly polarized photon there exists a vector which is *fixed by the physics* and which dictates how the photon will interact with a polarization analyser and (most importantly) through which linear polarization channel it will pass. Let’s further imagine that this vector can point in any direction within a circle orthogonal to the direction of travel. We presume that if the vector points in any direction within $\pm 45^\circ$ of the (arbitrary) laboratory vertical axis, it will pass through the vertical channel. Likewise, if the vector points in any direction within $\pm 45^\circ$ of the horizontal axis, it will pass through the horizontal channel.

To conserve angular momentum, we suppose that in each of the entangled photons these vectors are opposed. For example, if the vector for photon 1 points ‘north’ and falls within $\pm 45^\circ$ of the vertical axis, then the vector for photon 2 points ‘south’, but of course this will also fall within $\pm 45^\circ$ of the vertical axis (see Fig. A8.1). We can see immediately that with this arrangement of the analysers, we will only ever get the measurement outcomes vertical–vertical and horizontal–horizontal, entirely consistent with Eqs. (12.9) and (12.10).

Moreover, we can determine the probabilities P_{vv} and P_{hh} simply as the ratio of the areas of the two segments (each subtended by an angle $90^\circ = \frac{1}{2}\pi$ radians) to the total area of the circle,

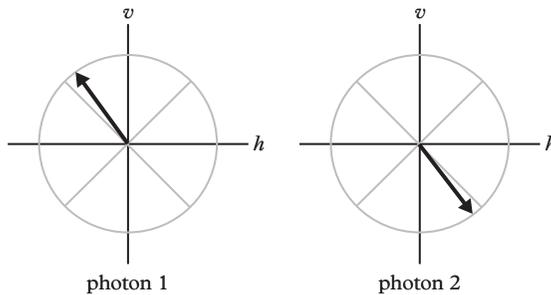


Figure A8.1 In this simple hidden variable extension, each circularly polarized photon possesses a hidden vector which predetermines its linear polarization state. In this example, the two photons both produce vertical polarization outcomes.

$$P_{vv} = P_{hh} = \frac{2 \times \frac{1}{2} \left(\frac{\pi}{2}\right) r^2}{\pi r^2} = \frac{1}{2}, \tag{A8.1}$$

which is also entirely consistent with the quantum-mechanical predictions. It's worth emphasizing once again that in this scenario we assume that linear polarization measurements performed on photon 1 in no way affect the outcome of any simultaneous or subsequent measurement of the linear polarization state of photon 2, and vice versa.

The key question is: What happens in this scheme when we rotate one (or both) of the polarization analysers? Remember that the hidden vectors are supposed to be fixed by the physics and predetermined. Their orientation in space cannot therefore be affected by the orientations of the analysers. Again, we don't need elaborate mathematics to figure out what happens. As we rotate one of the analysers clockwise (by an angle φ measured from the vertical), for a given orientation of the vectors we *reduce* the probability of a vertical-vertical result and *increase* (from 0) the probability of a vertical-horizontal result (see Fig. A8.2):

$$P_{vv}(\varphi) = \frac{1}{2} - \frac{2 \times \frac{1}{2} \varphi r^2}{\pi r^2} = \frac{1}{2} - \frac{\varphi}{\pi} \tag{A8.2}$$

and

$$P_{vh}(\varphi) = \frac{\varphi}{\pi}. \tag{A8.3}$$

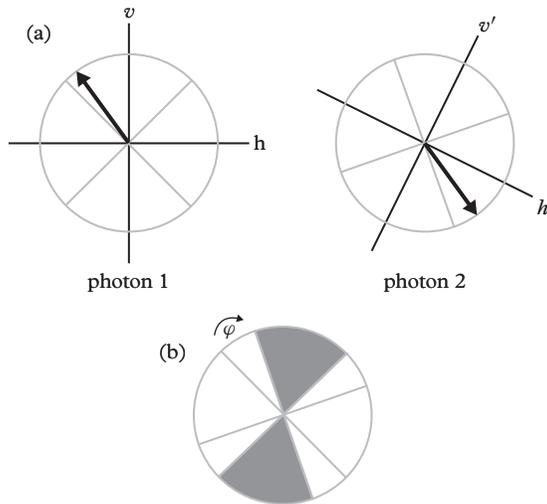


Figure A8.2 As the relative orientation of the polarization analysers is rotated as shown in (a), the area corresponding to vertical-vertical measurement outcomes—the shaded area in (b)—is reduced by an amount φ/π . The same vectors in Figure A8.1 now produce a vertical-horizontal outcome.

Similarly,

$$P_{hv}(\varphi) = \frac{\varphi}{\pi}. \quad (\text{A8.4})$$

And

$$P_{hh}(\varphi) = \frac{1}{2} - \frac{\varphi}{\pi}. \quad (\text{A8.5})$$

It is straightforward to show that this simple hidden variables model conforms to the demands of Bell's inequality. From Eq. (12.23) we have

$$P_{vh}(\varphi_1) + P_{vh}(\varphi_2) \geq P_{vh}(\varphi_3), \quad (\text{A8.6})$$

Which, for the particular combination of angles $\varphi_1 = 22\frac{1}{2}^\circ (\pi/8)$, $\varphi_2 = 22\frac{1}{2}^\circ (\pi/8)$, and $\varphi_3 = 45^\circ (\pi/4)$, gives

$$\frac{1}{8} + \frac{1}{8} = \frac{1}{4} \geq \frac{1}{4}. \quad (\text{A8.7})$$

From (12.16), the correlation function $C_{12}(\varphi)$ is

$$C_{12}(\varphi) = P_{vv}(\varphi) - P_{vh}(\varphi) - P_{hv}(\varphi) + P_{hh}(\varphi), \quad (\text{A8.8})$$

and from Eqs. (A8.2)–(A8.5) we see that this declines linearly with increasing φ :

$$\begin{aligned} C_{12}(\varphi) &= \frac{1}{2} - \frac{\varphi}{\pi} - \frac{\varphi}{\pi} - \frac{\varphi}{\pi} + \frac{1}{2} - \frac{\varphi}{\pi} \\ &= 1 - \frac{4\varphi}{\pi}. \end{aligned} \quad (\text{A8.9})$$

We see from this that the hidden variables prediction for $C_{12}(\varphi)$ is consistent with the quantum-mechanical prediction $\cos 2\varphi$ only for the difference angles $\varphi = 0$, $\varphi = \pi/4$ (45°), and $\varphi = \pi/2$ (90°), corresponding to $C_{12}(\varphi) = +1, 0$, and -1 . The predictions are inconsistent for all other angles, with the greatest disagreement occurring for $\varphi = \pi/8$ ($22\frac{1}{2}^\circ$) and $\varphi = 3\pi/8$ ($67\frac{1}{2}^\circ$), for which the hidden variables theory predicts $C_{12}(\varphi) = \frac{1}{2}$ and $C_{12}(\varphi) = -\frac{1}{2}$, compared with the quantum-mechanical predictions $1/\sqrt{2}$ and $-1/\sqrt{2}$ (see Fig. 12.3).

Index

n after a page number indicates a footnote or an endnote. *f* after a page number indicates a figure.

A

Aharonov, Yakir 249
Anderson, Carl 199
angular momentum 60, 282–7;
 see also orbital angular
 momentum, spin angular
 momentum
anomalous Zeeman effect 90,
 158, 164
anschaulichkeit 112
anticommutators 147
antimatter/antiparticles 199
Aspect, Alain 257, 260
atoms 14–15
axioms 207–8

B

Balmer, Johann Jakob 58
Balmer formula 58–9, 65, 67
Balmer series 59, 67
Bell, John 231, 249, 250, 252
Bell's inequality 249, 255–61,
 293
Bernoulli, Daniel 14
Bertlmann, Reinhard 252*n*
'Bertlmann's Socks and the
 Nature of Reality' (John
 Bell) 252, 253*f*
birefringent crystals 225
black body radiation 20–1
Bloch, Felix 92
Bohm, David 248–9, 262*n*
Bohr, Niels 57–8, 59, 61, 70,
 133, 152–4, 163–4, 181,
 203, 208, 220, 222, 223,
 247, 266
Bohr radius 65
Bohr's trilogy ('On the
 Constitution of Atoms and
 Molecules') 61, 69
Bohr–Einstein debate 222, 266
Bohr–Sommerfeld model 90,
 107, 111
Boltzmann, Ludwig 14, 15,
 25, 28

Boltzmann's constant 21
Born, Max 87, 115, 116, 120,
 126–8, 129–30, 203, 221
Born rule 214, 220, 221
Bose–Einstein
 condensation 177
bosons 177
bra vectors 210, 211
Brackett series 67
Brougham, Henry 15*n*

C

calcite crystals 225
Carlsberg Foundation 57
Carnot, Sadi 12
Cartan, Élie 84
Cauchy–Schwarz
 inequality 144, 280–1
cavity modes 269–71
cavity radiation 20, 24, 33
centrifugal potential
 energy 285
circularly polarized light 224*f*,
 225–7, 230*n*, 244, 270,
 291–3
Clapeyron, Émile 12
classical mechanics 3–10
classical wave equation 10,
 92–5
Clauser, John 257
Clausius, Rudolf 12, 14
Clebsch–Gordan
 coefficients 172
Clebsch–Gordan series 158
cloud chamber 134*n*
'collapse of the
 wavefunction' 237, 266
commutation relations /
 commutators 135–6, 137,
 144–6, 148, 149
complementarity 152–4, 219,
 220
complementary variables
 136

complex conjugate
 transpose 140
composite states 234
composite systems 168–9
Compton, Arthur 73
Compton effect 152, 153
Compton scattering 73
configuration space 6, 112
constant of the motion 149
Copenhagen interpretation
 ('Kopenhagener Geist')
 220, 222, 223, 239,
 243, 249
correlation functions 245, 293
Coulomb's law 62
covariance 147

D

Dalibard, Jean 257
Darwin, Charles (grandson of
 the 'real' Darwin) 57
Davisson, Clinton 85
de Broglie, Louis 35, 73–4,
 79–80, 81, 82, 84, 85,
 86–7, 203
de Broglie, Maurice 74
Debye, Pieter 73, 92, 107
determinant (matrices) 288
differential scattering
 cross-section 119, 124–6
Dirac, Paul vii, 116, 164, 179,
 181, 187, 189*n*, 190, 191*n*,
 197, 198, 199, 203, 204,
 206, 207, 208–9, 211,
 235, 247
Doppler effect 44
Dulong, Pierre 55
Dyson, Freeman vii, 200

E

Ehrenfest, Paul 163
Ehrenfest's theorem 150
eigenfunctions 98
eigenstates 211

- eigenvalue equations 98
 Einstein, Albert 32–3, 36, 41–2, 45, 46, 50, 51, 55, 56, 85, 87, 127, 208, 220, 221–3, 238–9, 243, 246–8, 249, 266
 Einstein separability 247, 256
 elastic scattering 118, 121
 electric dipole
 transitions 170–1
 electrodynamics 35–6, 41–5
 electromagnetic waves 10
 electron shells 159
 electron spin 157, 164–5, 171–2, 179, 181, 192–4, 203
 electron waves 120, 121*f*
 electrons 20, 70*n*, 176, 200
 electrostatic units 201*n*
 energy
 classical mechanics 4–5
 relativistic
 electrodynamics 186–7
 special relativity 44, 45–9, 50–2
 energy–time uncertainty relation 148–51, 223
 entangled particles 246
 entropy 12–13, 14, 15, 28–31
 EPR (Einstein, Podolsky, Rosen) argument 246–7, 248
 equipartition theorem 24
 ETH (Eidgenössische Technische Hochschule) 92*n*
 ether (hypothetical space-filling matter) 11
 Euler formulation 4
 expansion theorem 137–8, 228, 232
 expectation values 139
- F**
 Faraday, Michael 10
 fermions 176–7
 Feynman, Richard viii, 200, 262*n*
 first law of thermodynamics 23
 FitzGerald, George 35, 40
 fluxes 118, 125
 Foley, H. M. 200
 force (classical mechanics) 3–4
 Fowler, Alfred 67
- frames of reference 37–8, 40–1, 42–4, 45–9
 Freedman, Stuart 257
- G**
 gamma rays 152
 Geiger, Hans 56
 general theory of relativity 148
 Gerlach, Walther 159
 Germer, Lester 85
Gespensterfeld 127
 gluons 177
 Gordon, Walter 100, 179
 Goudsmit, Samuel 163
 Grangier, Philippe 257
 gravity (classical mechanics) 9
 ‘guiding field’ (*Führungsfeld*) 247
 gyromagnetic ratio 197, 284
- H**
 Habicht, Conrad 50
 Hamilton, William Rowan 5–6
 Hamiltonian mechanics 6–7, 14
 Hamiltonian operator 97–8, 216, 217
 Hankel functions 122–4
 Hansen, Hans Marius 58
 Heisenberg, Werner 115, 116, 117, 133–5, 136–7, 151–2, 153*n*, 162, 163–4, 165, 176, 179, 181, 198, 203, 219, 220, 221
 helium atom 158, 166–7, 173–6
 Hermann, Jakob 4
 Hermitian conjugate 211
 Hermitian operators 139–40, 204, 212
 Hertzprung–Russell diagram 165*n*
 hidden variables theories 248, 259–61, 266, 291–3
 Hilbert, David 205, 206, 207
 Hilbert spaces 206, 209–10
 Hilbert’s problems 207
 Hiley, Basil 262–3*n*
 Hiroshima 51
 Hooke, Robert 9
 Huygens, Christiaan 9
 hydrogen atom 100–107, 275–7, 278–9
- I**
 inner product (Hilbert spaces) 210
 ‘inner quantum number’ 158
 International Congress of Mathematicians (Paris 1900) 207
 isotropic scattering 119
- J**
 Jammer, Max 237
 Jeans, James 21
 Jordan, Pascual 116
 Joule, James 12
- K**
 Kelvin (William Thomson) 1, 12, 15*n*
 Kepler’s third law of planetary motion 63
 ket vectors 209, 210
 kinetic energy 4, 5, 48–9, 50, 62, 63
 kinetic theory of gases 14
 Kirchhoff, Gustav 20–1
 Klein, Oskar 100, 154, 179, 181
 Klein–Gordon equation 100, 179, 180–1, 188–91
 Kochen, Simon 256
 Kohlrausch, Rudolf 10
 Kossel, Walther 159
 Kronecker delta 210
 Kronig, Ralph de Laer 162–3, 164
 Kudar, Johann 181
 Kuhn, Thomas 15
 Kurlbaum, Ferdinand 22
 Kusch, Polykarp 200
- L**
 ‘la Comédie Française’ 85
 Lagrange, Joseph-Louis 5
 Laguerre polynomials 104, 276
 Lamb, Willis 200
 Lamb shift 200
 Landau, Basil 52
 Landé, Alfred 159
 Landé *g*-factor 162, 197, 200
 Langevin, Paul 84, 85
 Larmor, Joseph 40
 Legendre polynomials 102
 Leggett, Anthony 261
 Leggett’s inequality 261

- Leibniz, Gottfried 2, 4
 Lewis, Gilbert 33, 52
 light waves (classical mechanics) 9–11
 light-quantum hypothesis 32–3, 45
 linear Hertzian oscillators 20
 Liouville, Joseph 8*n*
 Lorentz, Hendrik 35, 40, 134, 163
 Lorentz factor 35, 46, 47, 48, 49–50
 Lorentz transformation 40–1, 43, 51, 77, 78–9, 82–4, 272–4
 Lorentz–FitzGerald contraction 35
 luminosity 119
 Lyman series 67
- M**
- Mach, Ernst 8, 15
 magnetic dipole moment 197, 284
 magnetic quantum number 90
 Malus's law 226
 Manchester Memorandum (Niels Bohr) 58
 Mandelstam, Leonid 151
 'many worlds' interpretation 266
 Marsden, Ernest 56
 mass
 classical mechanics 8
 special relativity 50, 51–2
 mass defect 51
Mathematical Foundations of Quantum Mechanics (John von Neumann) 206, 209, 231, 232, 248, 262*n*
Mathematical Principles of Natural Philosophy (Isaac Newton) 2, 3, 4*n*, 8, 9
 matrices 288–90
 matrix mechanics 115–17, 133, 205
 matter waves 84–6, 203
 Mauguin, Charles 84, 85
 Maxwell, James Clerk 10–11, 14
 Maxwell's Demon 237
 Maxwell's equations 10, 42, 186
 Maxwell–Boltzmann distribution 14
 measurement; *see* quantum measurement
 measurement eigenstates 234, 236*n*
 measurement eigenvalues 235
 measurement operator 232, 235
 Mermin, N. David 262*n*
 Michelson, Albert 1, 11, 89
 Michelson–Morley experiment 11, 35, 36, 43
 Millikan, Robert 33, 73
 momentum
 classical mechanics 3, 6
 relativistic electrodynamics 186–7
 special relativity 74–7
 momentum operator 98
 Moore, Walter 99
 Morley, Edward 11, 89
 multi-electron atoms 158
- N**
- Nafe, John 200
 Nelson, Edward 200
 Nernst, Walther 56
 neutrons 52*n*, 176
New York Times 247
 Newton, Isaac 2, 3, 4*n*, 8, 9
 Newton's first law of motion 150
 Newton's second law of motion 3, 4*n*, 7, 150
 Nicholson, John 69
 no-go theorems 256
 noble gases 159
 non-local hidden variable theories 261
 normal Zeeman effect 158
 normalization (wavefunctions) 113–14
 normalized vectors 210
 nuclear fission 51
- O**
- observable quantities 142
 Omnès, Roland 207
 'On the Electrodynamics of Moving Bodies' (Albert Einstein) 41–5
Opticks (Isaac Newton) 9
 orbital angular momentum 60, 68–9, 89, 102, 107, 177, 193, 194, 197, 285, 286*f*
 orbital angular momentum quantum number 90, 158, 165
 orbital approximation 167
 Orion nebula 67*f*
 orthogonal states 244
 orthogonal/orthonormal vectors 210
 orthogonal/orthonormal wavefunctions 114–15, 278–9
 orthohelium 175
 oscillators 20, 23–31
 Ostwald, Wilhelm 15
- P**
- Pais, Abraham 25–6, 45, 68, 89–90
 parahelium 175
 partial wave amplitudes 124–6
 Paschen, Friedrich 21, 67
 Pauli, Wolfgang vii–viii, 116, 117, 128, 159, 161–2, 163, 164, 181, 203, 220, 247
 Pauli exclusion principle 161–2, 166, 176, 244
 Pauli spin matrices 183–5, 190, 191–2, 195, 289–90
 periodic table 159, 160*f*, 161*n*
 Perrin, Jean 84, 85
 Petit, Alexis 55
 Pfund series 67
 phase convention 228–9
 phase space 6, 7, 206
 phase waves 80
 photons 33, 74, 76, 82, 177, 223–4
 Pickering, Edward Charles 67
 Pierce, Charles Sanders 240*n*
 pilot waves 81
 Planck, Max 13, 15, 16, 19–20, 22, 23–5, 28, 30, 32, 33, 55, 203, 271
 Planck's constant 21
 Planck's radiation law 22–3, 25, 26, 30, 31, 64
 Planck–Einstein relation 33, 44, 45, 64, 73, 99, 203
 Podolsky, Boris 243, 246
 Poincaré, Henri 7–8, 40
 Poincaré recurrence 8, 15
 polarized light 224–7, 270
 Popper, Karl 243

- position–momentum
uncertainty relation 146–7
- positivism 240*n*
- positrons 199
- potential energy 4, 5, 7, 8, 62,
63, 97
- pragmatism 240*n*
- principal quantum number 90,
91
- principle of relativity 36,
37–40, 45
- Principles of Quantum Mechanics*
(Paul Dirac) 187, 204,
208–9, 235
- Pringsheim, Ernst 21
- probability densities 127–8
- product (matrices) 289
- projection amplitudes 213,
214, 228–30
- projection operators 213
- projection postulate 236
- projection probabilities 214,
225–7, 228, 236
- protons 52*n*, 60*n*, 70*n*, 176
- proton–proton chain 51
- Q**
- quanta 24–5, 32
- ‘Quantization as an Eigenvalue
Problem’ (Erwin
Schrödinger) 98
- quantum chromodynamics 52*n*
- quantum electrodynamics 200
- quantum indistinguishability
169–70
- quantum jumps 66
- quantum measurement 231–7
- ‘Quantum Mechanics of
Collision Phenomena’
(Max Born) 126
- quantum numbers 59, 66,
90–1, 158, 161
- quantum probabilities 127,
205, 206, 214, 238
- Quantum Theory* (David
Bohm) 248, 249
- quantum theory of gravity 148
- quarks 52*n*, 176
- R**
- Rabi, Isidor 200
- radial distribution
function 128, 129*f*
- radial wave equation 120–2,
275
- radial wavefunctions 105*f*, 106
- Rayleigh (William Strutt)
21, 24
- Rayleigh–Jeans law 21, 22,
23, 24
- relativistic wave equation 185,
188–97, 194
- relativity; *see* general theory of
relativity, principle of
relativity, special theory of
relativity
- Retherford, Robert 200
- Ritz, Walther 59
- Robertson, Howard 137
- Roger, Gérard 257
- Rosen, Nathan 243, 246
- Rosenfeld, Léon 197
- Rovelli, Carlo 265
- Rubens, Heinrich 22, 23
- Russell–Saunders coupling 165
- Rutherford, Ernest 56, 57,
69–70, 199
- Rydberg, Johannes 58
- Rydberg constant 59, 66, 67
- Rydberg formula 59–68, 116,
203
- S**
- Sampanthar, Sam 52
- scattering (particles) 117–19;
see also Compton
scattering
- scattering amplitude 120, 124
- scattering potential 121
- Schrödinger, Erwin 91–2, 98,
107, 111–13, 115, 117,
133, 203, 219, 221, 239,
246, 247
- Schrödinger uncertainty
relation 147
- Schrödinger wave equation 95,
96–107, 111, 149, 216
- Schrödinger’s cat 239
- Schwinger, Julian 200
- second law of
thermodynamics 12–15,
28
- self-rotation 162–6
- ‘shut up and calculate’
slogan 248, 262*n*
- simultaneity 37–8
- singlet states 171, 174, 175
- Smolin, Lee ix
- Solo, Han 266
- Solvay, Ernest 92, 181
- Solvay conference (1927) 181,
204, 220
- Solvay conference (1930) 222
- Sommerfeld, Arnold 89, 153*n*,
158
- special theory of relativity 35,
41–52, 73, 74–7, 179,
180, 203
- Specker, Ernst 256
- speed of light (*c*) 51–2
- spherical harmonics 102
- spherical polar coordinates 100
- spin angular momentum 194,
196, 197
- spin angular momentum
operators 183, 184
- spin angular momentum
quantum number 165,
171, 172
- spin quantum number 165,
171
- spin wavefunctions 172, 176,
181–2
- ‘spooky action at a
distance’ 221, 246
- standard deviation 142
- standing wave condition
269
- Stark effect 89
- state preparation 233
- state vectors 205, 215
- statistical methods 15–16,
25–6, 28–31, 127
- Stern, Otto 159, 163, 222
- Stoner, Edmund 159, 161
- Strutt, William; *see* Rayleigh
- sum (matrices) 289
- superposition
(wavefunctions) 137
- superposition states 212–14
- Szilard, Leo 237
- T**
- Tamm, Igor 151
- Taylor series 49
- theorem of phase harmony 80
- thermodynamics 12–14
- Thomas, Llewellyn Hilleth
164
- Thomson, George P. 85
- Thomson, Joseph John
(J. J.) 56, 85
- Thomson, William; *see* Kelvin
- thought experiments
(*gedankenexperiments*)

quantum theory 222, 231
 relativity theory 39–40, 45–7
 time evolution operator 216–17
 time-dependent Schrödinger
 wave equation 99–100,
 216
 Tolman, Richard 52
 Tomonaga, Sin-Itiro 200
 trace (matrices) 288
 transpose (matrices) 288
 triplet states 172, 175
 two-slit apparatus 85*f*

U

Uhlenbeck, George 163, 164
 ultraviolet catastrophe 21
 uncertainty principle 136, 153,
 212
 uncertainty relations 146–51
 unit vectors 215
 unitary operators 216

V

variance 141–4
 vector spaces 206
 ‘Vertauschungs’ relations 187,
 193
 vertical/horizontal
 polarization 225, 226*f*
 virial theorem 63
 von Hevesy, George 57
 von Neumann, John vii, 205,
 206, 207, 209, 231, 232,
 236–7, 248, 262*n*

W

W particle 177
 wave equation; *see* classical
 wave equation, relativistic
 wave equation,
 Schrödinger wave equation
 wave mechanics 92, 96,
 111–12, 120, 133, 205

wave packets 111, 134
 wave–particle duality 74, 81–6,
 96
 wavefunctions 10, 95,
 112–115, 137–8, 204
 Weber, Wilhem 10
 Weyl, Herman 103
 Wien, Wilhelm 21, 103, 153*n*
 Wien’s law 21, 22
 Wigner, Eugene 237
 Wigner’s friend 238
 Wilson, Charles 134

Y

Young, Thomas 9–10, 15*n*

Z

Z particle 177
 Zeeman effect; *see* anomalous
 Zeeman effect
 Zermelo, Ernst 15