

1

Introduction

1.1 What this book is about

Quantum mechanics is an extraordinarily successful theory. The quantum mechanical description of the structures and spectra of atoms and molecules is virtually complete, and in principle, this provides the basis for understanding all of chemistry. Quantum mechanics gives detailed insight into many thermal, electrical, magnetic, optical, and elastic properties of condensed materials, including superconductivity, superfluidity, and Bose-Einstein condensation. Quantum mechanics underlies the theory of nuclear structure, nuclear reactions, and radioactive decay. Quantum electrodynamics (QED), an outgrowth of quantum mechanics and special relativity, is a very successful and detailed description of the interaction of charged leptons (i.e., electrons, muons, and tau leptons) with the electromagnetic radiation field. More generally, relativistic quantum field theory, the extension of quantum mechanics to relativistic fields, is the basis for all successful theoretical attempts so far to describe the phenomena of elementary particle physics.

We assume that you, the reader, have some elementary knowledge of quantum mechanics and that you know something about the historical development of the subject and its main principles and methods. We take advantage of this background, after a brief mathematical review in Chapter 2, by stating the rules of quantum mechanics in Chapter 3. An advantage of this approach is that all the rules are set forth in one place so that we can focus on them. In Chapter 3 we also describe application of the rules to several real physical situations, most significantly experiments with photon polarizations. Following some development of wave mechanics (Chapter 4), we illustrate the rules with additional examples (Chapter 5). We then develop the theory further in subsequent chapters, giving as many examples as we can from the physical world.

Our choice of topics is determined to a large extent by diverse student needs. Some students plan a career in theoretical physics, but most will work in experimental physics or will use quantum mechanics in some other branch of science or technology. Many will never take a subsequent course in elementary particle physics or quantum field theory. Yet most students want to know, and should know, something about the most interesting and important modern developments in quantum physics, even if time or preparation does not permit going into full detail about many topics. Thus, in addition to standard material, which can be found in a large number of existing textbooks, we include discussions of Bell’s inequality and photon polarization correlations, neutron interferometry, the Aharonov-Bohm effect, neutrino oscillations, the path integral method, second quantization for fermions, the stability of matter, quantization of the electromagnetic radiation field, the Casimir-Polder effect, the Lamb shift, the adiabatic theorem and geometric phases, relativistic wave equations and especially the Dirac equation,

the Dirac field, elementary QED, and a lengthy chapter on quantum mechanics of weak interactions, including an introduction to the electroweak standard model. The choice of topics is also influenced by my background and experience: I was trained as an experimentalist and have spent my entire research career in experimental physics.

The rules of quantum mechanics are remarkably successful in accounting for all experimental results to which they have been applied. However, because of the unique way in which probabilistic concepts appear, particularly in one rule (the so-called collapse postulate), controversy about the foundations of quantum mechanics has existed from the very beginning, and it continues today [see, e.g., Laloe (2012)]. Indeed, if we insist that quantum mechanics should apply not only to a microscopic system such as an electron or an atom but also to the macroscopic apparatus employed to measure that system's properties and the environment that is coupled to the apparatus, the collapse postulate is in conflict with another essential rule that describes how an isolated quantum mechanical system evolves continuously in time. This thorny issue is called the *quantum measurement problem*, and it has troubled many thoughtful persons, including two of the great founders of quantum theory, Albert Einstein and Erwin Schrodinger, and in more recent times the distinguished physicists John S. Bell and Stephen L. Adler, among many others. A summary of the quantum measurement problem and of several attempts to resolve it is given in Chapter 25.

Before we start, let us remark briefly on notation and units. Throughout this book, when the symbol e refers to electric charge it means the *magnitude* of the electronic charge, a positive quantity. The actual charge of the electron is $-e$. If we refer to a nonspecific electric charge that might or might not be e or $-e$, we use the symbol q .

It is not practical for us to work with a single system of units. Instead, we try to employ units that are most appropriate for the topic at hand. Initially, this may seem confusing and discouraging to the student, but it is a fact of life that a practicing physicist must learn to be conversant with several different unit systems. For the most part, we use Heaviside-Lorentz units (hlu system) for general discussions of nonrelativistic quantum mechanics. The hlu and cgs systems are the same, except that if a given electric charge has numerical value q_{cgs} in the cgs system, it has the value $q_{\text{hlu}} = \sqrt{4\pi}q_{\text{cgs}}$ in the hlu system, and similarly for currents, magnetic moments, electric dipole moments, and other electromagnetic sources. On the other hand, if a given electric field has numerical value $\mathcal{E}_{\text{cgs}}$ in the cgs system, it has numerical value $\mathcal{E}_{\text{hlu}} = \mathcal{E}_{\text{cgs}}/\sqrt{4\pi}$ in the hlu system, and similarly for magnetic fields and scalar and vector potentials. We employ atomic units (defined in Section 8.5) for atomic and molecular physics and natural units for relativistic quantum mechanics and field theory, with hlu conventions for electric and magnetic sources and fields. (This natural unit system is defined in Section 15.7 and is used extensively in Chapters 19–24). Although Système International (SI) units are familiar to many students and are convenient for practical engineering and technology, they are awkward and inconvenient for quantum mechanics, especially for relativistic quantum mechanics, so we avoid them.

1.2 A very brief summary of the antecedents of quantum mechanics

Although the invention of quantum mechanics occurred in the remarkably short time interval from 1925 through 1928, this burst of creativity was the culmination of a

twenty-five-year gestation period (1900–1925). During that era, the failures of classical physics to account for a wide range of important physical phenomena were revealed, and the need for radical new explanations of these phenomena became increasingly evident. In the following paragraphs we briefly summarize some of the most important achievements of the period from 1900 through 1925. [For a detailed history, see Jammer (1966)]. Here and in the rest of this book we encourage the reader to pay attention to the interplay between experiment and theory that has been so essential for the invention and development of quantum mechanics.

The question of how to account theoretically for the frequency spectrum of black-body radiation had been discussed in the last decades of the nineteenth century, but it gained urgency by 1900 because of accurate measurements of the spectrum by a number of experimentalists, notably H. Rubens and F. Kurlbaum. In that era, the energy per unit volume per hertz of black-body radiation at frequency ν in a cavity at absolute temperature T was predicted by the classical Rayleigh-Jeans formula to be

$$u_\nu = \frac{8\pi\nu^2 k_B T}{c^3} \quad (\text{Rayleigh – Jeans formula}) \quad (1.1)$$

where k_B is Boltzmann’s constant, and c is the velocity of light. This formula not only disagreed with the observations of Rubens and Kurlbaum, but when integrated over all frequencies, it led to the nonsensical conclusion that the total energy of radiation in a cavity of any finite volume at any finite temperature is infinite. Max Planck (1900) introduced the quantum of action h in late 1900 to obtain a new formula¹ for u_ν :

$$u_\nu = \frac{8\pi h \nu^3}{c^3} \frac{1}{\exp(h\nu/k_B T) - 1} \quad (\text{Planck’s law}) \quad (1.2)$$

Planck’s law agrees with experiment, and in the limit where $k_B T / \nu \gg h$, it reduces to the Rayleigh-Jeans formula. Planck later called his great achievement an act of desperation, and for some years after 1900, he struggled without success to find an explanation for the existence of h within the laws of classical physics.

Albert Einstein recognized the significance of Planck’s law more deeply than Planck himself. Einstein was thus motivated to suggest a corpuscular description of electromagnetic radiation (Einstein 1905). He proposed that the corpuscles (later called *photons*) have energy $E = h\nu$, where ν is the radiation frequency, and he employed this idea in his theory of the photoelectric effect. Convincing experimental evidence was obtained in support of this theory by a number of investigators, most notably Robert Millikan, in the decade following 1905 (Millikan 1916). Nevertheless, many physicists found it difficult to reconcile the idea of discrete photons with the highly successful and universally accepted wave theory of classical electromagnetism. Thus Einstein’s corpuscular description gained adherents only very slowly. However, in 1923, Arthur H. Compton made careful observations of x-ray–electron scattering, and he gave a successful kinematic description of this scattering (now called the *Compton effect*) by considering the relativistic collision of a photon with an electron, where both are regarded as particles

¹ The presently accepted value of h , now called *Planck’s constant*, is $h = 6.62606957(29) \times 10^{-27}$ erg·s.

(Compton 1923). Compton showed that a photon not only carries energy $E = h\nu$ but also linear momentum; that is,

$$p = \frac{h\nu}{c} \quad (1.3)$$

His results finally convinced the community of physicists to accept wave-particle duality for electromagnetic radiation. What we mean by this duality is that electromagnetic radiation has wavelike properties or particlelike properties depending on what sort of observation is made.

The specific heats of solids presented a problem somewhat related to that of the black-body spectrum. In 1819, these specific heats were predicted classically by DuLong and Petit to be a constant independent of temperature. However, by the end of the nineteenth century, it became clear that while measured specific heats agree with the DuLong-Petit law at relatively high temperatures, they tend toward zero as $T \rightarrow 0$. This behavior was explained by Einstein (1911) and in more detail by Peter Debye (1912) as well as by Max Born and Theodore von Karman (Born and von Karman 1912, 1913). Their theory, which invoked quantization of lattice vibrations of solids, was a natural outgrowth of the early quantum theory of black-body radiation.

Ernest Rutherford used the results of alpha-particle scattering experiments to propose the nuclear atom model (Rutherford 1911). Needless to say, an atom in this model consists of a massive and very compact nucleus about which atomic electrons circulate in orbital motion. According to classical physics, such electrons should radiate electromagnetic waves because of their centripetal acceleration, and a simple classical estimate shows that they should lose energy and spiral into the nucleus in times of order 10^{-15} s. However, atoms are stable, so it is obvious that the classical description is very wrong. In 1913, Niels Bohr recognized this, as well as the fact that no combination of fundamental constants in classical physics can yield a natural length scale for an atom, whereas $4\pi\hbar^2/m_e e^2 \approx 10^{-8}$ cm does provide such a scale. (Here m_e is the electron mass, and e is the magnitude of electron charge in the hlu system.) Employing the concepts of quantized stationary (nonradiating) orbits and radiative transitions between them, where h plays a crucial role, Bohr constructed his model of atomic hydrogen (Bohr 1913). He thereby successfully accounted for the frequencies of optical transitions in atomic hydrogen and in singly ionized helium. His model quickly gained wide acceptance in part because of convincing supportive evidence from the experiments of James Franck and Gustav Herz (Franck and Herz 1914). Here electrons from a thermionic source were accelerated in an evacuated tube containing a low density of atomic vapor (e.g., sodium, potassium, thallium, mercury, etc.). If the electron kinetic energy was sufficiently low, only elastic collisions between electrons and atoms occurred. However, if the electron kinetic energy was high enough to excite a transition from the ground state of an atom to an excited state, the electron suffered an inelastic collision with corresponding energy loss, and fluorescence was observed as the excited atom decayed back to the ground state.

Bohr's model was elaborated by Arnold Sommerfeld (1916), who derived a formula for the fine-structure splittings in hydrogen and singly ionized helium by applying quantization conditions to classical Keplerian orbits of the electron and by including an important relativistic correction. Sommerfeld's formula agreed (albeit fortuitously) with spectroscopic observations of the fine structure and thus the Bohr-Sommerfeld model was taken seriously for about a decade as a plausible way to understand atomic structure.

The fund of experimental data concerning atomic spectra grew very rapidly in the first decades of the twentieth century, thanks to the efforts of many optical and x-ray spectroscopists. Attention naturally was drawn to the problem of assigning Bohr-Sommerfeld quantum numbers to hundreds of newly observed energy levels in scores of atoms. Of special interest were the quantum numbers of atoms in their ground states because this was obviously related to the role of atomic structure in building up the periodic table. Here Edmund C. Stoner made a valuable contribution in October 1924 by publishing an authoritative classification of such quantum numbers (Stoner 1924). Stoner's conclusions came to the attention of Wolfgang Pauli, who used them to formulate the extremely important exclusion principle at the end of 1924 (Pauli 1925).

Observations and analyses of the Zeeman effect played an especially significant role in the elucidation of atomic energy level quantum numbers. Following Peter Zeeman's pioneering measurements of the splitting of sodium spectral lines in a magnetic field (Zeeman 1897), Henrik A. Lorentz gave what appeared to be a correct theoretical explanation based on classical electrodynamics in the same year (Lorentz 1897). This was called the *normal* Zeeman effect. However, as more observations with higher resolution were carried out on many spectral lines in various atoms, it became apparent that the normal Zeeman effect is the exception rather than the rule. Instead, the *anomalous* Zeeman effect is typical, in which more complicated patterns of level splittings occur. For years, the anomalous Zeeman effect remained a mystery because all efforts to explain it failed. Finally, the puzzle was resolved with invention of the concept of electron spin by George Uhlenbeck and Samuel Goudsmit in November 1925 (Uhlenbeck and Goudsmit 1925, 1926). Earlier in 1925, Ralph Kronig had conceived of the same idea, but he was discouraged by adverse criticism and withdrew his proposal. [For a brief history of electron spin, see Commins (2012).]

Next we turn to the phenomenon of wave-particle duality for *material particles* (i.e., electrons, protons, atoms, etc.). First, let us recall relation (1.3) between momentum and frequency established by Compton for the photon. Employing the familiar expression $\lambda = c/\nu$ relating wavelength and frequency, we see that (1.3) implies

$$\lambda = \frac{h}{p} \quad (1.4)$$

In 1923, Louis de Broglie (1923, 1924) made the extremely important suggestion that each material particle is associated with a wave such that if the momentum of the particle is p , the wavelength of the corresponding “matter” wave is also given by (1.4). Fragmentary experimental evidence supporting de Broglie's hypothesis was already available in 1921 from results obtained by C. Davisson and C. H. Kunsman on the scattering of electrons from a nickel surface (Davisson 1921). By 1927, Davisson and L. Germer (1927) and, independently, G. Thomson and A. Reid (Thomson and Reid 1927; Thomson 1928) provided convincing evidence for relation (1.4) from electron diffraction experiments. Since then, the validity of (1.4) for material particles has been demonstrated precisely in many different experiments using a variety of material particles (e.g., neutrons and neutral atoms) as well as electrons.

Even in the absence of any formal quantitative theory, it is natural to assume that a particle is most likely to be located where the amplitude of its corresponding de Broglie wave packet is large. However, if the momentum and hence the wavelength are reasonably well defined, the wave packet must extend over many wavelengths, in which case the position of the particle is

quite uncertain. Conversely, if the position is well defined, the wave packet must be confined to a small region of space, and therefore, it must be a superposition of components with many different wavelengths. Hence the momentum is very uncertain. The de Broglie relation (1.4) thus implies that it is impossible to determine simultaneously and precisely the position and the conjugate momentum of a particle.

This qualitative statement is made more precise by the uncertainty principle, which was formulated by Werner Heisenberg (1927) from consideration of a variety of thought experiments in which one tries to measure the position and momentum of a particle but where relation (1.4) applies not only to the particle in question but also to a photon that might be used in the measurement process. According to the uncertainty principle, the uncertainties Δx and Δp_x associated with a simultaneous measurement of coordinate x and conjugate momentum p_x , respectively, satisfy the inequality

$$\Delta x \Delta p_x \geq \frac{\hbar}{2} \quad (1.5)$$

where $\hbar = h/2\pi$. Although in classical mechanics the state at any given time of an isolated system of N particles, each with f degrees of freedom, is determined by specifying Nf generalized coordinates and Nf corresponding generalized momenta, the uncertainty principle tells us that this specification cannot be done precisely. A coordinate and the corresponding momentum are incompatible observables.

Intuitively, it is clear that because not only material particles but also photons obey the de Broglie relation (1.4), there should be an uncertainty principle for the electromagnetic field. Indeed, this is so (Jordan and Pauli 1928), although the uncertainty relation for electromagnetic field components is more complicated than for nonrelativistic material particles. We need not be concerned with such complications here. The main point for our present discussion is that the classical prescription for specifying a state of the electromagnetic field at any given time, by giving each component of the electric and magnetic fields at every point in space, cannot always be achieved.

It is easy to see from the de Broglie relation and the uncertainty principle that the Bohr-Sommerfeld model has a fatal defect, for in that model one starts in any given situation by finding the possible classical orbits of an electron or electrons and then selects from those orbits the ones that satisfy the Bohr-Sommerfeld quantization conditions. However, given the incompatibility of coordinate and conjugate momentum, and specifically the uncertainty relation (1.5), such orbits are in general not observable and indeed have no meaning, especially for states such as a ground state, that have small quantum numbers. In fact, looking back on the Bohr-Sommerfeld model from the viewpoint of quantum mechanics, and using the Wentzel-Kramers-Brillouin (WKB) approximation, one can show that the Bohr-Sommerfeld quantization conditions are valid only in the limiting case in which the potential energy varies very slowly over distances comparable with the linear dimensions of an electron wave packet [see, e.g., equation (9.18)].

Although the Bohr-Sommerfeld model was recognized for this and other reasons to be defective and was eventually replaced by quantum mechanics, it turned out that the Bohr-Sommerfeld quantum numbers did not have to be discarded wholesale; rather, some of these numbers could be retained if given new interpretations and new names. Consequently, after quantum mechanics was invented, the results of many analyses of atomic and molecular

spectra carried out before 1925 could be salvaged, including most interpretations of Zeeman-effect data, Stoner's very useful contribution, and the exclusion principle itself.

We have seen in this section that deep and broad flaws appeared in the classical picture of the atomic world in the first quarter of the twentieth century. These flaws were so fundamental and serious that it would be necessary to replace the entire classical edifice with a radically different theory – quantum mechanics. It should be no surprise that these radically new concepts required a new mathematical language that was quite different from the mathematical language of classical physics. It turned out that the natural mathematical language of quantum mechanics is the theory of linear vector spaces and, in particular, Hilbert spaces. Therefore, before we discuss the rules of quantum mechanics in Chapter 3, we review and summarize some of the most important features of Hilbert spaces in Chapter 2.

2

Mathematical Review

In this chapter we summarize the most important definitions and theorems concerning Hilbert spaces that are relevant for quantum mechanics. Much of the material that follows is quite elementary and is probably well known to most readers. We discuss it mainly to establish a common language and notation. The reader will notice as we proceed that our standards of rigor are low and would be scorned by a proper mathematician. For example, we omit any discussion of convergence when considering infinite-dimensional spaces.

2.1 Linear vector spaces

A linear vector space S consists of certain elements $|u\rangle, |v\rangle, \dots$ called *vectors* together with a field of ordinary numbers (sometimes called *c-numbers*) $a, b, c, \dots$. In quantum mechanics, the latter are the complex numbers, and we deal with complex vector spaces. The vectors $|u\rangle, |v\rangle, \dots$ and the numbers $a, b, c, \dots$ satisfy the following rules:

1. *Vector addition is defined.* If $|u\rangle$ and $|v\rangle$ are members of S , there exists another vector $|w\rangle$, also a member of S , such that

$$|w\rangle = |u\rangle + |v\rangle \tag{2.1}$$

2. *Vector addition is commutative.*

$$|u\rangle + |v\rangle = |v\rangle + |u\rangle \quad \text{for all } |u\rangle, |v\rangle \tag{2.2}$$

3. *There exists a null vector $|0\rangle$ or simply 0 such that*

$$|u\rangle + |0\rangle = |0\rangle + |u\rangle = |u\rangle \quad \text{for any } |u\rangle \tag{2.3}$$

4. *Multiplication of a vector $|u\rangle$ by any c-number a is defined.*

$$|u'\rangle = a|u\rangle \tag{2.4}$$

- is in the same “direction” (along the same ray) as $|u\rangle$.
5. *The following distributive law holds.*

$$a(|u\rangle + |v\rangle) = a|u\rangle + a|v\rangle \tag{2.5}$$

2.2 Subspaces

A vector space may contain subspaces. A *subspace* is a subclass of the space, itself having the properties of a vector space. For example, ordinary Euclidean 3-space contains as subspaces all the straight lines passing through the origin and all the two-dimensional planes that pass through the origin. All subspaces possess the null vector in common. They may or may not possess other vectors in common. If they do not, they are said to be *orthogonal subspaces*.

2.3 Linear independence and dimensionality

Vectors $|u_1\rangle, |u_2\rangle, \dots, |u_n\rangle$ are by definition linearly independent if and only if the equation

$$a_1 |u_1\rangle + a_2 |u_2\rangle + \dots + a_n |u_n\rangle = 0 \tag{2.6}$$

has no solution except for the trivial solution

$$a_1 = a_2 = \dots = a_n = 0 \tag{2.7}$$

Suppose that in a certain space S there are n linearly independent vectors $|u_1\rangle, |u_2\rangle, \dots, |u_n\rangle$, but any $n + 1$ vectors are linearly dependent. Then, by definition, the space is n -dimensional. The number n may be finite, denumerably infinite, or even continuously infinite. In most of the following discussion, we pretend that the space in question has finite n , but the results we obtain can be extended in a natural way to the other two cases. Particular problems associated with infinite dimensionality will be dealt with as we come to them (see, e.g., Section 2.14).

In an n -dimensional space where n is finite, n linearly independent vectors $|u_1\rangle, |u_2\rangle, \dots, |u_n\rangle$ are said to span the space or form a basis for the space. This means that any vector $|w\rangle$ can be expressed as a linear combination of the $|u_i\rangle$; that is,

$$|w\rangle = \sum_{i=1}^n a_i |u_i\rangle \tag{2.8}$$

where the a_i are complex numbers.

2.4 Unitary spaces: The scalar product

A *unitary space* is one in which for any two vectors $|u\rangle, |v\rangle$ the scalar product $\langle u | v \rangle$ is defined as a complex number with the following properties:

1. $\langle u | v \rangle = \overline{\langle v | u \rangle}$, where the bar means complex conjugate (thus $\langle u | u \rangle$ is real).

2. $\langle u | av \rangle = a \langle u | v \rangle$. Thus,

$$\langle au | v \rangle = \overline{\langle v | au \rangle} = a^* \langle u | v \rangle$$

where the asterisk means complex conjugate.

3. $\langle u | v + w \rangle = \langle u | v \rangle + \langle u | w \rangle$.

4. $\langle u | u \rangle > 0$ unless $|u\rangle = 0$.

We also use the following terminology:

$$\langle u | u \rangle = \text{norm of } |u\rangle \quad \sqrt{\langle u | u \rangle} = \text{length of } |u\rangle$$

Also, if $|w\rangle \neq 0$, $|u\rangle \neq 0$, but $\langle w | u \rangle = 0$, $|u\rangle$ and $|w\rangle$ are said to be *orthogonal*.

2.5 Formation of an orthonormal basis: Completeness – definition of Hilbert spaces

Given a set of linearly independent vectors $|u_1\rangle, |u_2\rangle, \dots, |u_n\rangle$ in a unitary space, we can construct an orthonormal basis as follows: form the unit vector

$$|\phi_1\rangle = \frac{|u_1\rangle}{\sqrt{\langle u_1 | u_1 \rangle}} \tag{2.9}$$

Next, form

$$|\phi_2\rangle = a|u_2\rangle + b|\phi_1\rangle \tag{2.10}$$

with a, b chosen so that $\langle \phi_1 | \phi_2 \rangle = 0$ and $\langle \phi_2 | \phi_2 \rangle = 1$. That is,

$$\langle \phi_1 | \phi_2 \rangle = a \langle \phi_1 | u_2 \rangle + b = 0 \tag{2.11}$$

which yields $b = -a \langle \phi_1 | u_2 \rangle$ and thus

$$|\phi_2\rangle = a[|u_2\rangle - |\phi_1\rangle \langle \phi_1 | u_2 \rangle] \tag{2.12}$$

with a chosen so that $\langle \phi_2 | \phi_2 \rangle = 1$. Next, define $|\phi_3\rangle$ by

$$|\phi_3\rangle = a'[|u_3\rangle - |\phi_1\rangle \langle \phi_1 | u_3 \rangle - |\phi_2\rangle \langle \phi_2 | u_3 \rangle] \tag{2.13}$$

with a' chosen to normalize $|\phi_3\rangle$, and so on. The result of this *Schmidt process* is a basis of orthogonal unit vectors (orthonormal vectors) $|\phi_1\rangle, \dots, |\phi_n\rangle$.

A set of basis vectors in a unitary space is said to be *complete* if any vector in the space can be expressed as a linear combination of the basis members. For finite-dimensional vector spaces, any set that spans the space is complete. If the space has infinite dimensionality,