

# Chapter 1

## Introduction

As a science or engineering major, you are about to embark on what may be the most important course of your undergraduate career. Quantum mechanics is the foundation on which our current picture of the structure of matter is built. For students, an introductory course in quantum mechanics can be difficult and frustrating. When you studied Newtonian physics, it was easy to envision experiments involving the motion of particles moving under the influence of forces. Electromagnetism and optics were a bit more abstract once the concept of fields was introduced, but you are familiar with many optical effects such as colors of thin films, the rainbow, and diffraction from single or double slits. In quantum mechanics, it is more difficult conceptually to understand what is going on since you have very little day to day experience with the wave nature of matter. However, you can exploit your knowledge of both classical mechanics and optics to get a better feel for quantum mechanics.

I often begin a course in quantum mechanics asking students “Why and when is quantum mechanics needed to explain the dynamics of a particle moving in a potential. In other words, when does a classical particle description fail?” To understand and appreciate many aspects of quantum mechanics, you should always have this and a few other questions in the back of your mind. Question two might be “How is the quantum mechanical solution of a given problem related to the corresponding problem in classical mechanics?” Question three could be “Is there a problem in optics with which I am familiar that can shed light on a given problem in quantum mechanics?” You can add other questions to this list, but these are a good start.

Another key point in the study of quantum mechanics is not to lose track of the physics. There are many special functions, such as Hermite and Laguerre polynomials, that emerge from solutions of the Schrödinger equation, and the algebra can get a little complicated. As long as you remind yourself of the physical nature of the solution rather than the mathematical details, you will be in great

shape. The reason for this is that most of the solutions share many common features. The specific form of the solutions may change, but the overall qualitative nature of the solutions is remarkably similar for many different problems. Moreover, with the availability of symbolic manipulation programs such as Mathematica, Maple, or Matlab, it is now easy to plot and evaluate any of these functions using a few keystrokes.

To help introduce the subject matter, I will present a very broad, qualitative overview of the way in which quantum mechanics was born, a birth that took about 25–30 years. I will not worry about historical accuracy here, but simply try to give you a reasonable picture of the manner in which it became appreciated that a wave description of matter was needed in certain limits.

## 1.1 Electromagnetic Waves

Quantum mechanics is a wave theory. Since the wave properties of matter have many similarities to the wave properties of electromagnetic radiation, it won't hurt to review some of the fundamental properties of the electromagnetic field. The possible existence of electromagnetic waves followed from Maxwell's equations. By combining the equations of electromagnetism, Maxwell arrived at a wave equation, in which the wave propagation speed in vacuum was equal to  $v = 1/\sqrt{\epsilon_0\mu_0}$ , where  $\epsilon_0$  is the permittivity and  $\mu_0$  the permeability of free space. Since these were known quantities in the nineteenth century, it was a simple matter to calculate  $v$ , which turned out to equal the speed of light. This result led Maxwell to conjecture that light was an electromagnetic phenomenon.

All electromagnetic waves travel in vacuum with speed  $c = 2.99792458 \times 10^8$  m/s, now *defined* as the speed of light. What distinguishes one type of electromagnetic wave from another is its wavelength  $\lambda$  or frequency  $f$ , which are related by

$$c = f\lambda. \quad (1.1)$$

Instead of characterizing a wave by its frequency (which has units of cycles per second or Hz) and wavelength, we can equally well specify the *angular frequency*,  $\omega = 2\pi f$ , (which has units of radians per second or  $s^{-1}$ ) and the magnitude of the *propagation vector* (or *wave vector*), defined by  $k = 2\pi/\lambda$ . With these definitions, Eq. (1.1) can be replaced by

$$\omega = ck, \quad (1.2)$$

which is known as a *dispersion relation*, relating frequency to wave vector. For electromagnetic radiation in vacuum the dispersion relation is linear.

The *source* of electromagnetic waves is oscillating or accelerating charges, which give rise to propagating electric and magnetic fields. The wave equation for the electric field vector  $\mathbf{E}(\mathbf{R}, t)$  at position  $\mathbf{R}$  at time  $t$  in vacuum is

$$\nabla^2 \mathbf{E}(\mathbf{R}, t) = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}(\mathbf{R}, t)}{\partial t^2}. \quad (1.3)$$

The simplest solution of this equation is also the most important, since it is a building block solution from which all other solutions can be constructed. The building block solution of the wave equation is the infinite, monochromatic, plane wave solution, having an electric field vector given by

$$\mathbf{E}(\mathbf{R}, t) = \epsilon E_0 \cos(\mathbf{k} \cdot \mathbf{R} - \omega t), \quad (1.4)$$

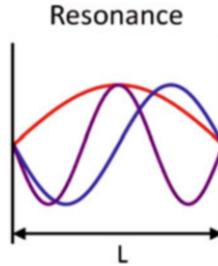
where  $E_0$  is the field amplitude and the *polarization* of the field is specified by a unit vector  $\epsilon$  that is perpendicular to the propagation vector  $\mathbf{k}$  of the field. There are two independent field polarizations possible for each propagation vector. The fact that  $\epsilon \cdot \mathbf{k} = 0$  follows from the requirement that  $\nabla \cdot \mathbf{E}(\mathbf{R}, t) = 0$  in vacuum. The field (1.4) corresponds to a wave that is infinite in extent and propagates in the  $\mathbf{k}$  direction. Of course, no such wave exists in nature since it would uniformly fill all space.

As a consequence of the linearity of the wave equation, the sum of any two solutions of the wave equation is also a solution. This is known as a *superposition principle*. It is not difficult to visualize how waves add together. If, at the same time you are putting your finger in and out of water in a lake, someone else is also putting their finger in and out of the water, the resulting wave results from the actions of *both* your fingers. The important thing to remember is that it is the *displacements* or *amplitudes* of the waves that add, not their intensities. The *intensity* of a wave is proportional to the square of the wave amplitude.

For the time being, let us consider two waves having the same frequency and the same amplitude. If the two waves propagate in opposite directions, the total electric field vector is

$$\begin{aligned} \mathbf{E}(\mathbf{R}, t) &= \epsilon E_0 [\cos(\mathbf{k} \cdot \mathbf{R} - \omega t) + \cos(-\mathbf{k} \cdot \mathbf{R} - \omega t)] \\ &= 2\epsilon E_0 \cos(\omega t) \cos(\mathbf{k} \cdot \mathbf{R}), \end{aligned} \quad (1.5)$$

a *standing wave* pattern. If you plot the wave amplitude at several different times as a function of position, you will find that there is an envelope for the wave that is *fixed* in space—the wave “stands” there and oscillates within the envelope. The points of zero amplitude are called *nodes* of the field and the maxima (or minima) are called *antinodes* of the field. For a standing wave field in the  $x$  direction that is confined between two parallel mirrors separated by a distance  $L$ , the standing wave pattern will “fit in” for wavelengths equal to  $2L$ ,  $L$ ,  $2L/3$ ,  $L/2$ , etc., as shown in



**Fig. 1.1** Resonance involving standing waves with clamped endpoints

Fig. 1.1. In this limit, the tangential component of the electric field vanishes at the mirrors, as required by the boundary conditions on the field. The condition

$$\lambda_n = 2L/n; \quad n = 1, 2, 3, \dots \quad (1.6)$$

is known as a *resonance* condition.

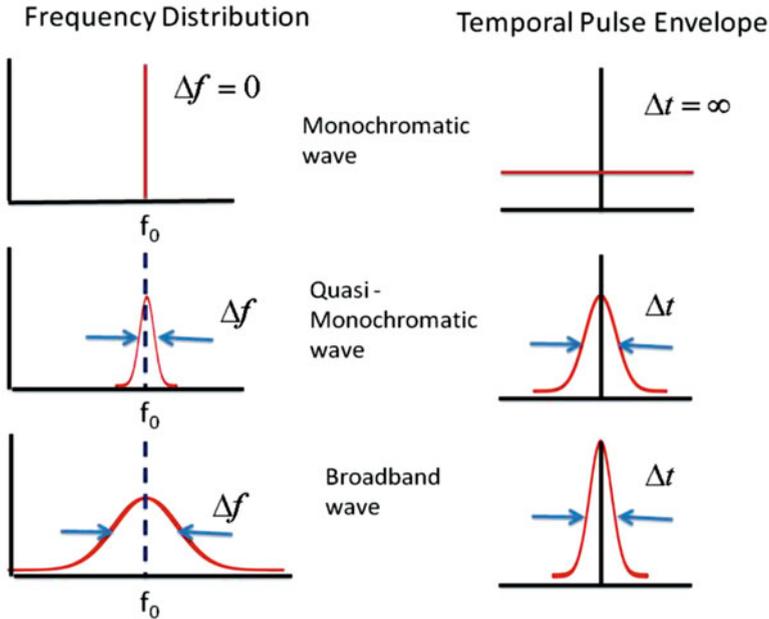
### 1.1.1 Radiation Pulses

To get a *pulse* of radiation, it is necessary to add together monochromatic waves having a *continuous* distribution of frequencies. It is *easy* to create a pulse of radiation. Simply turn a laser or other light source on and off and you have created a pulse. Why worry about the frequencies contained in the pulse? It turns out that it is important, even central, to understand this concept if you are going to have some idea of what quantum mechanics is about.

Let us assume the pulse is propagating in the  $x$  direction and has a duration  $\Delta t$  corresponding to a spatial width  $\Delta x = c\Delta t$ . Clearly this cannot be a monochromatic wave since a monochromatic wave is not localized. Instead, the pulse must be a superposition of waves having a range of frequencies  $\Delta f$  centered around some average frequency  $f_0$ . Using the theory of Fourier analysis it is possible to show that  $\Delta f$  and  $\Delta t$  are related by

$$\Delta f \Delta t \approx \frac{1}{2\pi}. \quad (1.7)$$

The quantity  $\Delta f$  is known as the *spectral width* of the pulse. If the frequency of the pulse is known *precisely*, as in a monochromatic wave, there is no *range* of frequencies and  $\Delta f = 0$ . In general there is a central frequency  $f_0$  in a pulse and a range of frequencies  $\Delta f$  about that central frequency as shown in Fig. 1.2. If  $\Delta f \ll f_0$ , then the frequency is pretty well-defined and the field is said to be *quasi-monochromatic*. The field associated with the light of a green laser pointer



**Fig. 1.2** Illustration of the equation  $\Delta f \Delta t \approx 1/2\pi$ . As the spectral frequency distribution of a pulse broadens, the temporal pulse width narrows

and most other laser fields are quasi-monochromatic, as are the fields from neon discharge tubes. Radio waves are also quasi-monochromatic, the central frequency is the frequency of the station (about 1000 kHz for AM and 100 MHz for FM) broadcasting the signal. On the other hand, if  $\Delta f$  is comparable with  $f_0$ , as in a light bulb, the source is said to be incoherent or broadband.

For example, consider a pulse of green laser light that has an ns ( $10^{-9}$  s) duration,  $\Delta t = 1$  ns. The spatial extent of this pulse is  $3 \times 10^8$  m/s  $\cdot$   $10^{-9}$  s = 30 cm. The frequency range in the pulse is given by

$$\Delta f = 1 / (2\pi 10^{-9} \text{ s}) = 1.6 \times 10^8 \text{ Hz.} \tag{1.8}$$

Since the light from the green laser pointer has a central wavelength of 532 nm, it has a central frequency of about  $f_0 = 5.6 \times 10^{14}$  Hz. As a consequence,  $\Delta f \ll f_0$  and this pulse is quasi-monochromatic. That is, a one nanosecond pulse of this light appears to have a single frequency or color if you look at it. On the other hand, if you try to make a one femtosecond (fs) pulse with this light ( $1 \text{ fs} = 10^{-15}$  s) which has a spatial extent of 300 nm, then

$$\Delta f = 1 / (2\pi 10^{-15} \text{ s}) = 1.6 \times 10^{14} \text{ Hz,} \tag{1.9}$$

which is comparable to the central frequency. As a result this is a broadband pulse which no longer appears green, but closer to white. Note that the spatial extent of the

pulse, 300 nm is comparable to the central wavelength. This is an equivalent test for broadband radiation. If the spatial extent is comparable to the central wavelength, the radiation is broadband; if it is much *larger* than the central wavelength, it is quasi-monochromatic.

### 1.1.2 Wave Diffraction

The final topic that we will need to know something about is wave diffraction. Diffraction is a purely wave phenomenon, but sometimes waves don't appear to diffract at all. If you shine a laser beam at a tree or a pencil, it will reflect off of these and not bend around them. Moreover, if you shine a laser beam through an open door, it will not bend around into the hallway. *Diffraction is only important when waves meet obstacles (including apertures or openings) that are comparable to or smaller than the central wavelength of the waves.* This is actually true for the laser beam itself!—it will stay a beam of approximately constant diameter only if its diameter is much greater than its wavelength, otherwise, it will spread significantly.

I can make this somewhat quantitative. Imagine there is a circular opening in a screen having diameter  $a$  through which a laser beam passes, or that it passes through a slit having width  $a$ , or that a laser beam having diameter  $a$  propagates in vacuum with no apertures present. In each of these cases, there is diffraction that leads to a spreading of the beam by an amount

$$\sin \theta \approx \frac{\lambda}{a}, \quad (1.10)$$

where  $\theta$  is the angle with which the beam spreads. If  $\lambda/a \ll 1$ , the spreading angle is very small and diffraction is relatively unimportant. *This is the limit where the wave acts as a particle, moving on straight lines.* On the other hand, when  $a$  gets comparable to a wavelength, the spreading of the beam becomes significant. If  $\lambda/a > 1$ , the spreading is over all angles. For example, if a pinhole is illuminated with light, you can see the diffracted light at any angle on the other side of the pinhole.

Even if the diffraction angle is small, the effects can get large over long distances. If a laser beam having diameter 1.0 cm and central wavelength 600 nm is sent to the moon, the diffraction angle is

$$\theta \approx \frac{\lambda}{a} = \frac{6 \times 10^{-7} \text{ m}}{10^{-2} \text{ m}} = 6 \times 10^{-5}. \quad (1.11)$$

By the time it gets to the moon the spot size diameter  $d$  is

$$d \approx R_{\text{EM}}\theta = 3.8 \times 10^8 \text{ m} \times 6 \times 10^{-5} = 2.3 \times 10^4 \text{ m} = 23 \text{ km}, \quad (1.12)$$

a lot larger than 1 cm! In this equation  $R_{\text{EM}}$  is the Earth-moon distance.

*The bottom line on diffraction. If you try to confine a wave to a distance less than or comparable to its wavelength, it will diffract significantly.*

## 1.2 BlackBody Spectrum: Origin of the Quantum Theory

Thermodynamics was developed in the nineteenth century and involves the study of the properties of vapors, liquids, and solids in terms of such parameters as temperature, pressure, volume, density, conductivity, etc. At the end of the nineteenth century, the theory of statistical mechanics was formulated in which the properties of systems of particles are explained in terms of their statistical properties. It was shown that the two theories of thermodynamics and statistical mechanics were consistent—one could explain macroscopic properties of vapors, liquids, and solids by considering them to be made up of a large number of particles following Newton’s laws. One result of this theory is that for a system of particles (or waves) in thermodynamic equilibrium at a temperature  $T$ , each particle or wave has  $(1/2)k_B T$  of energy for each “degree of freedom.” This is known as the *equipartition theorem*.

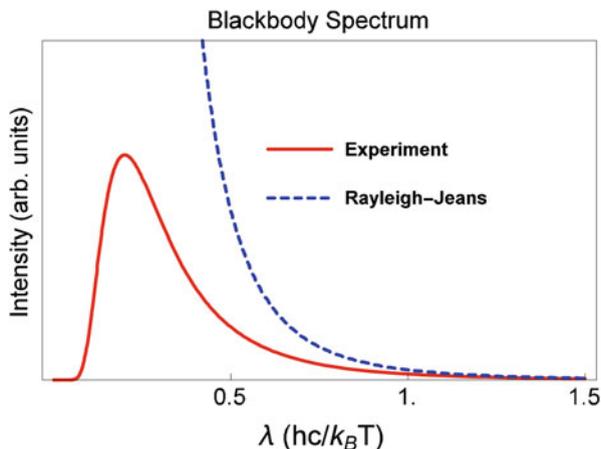
The quantity  $k_B = 1.38 \times 10^{-23} \text{ J/}^\circ\text{K}$  is Boltzmann’s constant and  $T$  is the absolute temperature in degrees Kelvin. A “degree of freedom” is related to an independent motion a particle can have (translation, rotation, vibration, etc.). For a free particle (no forces acting on it), there are three degrees of freedom, one for each independent direction of motion. For a transverse wave, there are two degrees of freedom, corresponding to the two possible independent directions for the polarization of the wave.

At the beginning of the twentieth century there was a problem in trying to formulate the theory of a *blackbody*. A blackbody is an object that, in equilibrium, absorbs and emits radiation at the same rate. At a given temperature, the radiation emitted by a blackbody is spread over a wide range of frequencies, but the peak intensity occurs at a wavelength  $\lambda_{\max}$  governed by Wien’s law,

$$\lambda_{\max} T = 2.90 \times 10^{-3} \text{ m} \cdot ^\circ\text{K}, \quad (1.13)$$

where  $T$  is the temperature in degrees Kelvin. For example, the surface of the sun is about  $5000^\circ\text{K}$ . If the sun is approximated as a blackbody, then  $\lambda_{\max} \approx 580 \text{ nm}$  is in the yellow part of the spectrum. As an object is heated it emits radiation at higher frequencies; an object that is “blue” hot is hotter than an object that is “red” hot.

The experimental curve for emission from a blackbody as a function of wavelength is shown in Fig. 1.3. To try to explain this result, Rayleigh and Jeans considered a model for a blackbody consisting of a cavity with a small hole in it. Standing wave patterns of radiation fill the cavity and each standing wave has  $k_B T$  of energy associated with it ( $k_B T/2$  for each independent polarization of the field). A standing wave pattern or *mode* is characterized by the number of (half) wavelengths in the  $x$ ,  $y$ , and  $z$  directions. There is a maximum wavelength  $\lambda = 2L$  in each direction, but there is *no* limit on the minimum wavelength. Since the number



**Fig. 1.3** Intensity per unit wavelength (in arbitrary units) as a function of wavelength [in units of  $(hc/k_B T)$ ] for a blackbody. The maximum occurs at  $\lambda_{\max} \approx hc / (5k_B T)$

of small wavelength (high frequency) modes that can fit is arbitrarily large and since each mode has  $k_B T$  of energy, the amount of energy needed to reach equilibrium becomes infinite. Another way of saying this is that the energy density (energy per unit frequency interval) approaches infinity as  $\lambda \sim 0$  or  $f \sim \infty$ . The fact that it does not take an infinite amount of energy to bring a blackbody into thermal equilibrium or that the energy density does not approach infinity at high frequencies is known as the “ultraviolet catastrophe”; theory and experiment were in disagreement for high or “ultraviolet” frequencies (smaller wavelengths) (see Fig. 1.3).

Planck tried to overcome this difference by making an *ad hoc* hypothesis.<sup>1</sup> In effect, he said, that to excite a mode of frequency  $f$ , a specific *minimum* amount of energy  $hf$  ( $h$  is defined below) is needed that must be provided in an *all or nothing* fashion by electrons oscillating in the cavity walls. This went against classical ideas that energy can be fed continuously to build up oscillation in a given mode. By adopting this theory, he found that he could explain the experimental data if he chose

$$h = 6.63 \times 10^{-34} \text{ J} \cdot \text{s} = 4.14 \times 10^{-15} \text{ eV} \cdot \text{s}, \quad (1.14)$$

which is now known as *Planck’s constant* [recall that one eV (electron-volt) is equal to  $1.60 \times 10^{-19}$  J (Joules)]. The quantum theory was born. Some details of the Planck solution are given in the Appendix.

In thermal equilibrium at temperature  $T$ , electrons in the walls of the cavity have a maximum energy of order of a few  $k_B T$ , so that they are capable of exciting only

<sup>1</sup>*On the law of the distribution of energy in the normal distribution*, Annalen der Physik **4**, 553–563 (1901).

those radiation modes having  $hf \lesssim k_B T$ . Quantitatively, it can be shown that the maximum in the blackbody spectrum as a function of frequency occurs at

$$f_{\max} = \frac{2.82k_B T}{h} = 5.88 \times 10^{10} T \text{ Hz}/^\circ\text{K}. \quad (1.15)$$

As a function of wavelength the maximum in the blackbody spectrum occurs at

$$\lambda_{\max} = \frac{hc}{4.965k_B T} = \frac{2.90 \times 10^{-3}}{T} \text{ m} \cdot ^\circ\text{K}. \quad (1.16)$$

The numerical factors appearing in these equations and the fact that

$$f_{\max} \lambda_{\max} \neq c \quad (1.17)$$

follow from the details of the Planck distribution law (see Appendix). Visible radiation having  $\lambda_{\max} = 500 \text{ nm}$  corresponds to a blackbody temperature of  $T \approx 6,000^\circ\text{K}$ , about the temperature of the surface of the sun. In the visible part of the spectrum,  $hf$  is of order of a few eV.

In 1907, Einstein also used the Planck hypothesis to explain a feature that had been observed in the specific heat of solids as a function of temperature.<sup>2</sup> Many solids have a specific heat at constant volume equal to  $3R$  at room temperature, where  $R = 8.31 \text{ J/mole}/^\circ\text{K}$  is the gas constant; however, some substances such as diamond have a much lower specific heat. Based on a model of solids composed of harmonic oscillators, Einstein used Planck's radiation law to show that the contributions to the specific heat of diamond from the oscillations are "frozen out" at room temperature. To fit the data on diamond, Einstein used a value of  $2.73 \times 10^{13} \text{ Hz}$  for the frequency of oscillations, which would imply that the vibrational degrees of freedom begin to diminish for temperatures below  $T = hf/k_B \approx 1300^\circ\text{K}$  and are effectively frozen out for temperatures below  $T = hf/(2.82k_B) \approx 460^\circ\text{K}$ .

### 1.3 Photoelectric Effect

The *all or nothing* idea surfaced again in Einstein's explanation of the photoelectric effect in Section 8 of his 1905 paper, *On a heuristic point of view concerning the production and transformation of light*.<sup>3</sup> It had been noticed by Hertz and J. J. Thomson in 1897 and 1899, respectively, that electrons could be ejected from metal surfaces when light was shined on the surfaces. Einstein gave his explanation of the effect in 1905 (this explanation and *not* relativity theory was noted in his Nobel Prize

<sup>2</sup>Planck theory of radiation and the theory of specific heats, *Annalen der Physik* **22**, 180–190 (1907).

<sup>3</sup>*Annalen der Physik* **17**, 132–148 (1905).

citation) and further experiments by Millikan confirmed Einstein's explanation. It turns out that Einstein's explanation is not really evidence for *photons*, even though the authors of many textbooks assert this to be the case. To understand the photoelectric effect, it is necessary to know that different metals have different *work functions*. The work function (typically on the order of several eV) is the minimum energy needed to extract an electron from a metal surface. In some sense it is a measure of the height of a hill that the electron must climb to get out of the metal. From a physical perspective, when an electron leaves the surface, it creates an image charge inside the surface that attracts the charge back to the surface; the work function is the energy needed to escape from this attractive force.

The experimental observations for the photoelectric effect can be summarized as follows:

1. When yellow light is shined on a specific metal, no electrons are ejected. Increasing the intensity of the light still does not lead to electrons being ejected.
2. When ultraviolet light is shined on the same metal, electrons are ejected. When the intensity of the ultraviolet radiation is increased, the number of electrons ejected increases, but the *maximum kinetic energy of the emitted electrons does not change*.

To explain these phenomena, Einstein stated that radiation consists of particles (subsequently given the name *photons* by the chemist Gilbert Lewis in 1926). For radiation of frequency  $f$ , the photons have energy  $hf$ . When such particles are incident on a metal surface, they excite the electrons by giving up their energy in an all or nothing fashion. Thus, if the work function is denoted by  $W$ , the frequency of the radiation must satisfy  $hf > W$  for the radiation to cause electrons to be ejected. If  $hf < W$ , the photons cannot excite the electrons, no matter what the intensity of the field—this explains why increasing the intensity of the yellow light does not lead to electrons being ejected. On the other hand, if  $hf > W$ , the photon can excite the electron with any extra energy going into kinetic energy  $KE$  of the electrons (some of which may be lost on collisions). Thus the maximum energy of the emitted electrons is

$$KE_{\max} = hf - W. \quad (1.18)$$

Increasing the intensity of the radiation does not change this maximum kinetic energy since it results from single photon events, it affects only the intensity of the emitted electrons.

## 1.4 Bohr Theory

By 1910, it was known that there existed negatively charged particles having charge  $-e$  equal to  $-1.6 \times 10^{-19}$  C and mass  $m$  equal to  $9.1 \times 10^{-31}$  kg. Moreover, the size of these particles could be estimated using theoretical arguments related to their

energy content to be about  $10^{-15}$  m. From atomic densities and measurements in gases, atoms were known to have a size of about  $10^{-10}$  m. Since matter is neutral, there must be positively charged particles in atoms that cancel the negative charge.

But how were the charges arranged? Was the positive charge spread out over a large sphere and the negative charge embedded in it or was there something like a planetary model for atoms? This question was put to rest in 1909 by Geiger and Madsen who collided alpha particles (helium nuclei that are produced in radioactivity) on thin metal foils. They found “back-scattering” that indicated the positive charges were small. In 1911 Rutherford analyzed the data and estimated the positive charges to have a size on the order of  $10^{-15}$  m. He then proposed a planetary model of the atom in which the electron orbited the positive nucleus.

However, there were problems in planetary—model—ville. It is easy to calculate that an accelerating electron in orbit around the nucleus should radiate its energy in about a nanosecond, yet atoms were stable. Moreover, the spectrum of radiation emitted by the hydrogen atom consisted of a number of discrete (quasi-monochromatic) lines, but accelerating electrons in different orbits should produce a broadband source of radiation. How could this be explained?

To explain the experimental data within the context of the planetary model, Bohr in 1913 came up with the following postulates:

1. The electron orbits the nucleus in *circular orbits* (back to Ptolemy and Copernicus) having discrete values of *angular momentum* given by

$$L = mvr = n\hbar, \quad (1.19)$$

where  $n = 1, 2, 3, 4, \dots$  (actually he formulated the law in terms of energy rather than angular momentum, but the two methods yield the same results). The quantity  $r$  is the radius of the orbit,  $v$  is the electron’s speed, and  $L$  is the magnitude of the angular momentum. The mass in this equation should actually be the reduced mass of the electron which is smaller than the electron mass by a factor of  $1/1.00054$ .

2. The electrons radiate energy only when they “jump” from a larger to a smaller orbit. The frequency of the radiation emitted is given by

$$f_{n_2, n_1} = \frac{E_{n_2} - E_{n_1}}{h}, \quad (1.20)$$

where  $E_{n_i}$  is the energy associated with an orbit having angular momentum  $L = n_i\hbar$ .

Although these are seemingly benign postulates, they have extraordinary consequences. To begin with, I combine the first postulate with the laws of electrical attraction and Newton’s second law. The magnitude of the electrostatic force on the electron produced by the proton is

$$F = \frac{1}{4\pi\epsilon_0} \frac{e^2}{r^2}. \quad (1.21)$$

Combing this with Newton's second law, I find

$$\frac{1}{4\pi\epsilon_0} \frac{e^2}{r^2} = ma = \frac{mv^2}{r}, \quad (1.22)$$

or

$$mvr = \frac{e^2}{4\pi\epsilon_0}. \quad (1.23)$$

Equations (1.19) and (1.23) can be solved for the allowed radii and speeds of the electron. The possible speeds are

$$v_n = \frac{\alpha_{FS}c}{n}, \quad (1.24)$$

where

$$\alpha_{FS} = \frac{1}{4\pi\epsilon_0} \frac{e^2}{\hbar c} \approx \frac{1}{137} \quad (1.25)$$

is known as the *fine-structure constant*. The allowed radii are

$$r_n = n^2 a_0, \quad (1.26)$$

where

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2} = \frac{\bar{\lambda}_c}{\alpha_{FS}} \approx 0.0529 \text{ nm} \quad (1.27)$$

is the *Bohr radius*,  $\bar{\lambda}_c = \lambda_c/2\pi$ , and

$$\lambda_c = h/mc = 2.43 \times 10^{-12} \text{ m} \quad (1.28)$$

is the *Compton wavelength* of the electron. The allowed energies are given by

$$E_n = \frac{mv_n^2}{2} - \frac{1}{4\pi\epsilon_0} \frac{e^2}{r_n} = -\frac{E_R}{n^2}, \quad (1.29)$$

where the *Rydberg energy*  $E_R$  is defined as

$$E_R = \frac{1}{2} mc^2 \alpha_{FS}^2 \approx 13.6 \text{ eV} = 2.18 \times 10^{-18} \text{ J}. \quad (1.30)$$

Equations (1.19), (1.24), (1.26), and (1.29) are amazing in that they predict that only orbits having *quantized* values of angular momentum, speed, radius, and

energy are permitted, totally in contradiction to any classical dynamic models. Moreover the smallest allowed radius  $r_1 = a_0$  gives the correct order of magnitude for the size of atoms. In addition, the energies could be used to explain the discrete nature of the spectrum of the hydrogen atom.

Bohr's second postulate addresses precisely this point. Since the energies are quantized, the frequencies emitted by an excited hydrogen atom are also quantized, having possible values

$$f_{nq} = \frac{E_n - E_q}{h} = \frac{E_R}{h} \left( \frac{1}{q^2} - \frac{1}{n^2} \right) = 3.2880 \times 10^{15} \left( \frac{1}{q^2} - \frac{1}{n^2} \right) \text{ Hz}, \quad (1.31)$$

with  $n > q$ . The corresponding wavelengths are

$$\lambda_{nq} = 91.18 \frac{n^2 q^2}{n^2 - q^2} \text{ nm}. \quad (1.32)$$

The different frequencies of the radiation emitted on transitions ending on the  $n = 1$  level are referred to as the *Lyman* series, those ending on  $n = 2$  as the *Balmer* series, and those ending on  $n = 3$  as the *Paschen* series. All Lyman transitions are in the ultraviolet or soft X-ray part of the spectrum. The three lowest frequency transitions in the Balmer series are in the visible (red— $\lambda_{32} = 656$  nm, blue-green— $\lambda_{42} = 486$  nm, and violet— $\lambda_{52} = 434$  nm), with the remaining Balmer transitions in the ultraviolet. All Paschen transitions and those terminating on levels having  $n > 3$  are in the infrared or lower frequency spectral range. The second postulate also “explained” the stability of the hydrogen atom, since an electron in the  $n = 1$  orbital had nowhere to go.

Thus, Bohr's theory explained in quantitative terms the spectrum of hydrogen. That is, the Bohr theory gave very good agreement with the experimental values for the frequencies of the emitted lines. Bohr theory leads to correct predictions for the emitted frequencies because Bohr got the energies right. The Bohr theory also predicts the correct characteristic values for the radius and speed of the electron in the various orbits. As we shall see, however, the radius and speed are not quantized in a correct theory of the hydrogen atom. Moreover, while angular momentum is quantized in the quantum theory, the allowed quantized values for the magnitude of the angular momentum differ somewhat from those given by the Bohr theory.

## 1.5 De Broglie Waves

It turns out that the photoelectric effect and quantization conditions of the Bohr theory of the hydrogen atom can be explained if one allows for the possibility that matter is described by a wave equation. The first suggestion of this type originates with the work of Louis de Broglie in 1923. The basic idea is to explain the stable

orbits of the electron in the hydrogen atom as a repeating wave pattern. Just what type of wave remains to be seen. De Broglie actually arrived at a wavelength for matter using two different approaches.

First let us consider the circular orbits of the electron in hydrogen, for which  $L_n = mv_n r_n = n\hbar$ . If we imagine that an orbit fits in  $n$  wavelengths of matter, then

$$n(\lambda_{dB})_n = 2\pi r_n = \frac{n\hbar}{mv_n}, \quad (1.33)$$

where  $\lambda_{dB}$  is the *de Broglie wavelength*. Solving for  $\lambda_{dB}$ , I find

$$(\lambda_{dB})_n = \frac{h}{mv_n}. \quad (1.34)$$

Thus, if we assign a wavelength to matter equal to Planck's constant divided by the momentum of the particle, then the wave will "repeat" in its circular orbit. If Planck's constant were equal to zero, there would be no wave-like properties to matter.

De Broglie reached this result in another manner starting with Einstein's theory of relativity, combined with aspects of the photoelectric effect. In special relativity the equation for the length of the momentum-energy 4-vector is

$$p^2 c^2 - E^2 = m^2 c^4, \quad (1.35)$$

where  $p$  is the momentum,  $m$  the mass, and  $E$  the energy of the particle. If I apply this equation to "particles" of light having zero mass, I find  $p = E/c$ , which relates the momentum and energy of light. Considering light of frequency  $f$  to be composed of photons having energy  $E = hf$ , we are led to the conclusion that the momentum of a photon is  $p = E/c = hf/c = h/\lambda$ , where  $\lambda$  is the wavelength of the light. Thus, for photons,  $\lambda = h/p$ . De Broglie then suggested that this is the correct prescription for assigning a wavelength to matter as well.

You can calculate the de Broglie wavelength of any macroscopic object and will find that it is extremely small compared with atomic dimensions. Since the de Broglie wavelength of macroscopic matter is much smaller than the size of an atom, macroscopic matter could exhibit its wave-like properties only if it encounters interaction potentials that vary on such a scale, an unlikely scenario. On the other hand, if you calculate the de Broglie wavelength of the electron in the ground state of hydrogen, you will find a value close to the Bohr radius. Since the electron is confined to a distance on the order of its wavelength, it exhibits wave-like properties.

In 1927, Davisson and Germer observed the diffraction and interference of electron waves interacting with a single crystal of nickel, providing confirmation of the de Broglie hypothesis. About the same time Schrödinger was formulating a wave theory of quantum mechanics. The wave theory of matter was born!

## 1.6 The Schrödinger Equation and Probability Waves

With de Broglie's hypothesis, it is not surprising that scientists tried to develop theories in which matter was described by a wave equation. In 1927, Schrödinger developed such a theory at the same time that Heisenberg was developing a theory based on matrices. Eventually the two theories were shown to be equivalent. I will focus on the Schrödinger approach in this discussion; the matrix approach is similar to Dirac's formalism, which is discussed in Chap. 11.

The Schrödinger equation is a partial differential equation for a function  $\psi(\mathbf{r}, t)$  which is called the *wave function*. Before thinking about how to interpret the wave function, it is useful to describe how to solve the Schrödinger equation. The Schrödinger equation is different than the wave equation for electromagnetic waves, although there are some similarities. To solve the Schrödinger equation, one must find the "building block" solutions, that is, those solutions analogous to the infinite plane monochromatic waves that served as the building block solutions of the wave equation in electromagnetism. The building block solutions of the Schrödinger equation are called *eigenfunctions*. Eigenfunctions or *eigenstates* are solutions of the Schrödinger equation for which  $|\psi(\mathbf{r}, t)|^2$  is constant in time. It is not obvious that such solutions exist, but it can be shown that they always can be found. The eigenfunctions are labeled by *eigenvalues* which correspond to dynamic constants of the motion such as energy, momentum, angular momentum, etc. Let me give you some examples.

A free particle is a particle on which no forces act. For such a particle both momentum  $\mathbf{p} = m\mathbf{v}$  and energy  $E = mv^2/2$  are constant. Thus momentum and energy can be used as eigenvalue labels for a free particle. But which is a more encompassing label? I can write the energy of the free particle as  $E = p^2/2m$ . If I give you the momentum, you can tell me the energy. On the other hand, if I give you the energy you can tell me the *magnitude* of the momentum, but not its direction. Thus momentum uniquely determines the eigenfunction since, to each momentum, there is associated exactly one eigenstate. On the other hand, for a given energy there is an infinite number of eigenstates since the momentum can be in any direction. In this case the eigenfunctions or eigenstates are said to be infinitely *degenerate*. We will see that whenever there is a degeneracy of this type, there is an underlying symmetry of nature. In this case there is translational symmetry, since the particle can move in any direction without experiencing a change the force acting on it (since there is no force). The eigenfunctions for the free particle are our old friends, infinite plane waves, but they differ from electromagnetic waves. The wavelength of the free particle waves is just the de Broglie wavelength  $\lambda = h/p$ —one momentum implies one wavelength. Thus it is the momentum that determines the wavelength of the free particle eigenfunctions.

For the bound states of the hydrogen atom, the electron is subjected to a force so that momentum is not constant and cannot be used as a label for the hydrogen atom eigenfunctions. On the other hand, angular momentum is conserved for any central force. For reasons to be discussed in Chap. 9, the three components of the angular momentum vector cannot be used to label the eigenfunctions; instead, the magnitude of the angular momentum and one of its components are used.

Let me return to the free particle. What is the significance of the wave function  $|\psi(\mathbf{r}, t)|^2$ ? The interpretation given is that  $|\psi(\mathbf{r}, t)|^2$  is the probability density (probability per unit volume) to find the particle at position  $\mathbf{r}$  at time  $t$ . To give you some idea of what this means, imagine throwing darts at a wall. You throw one dart and it hits somewhere on the wall. Now you throw a second and a third dart. Now you throw a million darts. After a million darts, you will have a pretty good idea where the next dart will land—in other words, you will know the probability distribution for where the darts will land. This will not tell you where the next dart will hit, only the probability. The probability of hitting a *specific* point is essentially equal to zero—you must talk about the probability density (in this case, probability per unit area). In quantum mechanics we have something similar. You prepare a quantum system in some initial state. At some later time you measure the particle somewhere. Now you start with an *identical* initial state, wait the same amount of time and measure the particle again. You repeat this a very large number of times and you will have the probability distribution to find the particle at a time  $t$  after it was prepared in this state. This will not tell you where the particle will be the next time you do the experiment—only the probability. If you send single particles, acting as waves, through the two-slit apparatus, each particles will set off only one detector on the screen. If you repeat this many times however, eventually you will build up the same type of interference pattern that occurs in the double slit experiment for light.<sup>4</sup>

Since the eigenfunctions for a free particle are infinite mono-energetic (single momentum or velocity) waves, the probability density associated with the free particle eigenfunction is constant over the entire universe! In other words, the eigenfunctions of the free particle are spread out over all space. Clearly no such state exists. Any free particles that we observe in nature are in a superposition state that is referred to as a *wave packet*, which is the matter analogue of a radiation pulse. As with radiation, we make up a superposition state by adding several waves or eigenfunctions together. But radiation pulses and wave packets differ in a fundamental way. All frequency components of a radiation pulse propagate at the same speed in vacuum, the speed of light. For a free particle wave packet, the different eigenfunctions that compose the wave packet correspond to different wavelengths, which, in turn, implies that they correspond to different momenta, since wavelength for matter waves is related to momentum. Thus different parts of a free particle wave packet move at different rates—the wave packet’s shape changes in time! In the case of matter waves, Eq. (1.7) is replaced by

$$\Delta x \Delta p \geq \hbar/2, \quad (1.36)$$

---

<sup>4</sup>A beautiful video of an experiment demonstrating the buildup of a two-slit interference pattern for electrons can be found at <http://www.hitachi.com/rd/portal/highlight/quantum/>.

where  $\Delta x$  is the spread of positions and  $\Delta p$  the spread of momenta in the wave packet. Equation (1.36) is a mathematical expression of the famous *Heisenberg Uncertainty Principle*.

Matter can be considered to be “particle-like” as long as it is not confined to a distance less than its de Broglie wavelength. If a particle is confined to a distance on the order of its de Broglie wavelength or encounters changes in potential energy that vary significantly in distances of order of its de Broglie wavelength, the particle exhibits wave-like properties.

## 1.7 Measurement and Superposition States

Measurement is one of the most difficult and frustrating features of quantum mechanics. The problem is that the measuring apparatus itself must be classical and not described by quantum mechanics. You will hear about “wave-function collapse” and the like, but quantum mechanics does not describe the dynamics in which a quantum state is measured by a classical apparatus. As long as you ask questions related to the probability of one or a succession of measurements, you need not run into any problems.

### 1.7.1 *What Is Truly Strange About Quantum Mechanics: Superposition States*

When I discussed electromagnetic waves, I showed that a superposition of plane wave states can result in a radiation pulse. There is nothing unusual about a radiation pulse. However, for quantum-mechanical waves, the superposition of eigenfunctions can sometimes lead to what at first appears to be rather strange results. The strangeness is most readily apparent if we look at the electrons in atoms or other quantum particles that are in a superposition of such *bound* states. Consider the electron in the hydrogen atom. It is possible for the electron to be in a superposition of two or more of its energy states (just as a free particle can be in a superposition of momentum eigenstates). Imagine that you prepare the electron in a superposition of its  $n = 1$  and  $n = 2$  energy states. What does this mean? If you measure the energy of the electron you will get *either*  $-13.6$  eV *or*  $-3.4$  eV. If you prepare the electron in the *same* manner and again measure the energy, you will get *either*  $-13.6$  eV *or*  $-3.4$  eV. After many measurements on identically prepared electrons, you will know the relative probability that the electron will be measured in the  $n = 1$  state or the  $n = 2$  state. You might think that sometimes the electron was prepared in the  $n = 1$  state and sometimes in the  $n = 2$  state, but this is not the case since it is assumed that the electron is always prepared in an *identical* manner which can yield either measurement—that is, it is in a superposition state of the  $n = 1$  and  $n = 2$  states. When you measure the energy, some people (myself excluded) like to say that

the wave function *collapses* into the state corresponding to the energy you measure. I would say that quantum mechanics just allows you to predict the probabilities for various measurements and cannot provide answers to questions regarding how the wave function evolves when a measurement is made.

The idea that the observer *forces* a quantum system into a given state has led to some interesting, but what I consider misleading, ideas. Perhaps the most famous is the question of “Who killed Schrödinger’s cat?” In the “cat” scenario, a cat is put into a box with a radioactive nucleus. If the nucleus decays, it emits a particle that activates a mechanism to release a poisonous gas. Since the nucleus is said to be in a superposition state of having decayed and not having decayed, the cat is also said to be in a superposition state of being dead and alive! Only by looking into the box does the observer know if the nucleus has decayed. In other words, the observation forces the nucleus into either its decayed or undecayed state. Thus it appears that looking into the box can result in the death of the cat.

My feeling is that this is all a lot of nonsense. The reason that it is nonsense is somewhat technical, however. When the nucleus decays, it emits a gamma ray, so the appropriate superposition state must include this radiated field as well. The transition from initial state to the decayed state plus gamma ray essentially occurs *instantaneously*, even though the *time* at which the decay occurs follows statistical laws. Thus when you look into the box, the cat is either alive or it is dead, but not in a superposition state. In other words, the health of the cat is a measure of whether or not the nucleus has decayed, but the cat itself is not in a superposition state. On many similarly prepared systems, you will simply find that the cat dies at different times according to some statistical law. You needn’t worry about being convicted of “catacide” if you open the box.

Note that if it were true that you could force the nucleus into a given state by observing it, you could prevent the nucleus from decaying simply by looking at it continuously! Since it starts in its initial state and you continuously force it to stay in its initial state by looking at it, it never decays. This is known as the Zeno effect, in analogue with Zeno’s paradoxes. It should be pointed out that it *is* possible to keep certain quantum systems in a given state by continuous observation. As long as you make measurements on the system on a time scale that is short compared with the time scale with which the system will evolve into a superposition state, you can keep the system in its initial state. For the decay of particles, however, the transition from initial to final states occurs essentially instantaneously and it is impossible to use continuous “measurements” to keep the particle in its undecayed state. I return to a discussion of the Zeno effect in Chap. 24.

### ***1.7.2 The EPR Paradox and Bell’s Theorem***

Although a single quantum system in a superposition state has no classical analogue and already represents a strange animal, things get *really* strange when we consider two, interacting quantum systems that are prepared in a specific manner.

### 1.7.2.1 The EPR Paradox

In 1935, Einstein, Podolsky, and Rosen (EPR) published a paper entitled *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete*, which has played an important role in the development of quantum mechanics.<sup>5</sup> For an excellent account of this paper and others related to it, see *Quantum Paradoxes and Physical Reality* by Franco Selleri.<sup>6</sup> In their paper, EPR first pose questions as to what constitutes a satisfactory theory: “*Is the theory correct? Is the description given by the theory complete?*” They then go on to define what they mean by an element of physical reality: *If, without in any way disturbing a system, we can predict with certainty (i.e. with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.*” Based on this definition and giving an example in which the wave function corresponds to an eigenfunction of one of two non-commuting operators,<sup>7</sup> they conclude that “*either (1) the quantum-mechanical description of reality given by the wave function is not complete or (2) when the operators corresponding to two physical quantities do not commute the two quantities cannot have simultaneous reality.*”

To illustrate the EPR paradox, one can consider the decay of a particle into two identical particles. The initial particle has no intrinsic spin angular momentum, whereas each of the emitted particles has a spin angular momentum of  $1/2$  (spin angular momentum is discussed in Chap. 12). The spin of the emitted particles can be either “up” or “down” relative to some quantization axis, but the sum of the components of spin relative to this axis must equal zero to conserve angular momentum (the spin of the original particle is zero so the total spin of the composite particles must be zero). The particles are emitted in opposite directions to conserve linear momentum and in a superposition of two states, a state in which one particle (particle *A*) has spin up and the other (particle *B*) has spin down plus a state in which one particle (particle *A*) has spin down and the other (particle *B*) has spin up, written symbolically as

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle), \quad (1.37)$$

where  $\uparrow$  refers to spin up and  $\downarrow$  to spin down. The spin states of the two particles are said to be *correlated* or *entangled*.

Consider the situation in which the particles are emitted in this correlated state and fly off so they are light years apart. According to quantum mechanics, the spin of each of the particles (by “spin,” I now mean either up or down) is not fixed until

<sup>5</sup>A. Einstein, B. Podolsky, and N. Rosen, *Physical Review* **47**, 777–780 (1935).

<sup>6</sup>Franco Selleri, *Quantum Paradoxes and Physical Reality* (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1990).

<sup>7</sup>The commutator of two operators is defined in Chap. 5.

it is measured. In other words, the spin of each of the particles can be either up or down. If you don't measure the spin of particle  $A$ , the spin of particle  $B$  will be either up or down when measured relative to the  $z$ -axis. However, if you measure the spin of particle ( $A$ ) to be up relative to the  $z$ -axis, you are guaranteed that particle ( $B$ ) will have its spin down relative to the  $z$ -axis.

Of course you are free to measure the spin component of particle  $A$  along *any* axis. If you measure the spin of particle  $A$  as up relative to the  $z$ -axis, the spin of particle  $B$  will have physical reality along the  $-z$  direction. On the other hand, if you measure the spin of particle  $A$  as up relative to the  $x$ -axis, the spin of particle  $B$  will have physical reality along the  $-x$  direction. This is a situation in which both states of particle  $B$  correspond to the *same* physical reality and these states correspond to eigenfunctions of non-commuting spin operators, which is not allowed in quantum mechanics. Since this outcome violates condition (2) stated above, the conclusion in EPR is that quantum mechanics is not a complete theory. This is known as the EPR paradox.

In the conventional description of quantum mechanics, there is no paradox. For these correlated states, the probability that you measure a spin down for the second particle after you have measured a spin up for the first particle is unity when both spins are measured relative to the same axis. Quantum mechanics does not tell you *why* this is the case or *how* it occurs, it just gives the probability for the outcome. The idea of one measurement *influencing* the other is not particularly useful or meaningful.

According to EPR, the state represented by Eq. (1.37) cannot represent the complete state of the system. Although it is nowhere mentioned in their paper, one often interprets this to imply that there must be additional labels or *hidden variables* needed in Eq. (1.37). In other words, there are some variables encoded in the quantum system that determine the outcome of the measurements. Quantum mechanics was inconsistent with EPR's element of reality, but a hidden variables theory might be. Which is right?

### 1.7.2.2 Bell's Theorem

The EPR paradox was and continues to be disturbing to some physicists. The idea that a measurement on one physical system can influence the outcome of a measurement on another system that is not causally connected with it can be somewhat unnerving. Motivated by the issues raised by the EPR paradox, John Bell in 1964 published a paper in which he discussed the idea of elements of reality.<sup>8</sup> Without any reference to quantum mechanics, he was able to prove that certain inequalities must be obeyed on measurements of correlated systems to be

---

<sup>8</sup>John Bell, *On the Einstein Podolsky Rosen Paradox*, *Physics* **1**, 195–200 (1964). See also, *Hidden variables and the two theorems of John Bell* by N. David Mermin in the *Review of Modern Physics*, **65**, 803–815 (1993).

consistent with the EPR idea of physical reality. In other words, if Bell's inequalities are violated, then physical observables may not have physical reality of the type described by EPR, independent of whether or not quantum mechanics is a valid theory. It turns out that there are now experiments in which Bell's inequalities are violated. Moreover, the results of these experiments are correctly predicted by quantum mechanics. I will give a proof of Bell's theorem and discuss examples in Chap. 13. Although Bell's theorem is generally accepted, there are some who question its validity.

Measurement in quantum mechanics is a subject that continues to attract a great deal of attention. At this level, maybe it's better not to worry about it too much and instead concentrate on mastering the basic elements of the quantum theory. In the words of Richard Feynman (or not),<sup>9</sup> *Shut up and calculate!*

## 1.8 Summary

In this chapter, I gave a qualitative introduction to several aspects of the quantum theory that we will encounter in the following chapters. As I stated, quantum mechanics is challenging the first go-around for many students. To help master the material it is important to try as many of the problems as possible. It can also help to consult other texts that may treat the material in a different fashion that you find more accessible. To begin, it may prove useful to review some mathematical concepts. You can skip Chap. 2 if you are familiar with these concepts.

## 1.9 Appendix: Blackbody Spectrum

Blackbody radiation is discussed in almost every quantum mechanics textbook, but the manner in which a blackbody achieves thermal equilibrium is almost never discussed. I will return to this question after discussing the equilibrium blackbody spectrum. There are essentially two ways to derive the blackbody spectrum. One employs the quantum statistics of a Bose gas plus some additional assumptions (radiation is described by Bose statistics and the energy of a photon having frequency  $\omega$  is  $\hbar\omega$ ). Since you will see such a treatment in your course on statistical mechanics, I give an alternative derivation, more along the lines given by Planck.

The first step is to get the distribution of radiation field modes in a box, each of whose sides has length  $L$ . There are two ways of doing this, both giving the same final *density of states* of the radiation field.

---

<sup>9</sup>N. D. Mermin, *Could Feynman have said this?*, Physics Today, May, 2004, page 10.

### 1.9.1 Box Normalization with Field Nodes on the Walls

In this case I assume that either the electric or magnetic field has a node on the perfectly conducting walls of a metal cavity which is bounded by  $0 \leq x \leq L$ ,  $0 \leq y \leq L$ ,  $0 \leq z \leq L$ . For the appropriate field to vanish at the boundaries, the field mode function must be of the form  $\sin(k_x x)$ ,  $\sin(k_y y)$ , or  $\sin(k_z z)$ , with

$$k_x = \frac{\pi n_x}{L}; \quad (1.38a)$$

$$k_y = \frac{\pi n_y}{L}; \quad (1.38b)$$

$$k_z = \frac{\pi n_z}{L}, \quad (1.38c)$$

where  $n_x, n_y, n_z$ , are *positive* integers [note that  $\sin(\frac{\pi x}{L})$  and  $\sin(-\frac{\pi x}{L})$  correspond to the same mode, differing only by a phase factor]. The field propagation vector is given by

$$\mathbf{k} = k_x \mathbf{u}_x + k_y \mathbf{u}_y + k_z \mathbf{u}_z, \quad (1.39)$$

where  $\mathbf{u}_j$  ( $j = x, y, z$ ) is a unit vector in the  $j$ -direction. The angular frequency  $\omega$  of a field mode is equal to  $kc$ , where  $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$ .

Thus, it follows from  $\omega = kc$  that

$$\omega_n = k_n c = \frac{\pi c n}{L}, \quad (1.40)$$

where

$$n = \sqrt{n_x^2 + n_y^2 + n_z^2}. \quad (1.41)$$

Each set of values  $(n_x, n_y, n_z)$  corresponds to a given mode of the field. If there are many modes, I can replace the discrete modes by a continuum, essentially making  $n$  a continuous variable with  $\omega = kc = \pi cn/L$ . In other words, the number of modes in a shell having inner radius  $n$  and outer radius  $n + dn$  is simply  $2\frac{1}{8} (4\pi n^2) dn$ . The factor of  $1/8$  is present since only the positive quadrant for  $(n_x, n_y, n_z)$  is allowed, the factor of  $2$  is added to allow for two polarization components for each spatial mode,  $(4\pi n^2)$  is the surface area of the shell, and  $dn$  is the width of the shell. It then follows from Eq. (1.40) that the number of modes having frequency between  $\omega$  and  $\omega + d\omega$  is

$$\mathcal{N}(\omega)d\omega = 2\frac{1}{8} (4\pi n^2) dn = \pi n^2 \frac{dn}{d\omega} d\omega = \frac{L^3 \omega^2 d\omega}{\pi^2 c^3}, \quad (1.42)$$

and the number of modes per unit volume between  $\omega$  and  $\omega + d\omega$ , denoted by  $n(\omega)d\omega$ , is

$$n(\omega)d\omega = \frac{\mathcal{N}(\omega)d\omega}{L^3} = \frac{\omega^2 d\omega}{\pi^2 c^3}. \quad (1.43)$$

### 1.9.2 Periodic Boundary Conditions

In this case I assume that either the electric or magnetic field is subject to *periodic boundary conditions* in which the field values repeat at spatial intervals  $L$  in each direction. Since the field phase of a monochromatic wave varies as  $e^{i\mathbf{k}\cdot\mathbf{r}}$  this implies that

$$k_x = \frac{2\pi n_x}{L}; \quad (1.44a)$$

$$k_y = \frac{2\pi n_y}{L}; \quad (1.44b)$$

$$k_z = \frac{2\pi n_z}{L}, \quad (1.44c)$$

where  $n_x, n_y, n_z$ , are integers, positive, negative or zero [now  $e^{i\frac{2\pi x}{L}}$  and  $e^{-i\frac{2\pi x}{L}}$  correspond to *different* modes]. The calculation proceeds as in box normalization, except that all quadrants are allowed. Thus,

$$\omega = kc = \frac{2\pi cn}{L}, \quad (1.45)$$

the number of modes in a shell having inner radius  $n$  and outer radius  $n + dn$  is  $2(4\pi n^2) dn$  (there is no factor of 1/8 is present since all quadrants for  $(n_x, n_y, n_z)$  are allowed). The number of modes having frequency between  $\omega$  and  $\omega + d\omega$  is

$$\mathcal{N}(\omega)d\omega = 2(4\pi n^2) dn = 8\pi n^2 dn = 8\pi n^2 \frac{dn}{d\omega} d\omega = \frac{L^3 \omega^2 d\omega}{\pi^2 c^3}, \quad (1.46)$$

such that the number of modes per unit volume,  $n(\omega)d\omega$ , between  $\omega$  and  $\omega + d\omega$  is

$$n(\omega)d\omega = \frac{\mathcal{N}(\omega)d\omega}{L^3} = \frac{\omega^2 d\omega}{\pi^2 c^3}, \quad (1.47)$$

in agreement with Eq. (1.43).

### 1.9.3 Rayleigh-Jeans Law

To get the Rayleigh-Jeans law one simply assigns an energy of  $k_B T$  to each mode, resulting in an energy density  $u(\omega)$  (energy per unit volume per unit frequency  $\omega$ ) given by

$$u(\omega) = k_B T \frac{\omega^2}{\pi^2 c^3}. \quad (1.48)$$

Clearly, the energy density diverges at large frequency; this is known as the *ultraviolet catastrophe*.

### 1.9.4 Planck's Solution

Implicit in the equipartition theorem is the assumption that each mode or degree of freedom can have a *continuous* energy distribution. In other words for the radiation modes in a cavity the energy distribution at temperature  $T$ , assumed to be a Boltzmann distribution, is

$$W(E) = \frac{1}{k_B T} e^{-E/k_B T}, \quad (1.49)$$

which has been normalized such that

$$\int_0^\infty W(E) dE = 1. \quad (1.50)$$

You can verify that

$$\bar{E} = \int_0^\infty E W(E) dE = k_B T, \quad (1.51)$$

is *independent of frequency*, for these modes.

To explain the blackbody spectrum, Planck conjectured that the energies of the modes were discrete rather than continuous. A mode having frequency  $\omega$  could have only those energies that are an integral multiple of a constant  $\hbar$  times  $\omega$ . In other words, there is now a *separate* energy distribution for *each* mode having frequency  $\omega$ . The probability of having energy  $n\hbar\omega$  in a mode having frequency  $\omega$  is given by

$$W(n, \omega) = A e^{-n\hbar\omega/k_B T}, \quad (1.52)$$

where  $A$  is a normalization constant that I do not have to specify. The average energy for a mode having frequency  $\omega$  is then

$$\langle W(\omega) \rangle = \frac{A \sum_{n=0}^{\infty} n \hbar \omega e^{-n \hbar \omega / k_B T}}{A \sum_{n=0}^{\infty} e^{-n \hbar \omega / k_B T}} = \hbar \omega \frac{\sum_{n=0}^{\infty} n x^n}{\sum_{n=0}^{\infty} x^n}, \quad (1.53)$$

where

$$x = \exp\left(-\frac{\hbar \omega}{k_B T}\right). \quad (1.54)$$

Using the fact that

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}; \quad (1.55a)$$

$$\sum_{n=0}^{\infty} n x^n = x \frac{d}{dx} \sum_{n=0}^{\infty} x^n = \frac{x}{(1-x)^2}, \quad (1.55b)$$

you can show that

$$\langle W(\omega) \rangle = \frac{x \hbar \omega}{(1-x)} = \frac{\hbar \omega}{\left(\frac{1}{x} - 1\right)} = \frac{\hbar \omega}{\left(e^{\hbar \omega / k_B T} - 1\right)}. \quad (1.56)$$

It follows that  $\langle W(\omega) \rangle \sim k_B T$  for  $\hbar \omega / k_B T \ll 1$ , reproducing the equipartition result, but  $\langle W(\omega) \rangle \sim \hbar \omega e^{-\hbar \omega / k_B T} \ll 1$  for  $\hbar \omega / k_B T \gg 1$ . In essence, Planck ruled out the possibility of exciting high frequency modes. The energy density per unit frequency is obtained by combining Eqs. (1.43) and (1.56),

$$u(\omega) = n(\omega) \langle W(\omega) \rangle = \frac{\hbar \omega^3}{\pi^2 c^3} \frac{1}{\left(e^{\hbar \omega / k_B T} - 1\right)}, \quad (1.57)$$

which is the Planck distribution. If  $\hbar$  is taken to be

$$\hbar = 1.055 \times 10^{-34} \text{ J} \cdot \text{s}, \quad (1.58)$$

Planck's distribution agrees with experiment.

The total energy density is given by

$$\begin{aligned} u &= \int_0^{\infty} \frac{\hbar \omega^3}{\pi^2 c^3} \frac{d\omega}{\left(e^{\hbar \omega / k_B T} - 1\right)} \\ &= \frac{\hbar}{\pi^2 c^3} \left(\frac{k_B T}{\hbar}\right)^4 \int_0^{\infty} \frac{x^3 dx}{(e^x - 1)} = \frac{4\sigma}{c} T^4, \end{aligned} \quad (1.59)$$

which is the *Stefan-Boltzmann Law* and

$$\sigma = \frac{k_B^4}{4\pi^2\hbar^3c^2} \int_0^\infty \frac{x^3 dx}{(e^x - 1)} \quad (1.60)$$

is Stefan's constant. The integral is tabulated

$$\int_0^\infty \frac{x^3 dx}{(e^x - 1)} = \frac{\pi^4}{15}, \quad (1.61)$$

such that

$$\sigma = \frac{\pi^2 k_B^4}{60\hbar^3c^2} = 5.67 \times 10^{-8} \text{W/m}^2 / (\text{°K})^4. \quad (1.62)$$

Equation (1.57) must also give Wien's displacement law; the wavelength corresponding to maximum emission multiplied by the temperature is a constant. The frequency that gives rise to the maximum in the energy density is obtained by setting  $du/d\omega = 0$ , yielding

$$\frac{3\omega_{\max}^2}{(e^{\hbar\omega_{\max}/k_B T} - 1)} - \frac{\hbar}{k_B T} e^{\hbar\omega_{\max}/k_B T} \frac{\omega_{\max}^3}{(e^{\hbar\omega_{\max}/k_B T} - 1)^2} = 0, \quad (1.63)$$

or

$$e^{\hbar\omega_{\max}/k_B T} \omega_{\max} = 3 \frac{k_B T}{\hbar} (e^{\hbar\omega_{\max}/k_B T} - 1); \quad (1.64a)$$

$$\omega_{\max} = 3 \frac{k_B T}{\hbar} (1 - e^{-\hbar\omega_{\max}/k_B T}). \quad (1.64b)$$

Setting

$$y = \frac{\hbar\omega_{\max}}{k_B T}, \quad (1.65)$$

I find that the maximum occurs for

$$y = 3(1 - e^{-y}), \quad (1.66)$$

which can be solved graphically to obtain  $y = 2.82$ , leading to

$$\hbar\omega_{\max} = 2.82k_B T; \quad f_{\max} = \frac{2.82k_B T}{h} = 5.88 \times 10^{10} T \text{ Hz/°K}. \quad (1.67)$$

If I set  $\lambda_{\max} = c/f_{\max}$ , I find

$$\lambda_{\max} T = \frac{cT}{f_{\max}} = 5.10 \times 10^{-3} \text{ m} \cdot ^\circ \text{K}, \quad (1.68)$$

which is not quite Wien's law,

$$\lambda_{\max} T = 2.9 \times 10^{-3} \text{ m} \cdot ^\circ \text{K}. \quad (1.69)$$

The reason for the difference is that Wien's law is derived from the energy density *per unit wavelength*  $w(\lambda)$  rather than energy density per unit frequency and the two methods give different maxima. That is, if you measure the energy density per unit frequency as a function of frequency and the energy density per unit wavelength as a function of wavelength,  $f_{\max} \lambda_{\max} \neq c$ . To see this I set

$$u(\omega)d\omega = -w(\lambda)d\lambda. \quad (1.70)$$

Since  $d\omega = -2\pi c d\lambda/\lambda^2$ , I find

$$\begin{aligned} w(\lambda) &= \frac{2\pi c}{\lambda^2} u\left(\frac{2\pi c}{\lambda}\right) \\ &= \frac{2\pi c}{\lambda^2} \frac{\hbar \left(\frac{2\pi c}{\lambda}\right)^3}{\pi^2 c^3} \frac{1}{(e^{hc/\lambda k_B T} - 1)} \\ &= \frac{8\pi hc}{\lambda^5} \frac{1}{(e^{hc/\lambda k_B T} - 1)}. \end{aligned} \quad (1.71)$$

Now, instead of Eq. (1.66), I must solve

$$y = 5(1 - e^{-y}). \quad (1.72)$$

The solution is  $y = hc/\lambda_{\max} k_B T = 4.965$ , such that

$$\lambda_{\max} T = \frac{hc}{4.965 k_B} = 2.90 \times 10^{-3} \text{ m} \cdot \text{K}, \quad (1.73)$$

which is Wien's law.

### 1.9.5 Approach to Equilibrium

Let me now return to the question of the approach to equilibrium. Consider first a box with two point particles, one at rest somewhere in the box and one having speed  $v_0$  and energy  $E_0 = mv_0^2/2$  located at some other position in the box. The

collisions between particles and between the particles and the walls are assumed to be perfectly elastic. You might think that, on average, each particle will have energy  $E_0/2$  if you wait long enough, but this is not true. There are many initial conditions that will not result in collisions between the particles. For a given set of initial conditions, the dynamics is perfectly determined.

As the number of particles in the box increases to a very large number, it becomes more and more likely that all the atoms undergo collisions at a rapid rate. In this case, the equilibrium distribution is Maxwellian. Returning to the modes of the radiation field in the cavity, there is no obvious way they can exchange energy to achieve equilibrium. They must exchange energy with the charges in the cavity walls. Thus one is faced with modeling this interaction and then having some model for the energy distribution of the charges in the cavity. As far as I know, no one has ever solved this approach to equilibrium in a satisfactory manner.

## 1.10 Problems

1. What is the ultraviolet catastrophe and how did it lead to Planck's quantum hypothesis? Specifically how did theory and experiment differ in describing the spectrum of a blackbody? What hypothesis did Planck make to minimize the contribution of the high frequency modes?
2. Describe the photoelectric effect experiments and Einstein's explanation for both the number and energy of the emitted electrons as a function of the frequency of the incident light.
3. What were Bohr's postulates in his theory of the hydrogen atom? How do these postulates explain the spectrum and stability of the hydrogen atom?
4. Draw an energy level diagram for the hydrogen atom and indicate on it the energy of the four lowest energy states. Also indicate the radius, electron velocity, and angular momentum of these states as given by the Bohr theory.
5. Calculate the wavelength of the  $n = 7$  to  $n = 2$  and of the  $n = 3$  to  $n = 2$  transitions in hydrogen. What frequency of radiation is needed to excite a hydrogen atom from its  $n = 1$  to  $n = 3$  state? How much energy is required to ionize a hydrogen atom from its  $n = 2$  state?
6. Radiation from the  $n = 2$  to  $n = 1$  state of hydrogen is incident on a metal having a work function of 2.4 eV. What is the maximum energy of the electrons emitted from the metal?
7. What is the significance of the equation  $\lambda_{dB} = h/p$ ? What is meant by the wave particle duality of matter? What determines when matter acts as a particle and when it acts as a wave?

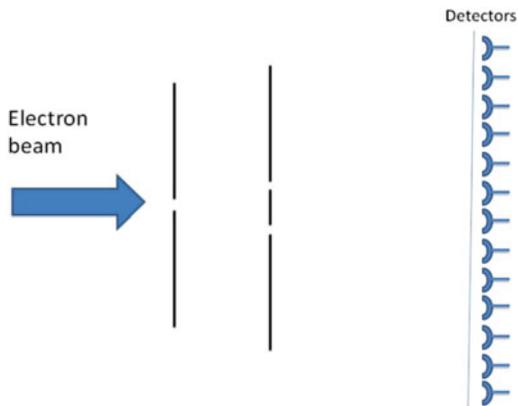


Fig. 1.4 Problem 1.10

8. Calculate the de Broglie wavelength for a particle of mass 1.0 g moving with a speed of 1.0 cm/yr. Calculate the de Broglie wavelength for the electron in the  $n = 1$  state of the Bohr atom.
9. In general terms, discuss the measurement process in quantum mechanics. Why is it necessary to make measurements on a large number of identically prepared systems to obtain  $|\psi(\mathbf{r}, t)|^2$ ? How does this differ from Newtonian mechanics? Why is a single particle in a superposition state an intrinsically quantum object?
10. In a two-slit experiment, particles are sent into the apparatus one particle at a time (see Fig. 1.4). How many detectors does a single particle trigger? If the experiment is repeated many times, what will a graph of the number of counts in a detector vs detector position look like? Explain.
11. When light is incident on a glass slab, some of the light is reflected. This is a wave-like phenomenon (if a classical particle encounters a change in potential, it simply slows down or speeds up with no reflection), even though this corresponds to a geometrical optics limit (neglect of diffraction). Why does a wave-like effect occur in this case? Is there any connection of this result with the rainbow?
12. The blackbody spectrum as a function of frequency  $u_f(f)$  can be obtained using  $u_f(f)df = u(\omega)d\omega = 2\pi u(2\pi f)df$ . Plot  $[cu_f(f)/4\pi] \times 10^{18}$  as a function of frequency for  $T = 2.73^\circ\text{K}$  and find the maximum frequency [ $cu_f(f)/4\pi$  is the power per unit area per unit frequency per unit solid angle—this corresponds to the flux incident on a detector per unit frequency per steradian (sr)]. This Planck distribution corresponds to the *cosmic microwave background*. What is the energy per unit volume of the cosmic microwave background?

## Extra Reading

There are many excellent texts on quantum mechanics. My suggestion is to go to your library and try to find texts that you find especially helpful. Some possible recommendations are:

### *Undergraduate Texts*

David Bohm, *Quantum Theory* (Prentiss Hall, New York., 1951).

Ashok Das and Adrian Melissinos, *Quantum Mechanics – A Modern Introduction*, (Gordan and Breach, New York, 1986).

Peter Fong, *Elementary Quantum Mechanics*, Expanded Edition (World Scientific, Hackensack, N.J., 2005).

Stephen Gasiorowicz, *Quantum Physics*, Second Edition (John Wiley and Sons, New York, 1996).

David Griffiths, *Introduction to Quantum Mechanics*, Second Edition (Pearson Prentiss Hall, Upper Saddle River, N.J., 2005).

Richard Liboff, *Introductory Quantum Mechanics* 4th Edition (Addison Wesley, San Francisco, 2003).

David Park: *Introduction to the Quantum Theory*, Second Edition (McGraw Hill, New York, 1974).

John Powell and Bernd Crasemann, *Quantum Mechanics* (Addison Wesley, Reading, MA, 1961).

David Saxon, *Elementary Quantum Mechanics*, (Holden-Day, San Francisco, 1968).

George Trigg, *Quantum Mechanics*, (Van Nostarnd, Princeton, N.J., 1968).

### *Graduate Texts*

Ernst Abers, *Quantum Mechanics* (Pearson Education, Upper Saddle River, N.J., 2004).

Claude Cohen-Tannoudji, Bernard Liu, and Franck Laloë, *Quantum Mechanics* (Wiley Interscience, Paris, 1977).

Robert Dicke and James Wittke, *Introduction to Quantum Mechanics*, (Addison Wesley, Reading, MA, 1960).

P. A. M. Dirac, *The Principles of Quantum Mechanics*, Fourth Edition (Oxford University Press, London, 1958).

Eugen Merzbacher, *Quantum Mechanics*, Third Edition (John Wiley and Sons, New York, 1998).

Albert Messiah, *Quantum Mechanics*, Vols. 1 and 2, (Dover Publications, New York, 2014).

Leonard Schiff, *Quantum Mechanics*, Third Edition (McGraw Hill, New York, 1968).