Chapter 7
From Classical Mechanics to Quantum Mechanics

7.1 Introduction

The chapter tackles the difficult problem of bridging the concepts of Classical Mechanics and Electromagnetism with those of Quantum Mechanics. The subject, which is fascinating *per se*, is illustrated within a historical perspective, covering the years from 1900, when Planck's solution of the black-body radiation was given, to 1926, when Schrödinger's paper was published.

At the end of the 1800s, the main branches of physics (mechanics, thermodynamics, kinetic theory, optics, electromagnetic theory) had been established firmly. The ability of the physical theories to interpret the experiments was such, that many believed that all the important laws of physics had been discovered: the task of physicists in the future years would be that of clarifying the details and improving the experimental methods. Fortunately, it was not so: the elaboration of the theoretical aspects and the refinement in the experimental techniques showed that the existing physical laws were unable to explain the outcome of some experiments, and could not be adjusted to incorporate the new experimental findings. In some cases, the theoretical formulations themselves led to paradoxes: a famous example is the *Gibbs entropy paradox* [70]. It was then necessary to elaborate new ideas, that eventually produced a consistent body generally referred to as *modern physics*. The elaboration stemming from the investigation of the microscopic particles led to the development of Quantum Mechanics, that stemming from investigations on high-velocity dynamics led to Special Relativity.

The chapter starts with the illustration of the planetary model of the atom, showing that the model is able to justify a number of experimental findings; this is followed by the description of experiments that can not be justified in full by the physical theories existing in the late 1800s: stability of the atoms, spectral lines of excited atoms, photoelectric effect, spectrum of the black-body radiation, Compton effect. The solutions that were proposed to explain such phenomena are then illustrated; they proved to be correct, although at the time they were suggested a comprehensive theory was still lacking. This part is concluded by a heuristic derivation of the time-independent Schrödinger equation, based upon the analogy between the variational principles of Mechanics and Geometrical Optics.

© Springer Science+Business Media New York 2015
M. Rudan, *Physics of Semiconductor Devices*,
DOI 10.1007/978-1-4939-1151-6_7

In the final part of the chapter the meaning of the wave function is given: for this, an analysis of the measuring process is carried out first, showing the necessity of describing the statistical distribution of the measured values of dynamical quantities when microscopic particles are dealt with; the connection with the similar situations involving massive bodies is also analyzed in detail. The chapter is concluded with the illustration of the probabilistic interpretation of the wave function.

7.2 Planetary Model of the Atom

Several experiments were carried out in the late 1800s and early 1900s, whose outcome was the determination of a number of fundamental constants of atomic physics; among them, the *electron charge-to-mass ratio* was measured by J. J. Thomson in 1897, and the *electron charge* was measured by R. Millikan in 1909 (Table D.1). A theory of the atom was proposed by E. Rutherford after a series of experiments in 1909–1914, that led to the measurement of the atomic radius r_a. The experiments consisted in measuring the broadening of beams of finely collimated α particles[1] passing through thin metal foils. The latter were typically made of gold sheets with a thickness of a few thousand atomic layers; the dynamics of the interaction between an α particle and the metal foil was treated by Classical Mechanics, using an interaction of the Coulomb type (Sect. 3.8). The outcome of the experiments led Rutherford to conceive the *planetary model* of the atom; this model depicts the atom as made of a small nucleus of atomic charge $Z\,q$ surrounded by Z electrons, where q is the absolute value of the electron charge and Z indicates the position of the element in the periodic table. The model assumes that, as the α-particles are rather heavy, they are deflected mainly by the nuclei[2] of the foil; the type of deflection implies that the majority of the α particles is deflected only once when crossing the foil, this indicating that the foil's atoms are placed far apart from each other. In fact, the deflection experiments made it possible to estimate[3] the atom's and nucleus' diameters, respectively, as

$$r_a \approx 0.1\ \text{nm}, \qquad r_e \approx 2\times 10^{-6}\sqrt{Z}\ \text{nm}, \qquad Z = 1, 2, \dots\ . \tag{7.1}$$

As Z ranges from 1 to about 100, the second relation in (7.1) shows that $r_e \ll r_a$ (compare with Table D.1).

The simplest atom is that of hydrogen. The planetary model depicts it as an electron moving nearby the proton under the effect of a Coulomb potential V, where the energy reference is chosen in such a way as to make $V(\infty) = 0$. Then, the theory

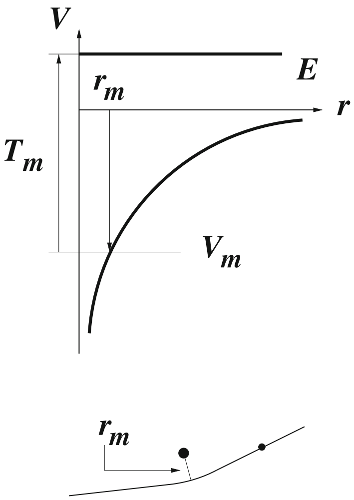
---

[1] These particles are obtained by ionizing helium atoms.

[2] The meaning of term "nucleus" in this context is clarified in Sect. 7.8.1.

[3] The estimate means that for scale lengths equal or larger than those indicated in (7.1), an atom or a nucleus can be considered as geometrical points having no internal structure. The electron's radius can be determined in a similar way using X-ray diffusion.

Fig. 7.1 Classical description of the electron's orbit for $E \geq 0$



of the Coulomb interaction in the attractive case applies (Sect. 3.13.6). As the proton is much more massive than the electron, its position can be approximated with that of the atom's center of mass, and placed in the origin; in summary, it is

$$V(\mathbf{r}) = V(r) = -\frac{q^2}{4\pi\,\varepsilon_0\,r}, \qquad T + V = \frac{1}{2}mu^2 - \frac{q^2}{4\pi\,\varepsilon_0\,r} = E = \text{const}, \quad (7.2)$$

where $\mathbf{r}$ is the electron position, u its velocity's module, T and E its kinetic and total energies, respectively, and ε_0 the permittivity of vacuum. The planetary model is extended to more complicated atoms by considering an outer electron moving nearby a core of net charge q embedding Z protons and $Z - 1$ electrons, or to even more complicated cases (*hydrogenic-like systems*). Observing that $T = E - V = E + |V| \geq 0$, and remembering the analysis of Sect. 3.8, one finds that two cases are possible: the first one, shown in Fig. 7.1, is $E \geq 0$, corresponding to $V \leq 0 \leq E$ and $r_{\max} = \infty$ (*free electron*). The second case is $E < 0$, corresponding to $V \leq E < 0$ and $r_{\max} < \infty$ (*bound electron*). For the qualitative reasoning to be carried out here, it is sufficient to consider the simpler case of a bound electron whose trajectory is circular (Fig. 7.2). In such a case, using the results of Sects. 3.7 and 3.13.6 after letting $Z_1 = Z_2 = 1$ and replacing s with r, yields $r = 4\pi\,\varepsilon_0\,M_B^2/(m\,q^2) = \text{const}$, where $M_B^2 = m^2\,r^2\,u^2$. Combining these relations with the first one in (7.2) shows that $F = |\mathbf{F}| = m\,|\mathbf{a}| = m\,u^2/r = 2\,T/r$ whence, using $F = q^2/(4\pi\,\varepsilon_0\,r^2) = -V/r$, one finds

$$T = -\frac{V}{2}, \qquad E = T + V = \frac{V}{2} = -\frac{q^2}{8\pi\,\varepsilon_0\,r} = \text{const} < 0. \quad (7.3)$$

It follows $\mathrm{d}E/\mathrm{d}r = |E|/r > 0$, that is, the total energy is larger at larger orbits (this is a particular case of the general theory worked out in Prob. 3.2).

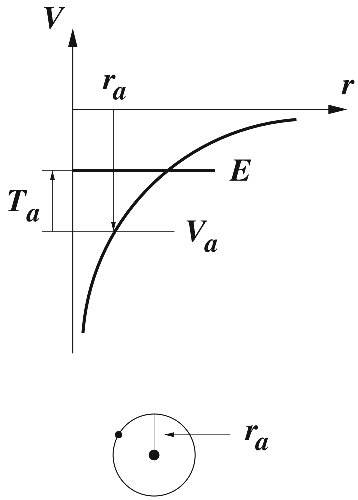


Fig. 7.2 Classical description of the electron's orbit for $E < 0$. For simplicity, a circular orbit is considered

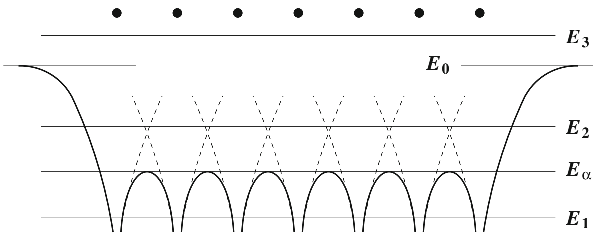
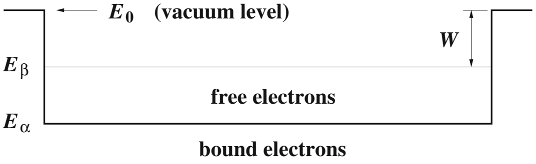


Fig. 7.3 Schematic description of the potential energy in a linear monatomic chain

Despite its simplicity, the planetary model is able to explain phenomena like the excitation and ionization of atoms; *excitation* corresponds to the absorption of energy from an external electromagnetic field, such that an initially-bound electron increases its energy from E_1 to E_2, where $E_1 < E_2 < 0$: the electron in the final state is still bound. The inverse process is the emission of energy in form of an electromagnetic radiation, so that the electron's total energy decreases from E_2 to E_1. In turn, *ionization* corresponds to the absorption of energy such that an initially-bound electron becomes free: $E_1 < 0$ and $E_2 \geq 0$. The inverse process is the capture of a free electron by an ionized atom, with an energy emission equal to $E_2 - E_1$.

The above reasoning can also be used to explain the behavior of systems more complicate than single atoms. For instance, consider a finite *linear monatomic chain*, namely, a system made of a finite number of identical atoms placed along a line, at equal mutual distances (Fig. 7.3), which can be thought of as a rudimental version of a crystal. The positions of the atoms are marked by the dots visible in the upper part of the figure. Let the chain be aligned with the x axis; if each nucleus is made to coincide with a local origin of the reference, the distance $r = \sqrt{x^2 + y^2 + z^2}$ in the expression of the potential energy becomes $|x|$; the potential energy pertaining to each nucleus is proportional to $1/|x|$ and is indicated with a dashed line in the figure. An electron placed at some position in the chain is subjected to the sum of

Fig. 7.4 The same structure of Fig. 7.3, where the peaks are replaced with the envelope



the potential energies; being the latter negative, the sum is lower than the individual contributions: in the figure, it is indicated by the continuous line, which for simplicity is drawn by adding up the contributions of the nearest-neighboring nuclei only. The chain, however, has a finite length; when the leftmost nucleus is considered, the potential energy on its left does not suffer any lowering: this creates the energy step visible in the figure. The same happens on the right side of the righmost nucleus. The shape of the potential energy thus obtained is able to qualitatively explain several features of crystals. For instance, consider the case where the only force acting on an electron inside the crystal derives from the potential energy of Fig. 7.3, namely, it is a conservative force. If the total energy of the electron under consideration is E_1, the electron's position is confined within the potential well dictated by the initial position of its motion. Thus, the electron oscillates within the well like in the example of Prob. 3.2, and its motion can not extend out of it; if all electrons of the crystal are bound, the material is an insulator. This reasoning implies that the situation where all electrons are bound is maintained also under the application of an external voltage; due to this, no electric current ensues.

If the total energy of the electron under consideration is E_2, the electron can move within the whole crystal; finally, if the total energy is E_3, the electron overcomes one or the other of the energy steps and moves into vacuum: for this reason, the minimum energy E_0 necessary for the electron to leave the crystal is called *vacuum level*. If the two ends of the crystal are connected to a voltage generator by suitable contacts, and an external voltage is applied, the material can carry an electric current, whose amplitude depends also on the number of electrons whose energy is sufficiently high.[4] It is worth noting that, although this is prohibited in the frame of Classical Mechanics, the electrons whose energy is of the type E_2 may also contribute to the current; in fact, Quantum Mechanics shows that they have a finite probability to penetrate the energy step and reach the contact. This phenomenon is called *tunnel effect* (Sect. 11.3.1).

To proceed it is convenient to give Fig. 7.3 a simpler appearance: in fact, considering that the interatomic distance is a fraction of a nanometer, the spatial extent in the x direction where each peak is placed is hardly visible in a macroscopic representation; for this reason, it is sufficient to graphically indicate the envelope E_α of the peaks. By the same token, the steps on the sides are described as discontinuities (Fig. 7.4). The electrons with $E < E_\alpha$ or $E \geq E_\alpha$ are called, respectively, *bound*

---

[4] The combination of the number of such electrons with other factors also determines whether the material is a conductor or a semiconductor (Chap. 18).

electrons and *free electrons*.[5] In the equilibrium condition the total energy of the electrons is prescribed; it follows that the majority of the electrons has an energy E lower than a given value E_β (which is not necessarily larger than E_α). Thus, the difference $W = E_0 - E_\beta$ is a measure of the energy that is necessary to extract an electron from the material: among other things, the model explains the existence of a minimum extraction energy of the electrons.[6]

7.3 Experiments Contradicting the Classical Laws

About 1900, experimental evidence was found for a number of phenomena that contradict the calculations based on the known physical laws of the time, that is, the laws of Analytical Mechanics, Electromagnetism, and Statistical Mechanics. A number of such phenomena are listed in this section.

Stability of the Atom

The solution of the electromagnetic equations shows that an accelerated electron radiates a power given by (5.72), namely, $q^2 \dot{v}^2/(6\pi\,\varepsilon_0\,c^3)$, with $\dot{v}$ the electron's acceleration. As discussed in Sect. 5.11.2, this is in contradiction with the planetary model of the atom (7.3), in which $T + V = E = \text{const}$: due to the radiated power, the electron should lose energy and, as a consequence, the atom should shrink. The possibility of an extremely slow, non-detectable shrinking must be ruled out: in fact, a perturbative calculation (Sect. 5.11.3) shows that due to the energy loss the atomic radius should decrease from the initial value, say, r, to the value r/e in about 10^{-8} s. This, however, is not observed (Sect. 9.7.2).

Spectral Lines of Excited Atoms

The planetary model explains the emission of electromagnetic radiation by excited atoms. The explanation, however, is qualitative only, because the model does not impose any constraint on the frequency of the emitted waves. In contrast, the experiments show that the waves emitted by, e.g., hydrogen atoms have frequencies ν of the form (*Balmer law*, 1885),[7]

$$\nu_{nm} = \nu_R \left(\frac{1}{n^2} - \frac{1}{m^2}\right), \qquad \nu_R \simeq 3.3 \times 10^{15}\ \mathrm{s}^{-1}, \tag{7.4}$$

---

[5] The description is qualitative; for instance, it does not consider the band structure of the solid (Sect. 17.6).

[6] When the material is a conductor, E_β coincides with the Fermi level (Sect. 15.8.1), and W is called *work function*; in a semiconductor, E_β coindices with the lower edge E_C of the conduction band (Sect. 17.6.5) and the minimum extraction energy (typically indicated with a symbol different from W) is called *electron affinity* (Sect. 22.2).

[7] The ratio $R = \nu_R/c \simeq 1.1 \times 10^5\ \mathrm{cm}^{-1}$ is called *Rydberg constant*. The formula was generalized in the 1880s to the hydrogenic-like atoms by Rydberg: the expression (7.4) of the frequencies must be multiplied by a constant that depends on the atom under consideration.

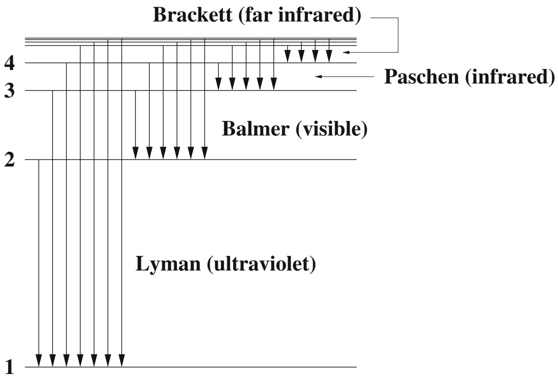


Fig. 7.5 Designation of the lower series of spectral lines (7.4)

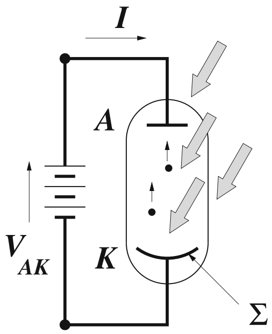


Fig. 7.6 Schematic cross-section of the apparatus used for measuring the photoelectric effect

where n, m are integers, $m > n \geq 1$. The emissions described by (7.4) are also called *spectral lines*. The lower series of spectral lines are shown in Fig. 7.5 along with their designations; the numbers in the figure correspond to n, m in (7.4). Another experimental finding shows that, instead of occurring with a single emission of frequency ν_{nm}, the release of electromagnetic energy by the atom may be accomplished in steps; if that happens, the frequencies associated to the individual steps fulfill a relation called *Ritz emission rule*: considering, e.g., two steps, it reads

$$\nu_{nm} = \nu_{nk} + \nu_{km}, \tag{7.5}$$

with ν_{nk}, ν_{km} the frequencies of the individual steps.

Photoelectric Effect

It is found that an impinging electromagnetic radiation extracts charges from a metal (H. Hertz, 1887) and that these charges are electrons (J. J. Thomson, 1899). The phenomenon is ascribed to the absorption of energy from the radiation: the electron absorbs an energy sufficiently large to be extracted from the metal. An electron thus extracted is also called *photoelectron*. A sketch of the measuring apparatus is given in Fig. 7.6, where two electrodes, anode (A) and cathode (K), are placed inside a vacuum

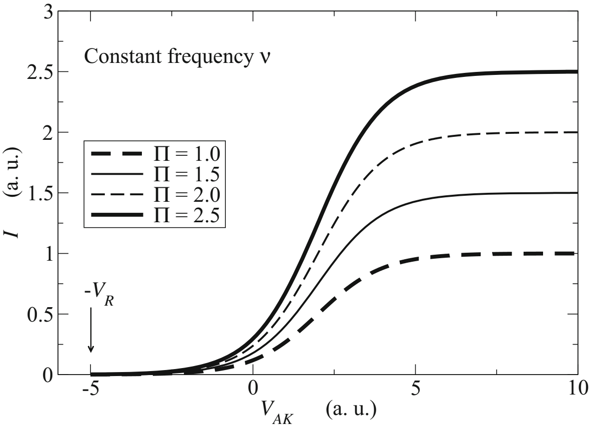


Fig. 7.7 The $I = I(V_{AK})$ curves, in arbitrary units, obtained from the photoelectric effect at constant frequency of the radiation, with the spectral power used as a parameter

tube in order to prevent interactions between the photoelectrons and the atmosphere. A voltage V_{AK} is applied to the electrodes, such that $V_{AK} > 0$ when the electric potential at the anode is larger than that at the cathode. A monochromatic radiation of a given intensity is made to impinge on the cathode, whose surface is marked with Σ, and the current I flowing in the tube is recorded. Important parameters are the radiation's frequency ν, the spectral intensity of the radiation, $\eta = \mathrm{d}E/(\mathrm{d}\Sigma\,\mathrm{d}t\,\mathrm{d}\nu)$, where $\mathrm{d}\Sigma$ is the surface element of the cathode, and the spectral power[8]

$$\Pi = \int_\Sigma \eta\,\mathrm{d}\Sigma = \frac{\mathrm{d}E}{\mathrm{d}t\,\mathrm{d}\nu}. \tag{7.6}$$

The outcome of the experiment is shown in arbitrary units in Figs. 7.7 and 7.8. The first one shows a set of the $I = I(V_{AK})$ curves at constant ν, with Π a parameter. When V_{AK} is positive and sufficiently high, it is expected that practically all electrons extracted from the cathode be driven to the anode; as a consequence, the slope of the curves should be negligible. Also, when the intensity of the radiation increases, the number of extracted electrons, and the current with it, should also increase. This is in fact confirmed by the curves of Fig. 7.7. When, instead, V_{AK} is negative, only the electrons leaving the cathode with a sufficiently high kinetic energy are able to reach the anode, whereas those whose initial kinetic energy is low are repelled towards the cathode by the electric field imposed by the reverse bias.[9] Considering for simplicity a one-dimensional case, energy conservation yields for an electron traveling from cathode to anode,

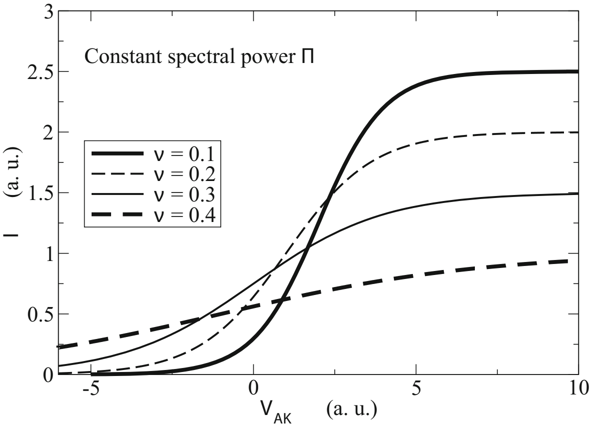
$$\frac{1}{2}m\,u_A^2 - \frac{1}{2}m\,u_K^2 = q\,V_{AK}, \tag{7.7}$$

---

[8] The units of η and Π are $[\eta] = \mathrm{J\,cm^{-2}}$ and $[\Pi] = \mathrm{J}$, respectively.

[9] The concentration of electrons in the vacuum tube is small enough not to influence the electric field; thus, the latter is due only to the value of V_{AK} and to the form of the electrodes.

Fig. 7.8 The $I = I(V_{AK})$ curves, in arbitrary units, obtained from the photoelectric effect at constant spectral power of the radiation, with frequency used as a parameter



where u_K is the electron's velocity at the cathode and u_A that at the anode. The *blocking voltage* $V_R > 0$ is the value $V_{AK} = -V_R$ such that $u_A = 0$; from (7.7) one obtains the relation

$$\frac{1}{2} m\, u_K^2 = q\, V_R, \tag{7.8}$$

which allows one to measure the kinetic energy of the most energetic electrons that are extracted from the cathode at given spectral power and frequency of the radiation.[10] Such electrons are those that inside the cathode have an energy in the vicinity of E_β (Fig. 7.4) and do not suffer energy losses while being extracted. If E_L is the energy that the most energetic electron absorbs from the radiation, its kinetic energy at the cathode is $(1/2)\, m\, u_K^2 = E_L - W$, with W the metal's work function, whence

$$q\, V_R = E_L - W, \tag{7.9}$$

so that the photoelectric effect provides in fact a method for measuring E_L. The classical model predicts that the blocking voltage should increase with Π; this, however, does not happen: as shown in Fig. 7.7, at a given frequency the blocking voltage is the same for all values of Π.

In addition, it is unexpectedly found that both I and V_R depend on the frequency ν (Fig. 7.8). In fact, the comparison between the experimental blocking voltages and (7.9) shows that the energy E_L that the electron absorbs from the electromagnetic field is proportional to the frequency,

$$E_L = h\, \nu, \tag{7.10}$$

with $h \simeq 6.626 \times 10^{-34}$ J s the *Planck constant*. If $h\, \nu < W$, no current is measured; this provides a threshold value for the frequency to be used in the experiment.

---

[10] The most energetic electrons succeed in overcoming the effect of the reverse bias and reach the vicinity of the anode; they constantly slow down along the trajectory, to the point that their velocity at the anode vanishes. Then, their motion reverses and they are driven back to the cathode.

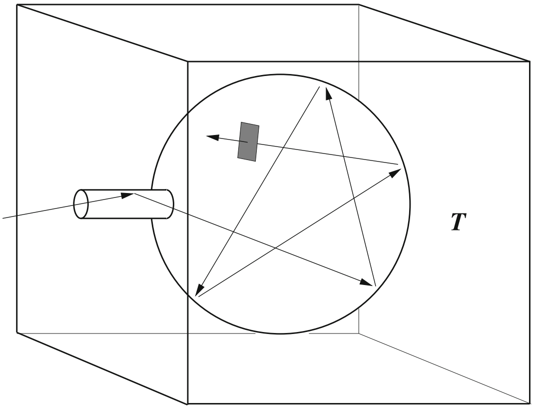


Fig. 7.9 The approximation to a black body consisting in a small hole in the wall of an enclosure kept at constant temperature. If a thermometer (represented by the *shaded area*) were suspended within the enclosure, it would indicate the same temperature T as the walls, irrespective of its position or orientation

Spectrum of the Black-Body Radiation

Consider a body at temperature T in equilibrium with an electromagnetic field. Due to the detailed-balance principle, the spectral intensity η_B emitted by the body, that is, the electromagnetic power emitted by it per unit surface element $\mathrm{d}\Sigma$ and unit frequency $\mathrm{d}\nu$, in the direction normal to $\mathrm{d}\Sigma$, fulfills the relation

$$\eta_B = \alpha\, \eta, \tag{7.11}$$

where η is the spectral intensity of the radiation (compare with (7.6)), and $0 \le \alpha \le 1$ the fraction of η absorbed by the body at frequency ν. By Kirchhoff's law (1859), for any body in thermal equilibrium with radiation it is

$$\frac{\eta_B}{\alpha} = K(\nu, T), \tag{7.12}$$

where K is a universal function of ν and T [85, Sect. 9–15]. A *black body* is a body such that $\alpha = 1$ at all frequencies; thus, for a black body at equilibrium with radiation it is $\eta_B = K$. A good approximation to a black body is a small hole in the wall of an enclosure kept at constant temperature, like that illustrated in Fig. 7.9: any radiation entering the hole has a negligible probability of escaping, due to multiple reflections at the walls; as a consequence, the hole acts like a perfect absorber. Thanks to $\eta_B = K$, the spectral intensity emitted by any black body has the same characteristics: in particular, it is not influenced by the form of the enclosure, the material of which the walls are made, or other bodies present in the enclosure. As a consequence, η_B, or any other function related to it, can be calculated by considering a convenient geometry of the problem and assuming that the radiation propagates *in vacuo*.

To proceed, consider a surface element $\mathrm{d}\Sigma$ of the black body, and connect a local Cartesian reference to it such that $\mathrm{d}x$ and $\mathrm{d}y$ belong to $\mathrm{d}\Sigma$; it follows

$$\eta_B = \frac{\mathrm{d}E}{\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}t\,\mathrm{d}\nu} = c\,\frac{\mathrm{d}E}{\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}\nu} = c\,u, \tag{7.13}$$

where u is the *spectral energy density* of the black body, that is, the energy per unit volume and frequency. The integral of u over the frequencies yields the energy density; remembering that equilibrium is assumed, one finds[11]

$$w_{\rm em}^{\rm eq}(T) = \int_0^\infty u(\nu, T)\ {\rm d}\nu. \tag{7.14}$$

In turn, the integral of u over the coordinates gives the equilibrium value of the spectral energy, whose general definition is given by (5.47). It is found experimentally (Stefan's law, 1879) that

$$\int_0^\infty \eta_B(\nu, T)\,{\rm d}\nu = \sigma\, T^4, \tag{7.15}$$

where $\sigma = 5.67 \times 10^{-12}\ {\rm W\,cm^{-2}\,K^{-4}}$ is the Stefan-Boltzmann constant. Combining (7.15) with (7.13) and (7.14) yields $w_{\rm em}^{\rm eq}(T) = \sigma\, T^4/c$.

The spectral energy density u can be calculated as the product of the number of monochromatic components of the electromagnetic field per unit volume and frequency, times the energy of each monochromatic component. The first factor is readily found by taking an enclosure of prismatic form like that of Sect. 15.9.4; the calculation yields $8\,\pi\,\nu^2/c^3$, which is obtained by dividing both sides of (15.74) by the enclosure's volume V. As for the energy of each monochromatic component, the only assumption possible in the frame of Classical Mechanics is that the energy of the electromagnetic field at equilibrium is distributed over the frequencies according to the Maxwell–Bolzmann distribution (6.14). Assuming that each monochromatic component is equivalent to a one-dimensional linear-harmonic oscillator, the energy to be associated to it is the average energy of a system with one degree of freedom; thus, letting $R = 1$ in (6.37), yields for the average energy the value $k_B\,T$. The product of the two factors thus found yields for the spectral energy density of the black body the expression

$$u(\nu, T) = 8\,\pi\,\frac{k_B T}{c^3}\,\nu^2, \tag{7.16}$$

called *Rayleigh-Jeans law*. Experimental results for u as a function of frequency are shown in Fig. 7.10, with temperature a parameter. The comparison with experiments shows that the parabolic behavior of (7.16) approximates the correct form of the curves only at low frequencies; clearly the result expressed by (7.16) can not be correct, because it makes the equilibrium energy density (7.14) to diverge.[12]

---

[11] Compare with the general definition (5.10) of $w_{\rm em}$, where the assumption of equilibrium is not made.

[12] This unphysical outcome is also called *ultraviolet catastrophe*.

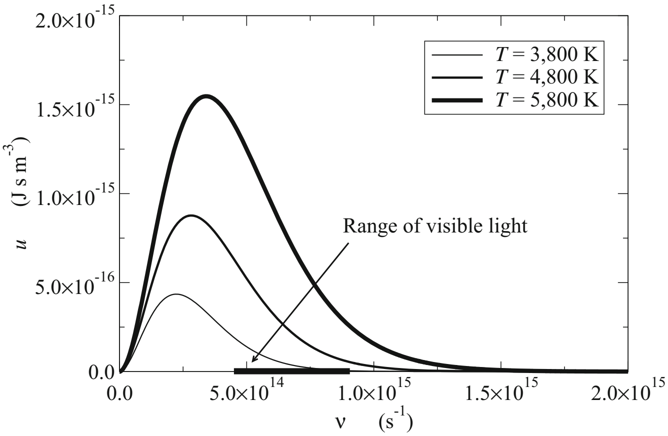


Fig. 7.10 Spectral energy density of the black body at different temperatures. The value $T = 5800$ K corresponds to the surface temperature of the sun

Compton Effect

When X-rays of a given frequency interact with atoms and are scattered with an angle ψ with respect to the direction of incidence, the frequency of the emerging rays is found to depend on ψ. This outcome is in contrast with the prediction of the electromagnetic theory, according to which the frequency of the scattered rays should be equal to that of the impinging ones. The dependence of frequency on the scattering angle is also called *Compton effect*.

The experimental setup for measuring the Compton effect is schematically shown in Fig. 7.11. The gray box in the middle of the figure is a piece of solid material, onto which the radiation impinges from the left (dark arrows); the vertical lines are the intersections of the constant-phase planes with the plane of the figure. The gray arrows on the right represent the part of the radiation that does not interact with the material and exits from it unaltered, while the white arrows indicate some of the directions of the rays scattered by the material. The circumferences are the intersections with the figure's plane of the spherical waves produced by the scattering. The origin of the measuring angle is aligned with the direction of the incoming radiation, so that $\psi = 0$ corresponds to the absence of scattering, $\psi = 2\pi$ to reflection.

7.4 Quantum Hypotheses

In the early 1900s, a number of hypotheses were made to solve the contradictions between the experimental evidence and the calculations based on the physical laws known at that time. The solutions thus found and the new concepts originating from them were eventually combined into a unified and coherent theory, Quantum Mechanics.

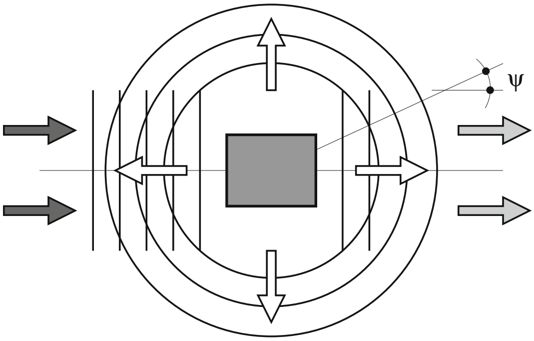


Fig. 7.11 Scheme of the experimental setup for measuring the Compton effect

In essence, the contradictions with the physical laws known in the early 1900s were found thanks to the refinement of the experimental techniques. Such refinements were in turn made available by the general advancement of science that had taken place in the preceding decades. Thanks to them, it was possible to start investigating the microscopic world, namely, the dynamics of elementary particles. A parallel improvement took place in the same years in the investigation of the dynamics at high velocities, and led to the concepts of Special Relativity (1905).[13]

7.4.1 Planck's Solution of the Black-Body Problem

To explain the features of the black-body radiation, Planck made in 1900 the hypothesis that a monochromatic electromagnetic energy is absorbed or emitted only in quantities that are integer multiples of a fixed quantity $h\nu$, where h is a suitable constant [82]. The occupation number then becomes

$$P_n = P_0 \exp(-n\beta\, h\, \nu), \qquad \beta = 1/(k_B\, T). \tag{7.17}$$

As a consequence, using the same procedure as in Sect. 6.6.2 after replacing the integrals with sums, yields for the average energy $\overline{n\, h\, \nu}$ the expression [14]

$$\overline{n\, h\, \nu} = \frac{\sum_{n=0}^{\infty} n\, h\, \nu\, P_n}{\sum_{n=0}^{\infty} P_n} = \frac{h\, \nu}{\exp(\beta\, h\, \nu) - 1}. \tag{7.18}$$

In contrast with the constant value $k_B\, T$ used in the determination of the Rayleigh-Jeans law, here the average energy of each monochromatic component depends on the

---

[13] As the particles' velocities that occur in solid-state physics are low, Special Relativity is not used in this book; the only exception is in the explanation of the Compton effect, illustrated in Sect. 7.4.3.

[14] The detailed calculation leading to (7.18) is shown in Prob. 6.1.

component's frequency. Multiplying (7.18) by the number $8\,\pi\,\nu^2/c^3$ of monochromatic components of the electromagnetic field per unit volume and frequency, found in Sect. 7.3, yields for the spectral energy density of the black body the expression

$$u(\nu, T) = 8\,\pi\,\frac{h\,\nu^3/c^3}{\exp\left[h\,\nu/(k_B\,T)\right] - 1}, \tag{7.19}$$

called *Planck law* (1900). The derivation of (7.19) involves one undetermined parameter, h. If the latter is made equal to the Planck constant introduced in the description of the photoelectric effect (Sect. 7.3), the resulting expression fits perfectly the experimental data like those of Fig. 7.10. Remembering that the spectral energy density of a black body in equilibrium is a universal function, it follows that h does not depend on the specific experiment, namely, it is a universal constant.

The low-frequency limit of (7.19), $h\,\nu \ll k_B\,T$, is independent of h and renders the Rayleigh–Jeans law (7.16).

7.4.2 Einstein's Solution of the Photoelectric Effect

In 1905, Einstein proposed the following explanation of the photoelectric effect: the transport of electromagnetic energy is *quantized*; specifically, a monochromatic electromagnetic wave of frequency ν is made of the flux of identical objects, called *photons*, each carrying the energy $h\,\nu$. In the interaction with a photon, an electron may absorb an energy up to $h\,\nu$. If the absorbed energy is exactly $h\nu$, the photon is annihilated [34].[15] This theory provides a correct explanation of the photoelectric effect: with reference to Fig. 7.7, the photoelectric current increases as the spectral power Π increases at constant ν, because the number of photons is larger: as a consequence, the number of photoelectrons is larger as well. In turn, with reference to Fig. 7.8, the blocking voltage V_R increases as ν increases at constant Π, because the photons are more energetic; however, they are fewer, which explains why the curves intersect each other: the spectral power, in fact, can be written as $\Pi = \mathrm{d}E/(\mathrm{d}t\,\mathrm{d}\nu) = h\,\nu\,[\mathrm{d}N/(\mathrm{d}t\,\mathrm{d}\nu)]$, where the quantity in brackets is the number of photons per unit time and frequency; as a consequence, the constraint $\Pi = \text{const}$ of the experiment of Fig. 7.8 makes the quantity in brackets to decrease when the photon energy $h\,\nu$ increases.

7.4.3 Explanation of the Compton Effect

The concept of photon, introduced in Sect. 7.4.2, explains the Compton effect by describing the interaction of the electron with the electromagnetic field as the collision

---

[15] Einstein's hypothesis is more general than Planck's: the latter, in fact, assumes that energy is quantized only in the absorption or emission events.

between the electron and a photon [19]. As the photon's velocity is c, its rest mass is zero (Sect. 3.13.7); in turn, the modulus of the photon's momentum is $p = E/c$, which is consistent with classical electromagnetism (compare with (5.43)).

The analysis of the electron-phonon collision is worked out assuming that the system made of the two particles under consideration is isolated; thus, the calculation is based upon the energy- and momentum-conservation equations, and the results of Sect. 3.13.8 hold. The dynamical quantities for the photon are given by

$$E = h\,\nu, \qquad p = \frac{E}{c} = \frac{h\,\nu}{c} = \frac{h}{\lambda}, \tag{7.20}$$

the second of which derives from (5.55) expressed *in vacuo*. Defining the *reduced Planck constant* $\hbar = h/(2\,\pi) \simeq 1.055 \times 10^{-34}$ J s, and using the modulus k of the wave vector, (7.20) becomes

$$E = \hbar\,2\pi\,\nu = \hbar\,\omega, \qquad p = \frac{\hbar}{\lambda/(2\,\pi)} = \hbar\,k. \tag{7.21}$$

The second relation of (7.21) in vector form reads

$$\mathbf{p} = \hbar\,\mathbf{k}. \tag{7.22}$$

Here the useful outcome of the analysis of Sect. 3.13.8 is (3.92), that relates the photon's energies prior and after the collision (E_a and E_b, respectively) with the deflection angle ψ (Fig. 3.7). Using $E = c\,h/\lambda$ in (3.92) yields

$$\lambda_b - \lambda_a = 2\,\lambda_0\,\sin^2\left(\frac{\psi}{2}\right), \qquad \lambda_0 = \frac{h}{m_0\,c}, \tag{7.23}$$

with $\lambda_0 \simeq 2.43 \times 10^{-12}$ m the *Compton wavelength* (1923). The frequency corresponding to it is $\nu_0 = c/\lambda_0 \simeq 1.2 \times 10^{20}$ Hz. The maximum difference in wavelength corresponds to the case of reflection, $\max(\lambda_b - \lambda_a) = 2\,\lambda_0$. Even in this case, the smallness of λ_0 makes the effect difficult to measure; in practice, the shift in wavelength is detectable only for sufficiently small values of λ_a, typically in the range of 10^{-10} m corresponding to the X-ray frequencies ($\nu \sim 10^{18}$ s^{-1}). Due to the large energy of the photon, the energy transferred to the electron brings the latter into a high-velocity regime; this, in turn, imposes the use of the relativistic expressions for describing the electron's dynamics.

7.4.4 Bohr's Hypothesis

The description of the monochromatic components of the electromagnetic field as a flow of identical photons with energy $h\,\nu$ lends itself to the explanation of the Balmer law (7.4). Such an explanation (*Bohr's hypothesis*, 1913) is based on the idea that, if

ν_{nm} is the frequency of the emitted radiation, the corresponding energy of the emitted photon is $h\,\nu_{nm}$; multiplying (7.4) by h and remembering that $m > n$ then yields

$$h\,\nu_{nm} = h\,\nu_R \left(\frac{1}{n^2} - \frac{1}{m^2}\right) = \left(-\frac{h\,\nu_R}{m^2}\right) - \left(-\frac{h\nu_R}{n^2}\right). \qquad (7.24)$$

As the left hand side is the energy of the emitted photon, the terms on right hand side can be recast as

$$E_m = -\frac{h\nu_R}{m^2}, \qquad E_n = -\frac{h\nu_R}{n^2}, \qquad E_n < E_m < 0\,; \qquad (7.25)$$

then, if E_m (E_n) is interpreted as the atom's energy before (after) emitting the photon, Balmer's law becomes the expression of energy conservation. From this, the emission rule of Ritz is easily explained; in fact, (7.5) is equivalent to

$$E_m - E_n = (E_m - E_k) + (E_k - E_n). \qquad (7.26)$$

Bohr's hypothesis is expressed more precisely by the following statements:

1. The energy variations of the atom are due to the electrons of the outer shell, that exchange energy with the electromagnetic field.
2. The total energy of a non-radiative state is quantized, namely, it is associated to an integer index: $E_n = -h\,\nu_R/n^2$, $n = 1, 2, \ldots$; the values of energy thus identified are called *energy levels*. The lowest level corresponds to $n = 1$ and is called *ground level* or *ground state*.
3. The total energy can vary only between the quantized levels by exchanging with the electromagnetic field a photon of energy $\nu_{nm} = (E_m - E_n)/h$.

It is interesting to note that, by combining Bohr's hypothesis with the planetary model of the atom, the quantization of the other dynamical quantities follows from that of energy; again, the case of a circular orbit is considered. By way of example, using $E_n = -h\,\nu_R/n^2$ in the second relation of (7.3) provides the quantization of the orbit's radius:

$$r = r_n = -\frac{q^2}{8\,\pi\,\varepsilon_0\,E_n} = \frac{q^2}{8\,\pi\,\varepsilon_0}\,\frac{n^2}{h\,\nu_R}. \qquad (7.27)$$

The smallest radius r_1 corresponds to the ground state $n = 1$; taking ν_R from (7.4) and the other constants from Table D.1 one finds $r_1 \simeq 0.05$ nm; despite the simplicity of the model, r_1 is fairly close to the experimental value r_a given in (7.1).

In turn, the velocity is quantized by combining (7.3) to obtain $T = -V/2 = -E$; replacing the expressions of T and E then yields

$$\frac{1}{2}\,m\,u^2 = \frac{h\,\nu_R}{n^2}, \qquad u = u_n = \sqrt{\frac{2\,h\,\nu_R}{m\,n^2}}. \qquad (7.28)$$

The largest velocity is found from (7.28) by letting $n = 1$ and using the minimum value for the mass, that is, the rest mass $m = m_0$. It turns out $u_1 \simeq 7 \times 10^{-3}\,c$;

as a consequence, the velocity of a bound electron belonging to the outer shell of the atom can be considered non relativistic. Thanks to this result, from now on the electron's mass will be identified with the rest mass. Finally, for the angular momentum $M = r\,p = r\,m\,u$ one finds

$$M = M_n = \frac{q^2\,n^2}{8\,\pi\,\varepsilon_0\,h\,\nu_R}\,m\,\sqrt{\frac{2\,h\,\nu_R}{m\,n^2}} = \frac{1}{2\,\pi}\left[\frac{q^2}{\varepsilon_0}\sqrt{\frac{m}{8\,h\,\nu_R}}\right]n. \tag{7.29}$$

The quantity in brackets in (7.29) has the same units as M, namely, an action (Sect. 1.5) and, replacing the constants, it turns out[16] to be equal to h. Using the reduced Planck constant it follows

$$M_n = n\,\hbar. \tag{7.30}$$

The Bohr hypothesis provides a coherent description of some atomic properties; yet it does not explain, for instance, the fact that the electron belonging to an orbit of energy $E_n = -h\,\nu_R/n^2$ does not radiate, in contrast to what is predicted by the electromagnetic theory (compare with the discussion in Sect. 7.3). Another phenomenon not explained by the hypothesis is the fact that only the ground state of the atom is stable, whereas the excited states are unstable and tend to decay to the ground state.

7.4.5 De Broglie's Hypothesis

The explanation of the Compton effect (Sect. 7.4.3) involves a description of the photon's dynamics in which the latter is treated like a particle having energy and momentum. Such mechanical properties are obtained from the wave properties of a monochromatic component of the electromagnetic field: the relations involved are (7.20) (or (7.21)), by which the photon energy is related to the frequency, and its momentum to the wave vector. It is worth specifying that such relations are applied to the asymptotic part of the motion, namely, when the photon behaves like a free particle. In 1924, de Broglie postulated that analogous relations should hold for the free motion of a real particle: in this case, the fundamental dynamic properties are energy and momentum, to which a frequency and a wavelength (or a wave vector) are associated by relations identical to (7.20), (7.21),[17]

$$\omega = 2\,\pi\,\nu = 2\,\pi\,\frac{E}{h} = \frac{E}{\hbar}, \qquad k = \frac{2\,\pi}{\lambda} = \frac{2\,\pi}{h/p} = \frac{p}{\hbar}, \qquad \mathbf{k} = \frac{\mathbf{p}}{\hbar}. \tag{7.31}$$

The usefulness of associating, e.g., a wavelength to a particle's motion lies in the possibility of qualitatively justifying the quantization of the mechanical properties

---

[16] This result shows that the physical constants appearing in (7.29) are not independent from each other. Among them, ν_R is considered the dependent one, while q, $m = m_0$, ε_0, and h are considered fundamental.

[17] The wavelength associated to the particle's momentum is called *de Broglie's wavelength*.

illustrated in Sect. 7.4.4. For this, consider the case of the circular orbit of the planetary motion, and associate a wavelength to the particle's momentum, $\lambda = h/p$. Such an association violates the prescription that (7.31) apply only to a free motion; however, if the orbit's radius is very large, such that $\lambda \ll r$, the orbit may be considered as locally linear and the concept of wavelength is applicable. Replacing $\lambda = h/p$ in (7.30) yields

$$2\pi r = n\lambda, \tag{7.32}$$

namely, the quantization of the mechanical properties implies that the orbit's length is an integer multiple of the wavelength associated to the particle. This outcome suggests that the formal description of quantization should be sought in the field of eigenvalue equations.

De Broglie also postulated that a function $\psi = \psi(\mathbf{r}, t)$, containing the parameters ω, $\mathbf{k}$ defined in (7.31), and called *wave function*, is associated to the particle's motion. Its meaning is provisionally left indefinite; as for its form, it is sensible to associate to the free motion, which is the simplest one, the simplest wave function, that is, the planar monochromatic wave. The latter is conveniently expressed in complex form as

$$\psi = A \exp[\mathrm{i}(\mathbf{k}\cdot\mathbf{r} - \omega t)], \tag{7.33}$$

where $A \neq 0$ is a complex constant, not specified. Due to (7.31), the constant wave vector $\mathbf{k}$ identifies the momentum of the particle, and the angular frequency ω identifies its total energy, which in a free motion coincides with the kinetic energy. It is worth pointing out that, despite its form, the wave function is not of electromagnetic nature; in fact, remembering that in a free motion it is $H = p^2/(2m) = E$, with H the Hamiltonian function, it follows

$$\hbar\omega = \frac{1}{2m}\hbar^2 k^2, \qquad \omega(\mathbf{k}) = \frac{\hbar}{2m}k^2, \tag{7.34}$$

which is different from the electromagnetic relation $\omega = c\,k$. By the same token it would not be correct to identify the particle's velocity with the phase velocity u_f derived from the electromagnetic definition; in fact, one has

$$u_f = \frac{\omega}{k} = \frac{E/\hbar}{p/\hbar} = \frac{p^2/(2m)}{p} = \frac{p}{2m}. \tag{7.35}$$

The proper definition of velocity is that deriving from Hamilton's Eqs. (1.42); its ith component reads in this case

$$u_i = \dot{x}_i = \frac{\partial H}{\partial p_i} = \frac{1}{\hbar}\frac{\partial H}{\partial k_i} = \frac{\partial\omega}{\partial k_i} = \frac{\hbar k_i}{m} = \frac{p_i}{m}. \tag{7.36}$$

The concepts introduced so far must now be extended to motions of a more general type. A sensible generalization is that of the conservative motion of a particle subjected to the force deriving from a potential energy $V(\mathbf{r})$. In this case the association described by (7.31) works only partially, because in a conservative motion the total energy is a constant, whereas momentum is generally not so. As a consequence, letting $\omega = E/\hbar$ yields for the wave function the form

$$\psi = w(\mathbf{r}) \exp(-\mathrm{i}\,\omega\,t), \tag{7.37}$$

which is still monochromatic but, in general, not planar. Its *spatial part* $w(\mathbf{r})$ reduces to $A\, \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})$ for the free motion. The function of the form (7.37) is postulated to be the wave function associated with the motion of a particle at constant energy $E = \hbar\omega$. While the time dependence of ψ is prescribed, its space dependence must be worked out, likely by solving a suitable equation involving the potential energy V, the particle's mass and, possibly, other parameters.

7.5 Heuristic Derivation of the Schrödinger Equation

The concept of wave function introduced in Sect. 7.4.5 has been extended from the case of a free motion, where the wave function is fully prescribed apart from the multiplicative constant A, to the case of a conservative motion (7.37), where only the time dependence of the wave function is known. It is then necessary to work out a general method for determining the spatial part $w(\mathbf{r})$. The starting point is the observation that w is able at most to provide information about the particle's trajectory, not about the particle's dynamics along the trajectory. One of the methods used in Classical Mechanics to determine the trajectories is based on the Maupertuis principle (Sect. 2.7); moreover, from the discussion carried out in Sect. 5.11.6 it turns out that the analogy between the Maupertuis principle and the Fermat principle of Geometrical Optics (compare with (5.80)) provides the basis for a generalization of the mechanical laws. The first of the two principles applies to a particle (or system of particles) subjected to a conservative force field prescribed by a potential energy $V(\mathbf{r})$, with E a given constant; the second one applies to a monochromatic ray ($\nu = \text{const}$) propagating in a medium whose properties are prescribed by the refraction index $n(\mathbf{r})$. The latter is related to frequency and wavelength by $n = c/(\lambda\,\nu)$ (compare with (5.55)); as a consequence, (5.80) can be rewritten as

$$\delta \int_{AB} \sqrt{E - V}\, \mathrm{d}s = 0, \qquad \delta \int_{AB} \frac{1}{\lambda}\, \mathrm{d}s = 0. \tag{7.38}$$

Considering that the variational principles hold apart from a multiplicative constant, the two expressions in (7.38) transform into each other by letting

$$\sqrt{E - V} = \frac{\alpha}{\lambda}, \tag{7.39}$$

where α is a constant that must not depend on the form of V or λ, nor on other parameters of the problem. For this reason, α is left unchanged also after removing the Geometrical-Optics approximation; when this happens, the Fermat principle is replaced with the Maxwell equations or, equivalently, with the wave equations for the electric field (4.64) and magnetic field (4.65). For simplicity, the latter equations are solved in a uniform medium with no charges in it, on account of the fact that α is not influenced by the medium's properties. Also, considering that in the uniform case (4.64) and (4.65) have the same structure, and that the function w under investigation is scalar, the analysis is limited to any scalar component C of $\mathbf{E}$ or $\mathbf{H}$; such a component fulfills the equation

$$\nabla^2 C - \frac{1}{u_f^2}\frac{\partial^2 C}{\partial t^2} = 0, \tag{7.40}$$

with $u_f = \text{const}$ the medium's phase velocity. Solving (7.40) by separation with $C(\mathbf{r}, t) = \eta(\mathbf{r})\,\theta(t)$ yields

$$\theta\,\nabla^2\eta = \frac{1}{u_f^2}\,\ddot{\theta}\,\eta, \qquad \frac{\nabla^2\eta}{\eta} = \frac{1}{u_f^2}\,\frac{\ddot{\theta}}{\theta} = -k^2, \tag{7.41}$$

where the separation constant $-k^2$ must be negative to prevent θ from diverging. As a consequence, k is real and can be assumed to be positive. The solution for the time factor is $\theta = \cos(\omega t + \varphi)$, where the phase φ depends on the initial conditions, and $\omega = 2\pi\,\nu = u_f\,k > 0$. It follows $k = 2\,\pi\,\nu/u_f = 2\,\pi/\lambda$, whence the spatial part of (7.40) reads

$$\nabla^2\eta + \frac{(2\,\pi)^2}{\lambda^2}\,\eta = 0, \tag{7.42}$$

namely, a Helmholtz equation (Sect. 4.7). By analogy, the equation for the spatial part w of the wave function is assumed to be

$$\nabla^2 w + \frac{(2\,\pi)^2}{\alpha^2}\,E\,w = 0, \tag{7.43}$$

which is obtained from (7.42) by replacing η with w and using (7.39) with $V = 0$. The value of α is determined by expressing E in (7.43) in terms of the de Broglie wavelength; using the symbol λ_{dB} for the latter to avoid confusion with the electromagnetic counterpart, one finds $E = p^2/(2\,m) = h^2/(2\,m\,\lambda_{\text{dB}}^2)$, namely,

$$\nabla^2 w + \frac{(2\,\pi)^2}{\alpha^2}\,\frac{h^2}{2\,m}\,\frac{1}{\lambda_{\text{dB}}^2}\,w = 0. \tag{7.44}$$

Equation (7.44) becomes identical to (7.42) by letting $\alpha^2 = h^2/(2m)$ whence, using the reduced Planck constant $\hbar$, (7.43) becomes $\nabla^2 w + (2\,m\,E/\hbar^2)\,w = 0$. Such a differential equation holds in a uniform medium; hence, the dynamical property involved is the kinetic energy of the particle. The extension to the case of a non-uniform medium is then obtained by using the general form $E - V$ of the kinetic energy in terms of the coordinates; in conclusion, the equation for the spatial part of the wave function in a conservative case is

$$\nabla^2 w + \frac{2\,m}{\hbar^2}\,(E - V)\,w = 0, \qquad -\frac{\hbar^2}{2\,m}\,\nabla^2 w + V\,w = E\,w. \tag{7.45}$$

The above is the *time-independent Schrödinger equation*. It is a homogenous equation, with E the eigenvalue and w the eigenfunction.[18] Although the derivation based on the analogy between mechanical and optical principles is easy to follow, it must be remarked that the step leading from (7.44) to (7.45) is not rigorous; in the electromagnetic case, in fact, Eq. (7.42) for the spatial part holds only in a uniform medium; when the latter is non uniform, instead, the right hand side of (7.42) is different from zero, even in a charge-free case, because it contains the gradient of the refraction index. As shown in Sect. 1.10.4, the actual method used in 1926 by Schrödinger for deriving (7.45) consists in seeking the constrained extremum of a functional generated by the Hamilton–Jacobi equation; in such a procedure, the hypothesis of a uniform medium is not necessary.

It is also worth noting that in the analogy between mechanical and optical principles the spatial part of the wave function, and also the wave function as a whole, is the analogue of a component of the electromagnetic field. From this standpoint, the analogue of the field's intensity is the wave function's square modulus. In the monochromatic case, the latter reads $|\psi|^2 = |w|^2$. This reasoning is useful in the discussion about the physical meaning of ψ.

7.6 Measurement

To make the wave function a useful tool for the description of the particles' dynamics it is necessary to connect the value taken by ψ, at a specific position and time, with some physical property of the particle (or system of particles) under investigation. To make such a connection it is in turn necessary to measure the property of interest; otherwise, the hypotheses illustrated in the previous sections would be relegated to a purely abstract level. In other terms, the meaning of the wave function can be given only by discussing the measuring process in some detail. The analysis is carried out below, following the line of [69]; in particular, a general formalism is sought which applies to both the macroscopic and microscopic bodies; the specific features brought about by the different size of the objects that are being measured are made evident by suitable examples.

The measurement of a dynamical variable A pertaining to a physical body is performed by making the body to interact with a measuring apparatus and recording the reading shown by the latter. For simplicity it is assumed that there is a finite number of possible outcomes of the measurement, say, $A_1, \dots, A_M$. The extension to the case of a continuous, infinitely extended set of outcomes can be incorporated into the theory at the cost of a more awkward notation. Letting A_i be the outcome of the measurement of A, consider the case where the body is later subjected to the measurement of another dynamical variable B. Assume that the outcome of such a measurement is B_j, out of the possible outcomes $B_1, \dots, B_N$. Next, the body is subjected to the measurement of a third variable C, this yielding the value C_k, and so on.

---

[18] The structure of (7.45) is illustrated in detail in Chap. 8.

As in general the dynamical variables depend on time, it is necessary to specify the time of each measurement. The most convenient choice is to assume that the time interval between a measurement and the next one is negligibly small, namely, that the measurement of B takes place immediately after that of A, similarly for that of C, and so on. The duration of each measurement is considered negligible as well. A special consequence of this choice is the following: if the measurement of A yielded A_i, and the measurement is repeated (namely, $B = A$), the outcome of the second measurement is again A_i.

Consider now the case where, after finding the numbers A_i, B_j, and C_k from the measurements of A, B, and C, respectively, the three variables are measured again, in any order. The experiments show that the results depend on the size of the body being measured. For a massive body the three numbers A_i, B_j, and C_k are always found. One concludes that the dynamical state of a massive body is not influenced by the interaction with the measuring apparatus or, more precisely, that if such an influence exists, it is so small that it can not be detected. As a consequence one may also say that the values of the dynamical variables are properties of the body that exist prior, during, and after each measurement.

The situation is different for a microscopic body. By way of example, consider the case of a measurement of B followed by a measurement of A, the first one yielding B_n, the second one yielding A_i. If the measurement of B is carried out again after that of A, the result is still one of the possible outcomes $B_1, \dots, B_N$, but it is not necessarily equal to B_n. In other terms, the individual outcome turns out to be unpredictable. For a microscopic body one concludes that the interaction with the measuring apparatus is not negligible. It is worth observing that the apparatus able to measure the dynamical variable A may also be conceived in such a way as to block all outcomes that are different from a specific one, say, A_i. In such a case the apparatus is termed *filter*. Using the concept of filter one may build up the statistical distribution of the outcomes, for instance by repeating a large number of times the experiment in which the measurement of B is carried out after filtering A_i. The statistics is built up by recording the fraction of cases in which the measurement of B carried out on an A_i-filtered body yields the result B_j, $j = 1, \dots, N$.

7.6.1 Probabilities

The fraction of measurements of the type described above, namely, of those that yield B_j after a measurement of A that has yielded A_i, will be indicated with the symbol $P(A_i \rightarrow B_j)$. Obviously the following hold:

$$0 \le P(A_i \rightarrow B_j) \le 1, \qquad \sum_{j=1}^{N} P(A_i \rightarrow B_j) = 1. \tag{7.46}$$

The first relation in (7.46) is due to the definition of $P(A_i \rightarrow B_j)$, the second one to the fact that the set of values $B_1, \dots, B_N$ encompasses all the possible outcomes of the measurement of B. It follows that $P(A_i \rightarrow B_j)$ is the probability that a measurement of the dynamical variable B, made on a particle that prior to the measurement is in the state A_i of the dynamical variable A, yields the value B_j. The possible combinations of $P(A_i \rightarrow B_j)$ are conveniently arranged in the form of an M-row×N-column matrix:

$$\mathbf{P}_{AB} = \begin{bmatrix} P(A_1 \rightarrow B_1) & \dots & P(A_1 \rightarrow B_N) \\ \vdots & & \vdots \\ P(A_i \rightarrow B_1) & \dots & P(A_i \rightarrow B_N) \\ \vdots & & \vdots \\ P(A_M \rightarrow B_1) & \dots & P(A_M \rightarrow B_N) \end{bmatrix}. \tag{7.47}$$

Due to (7.46), each row of $\mathbf{P}_{AB}$ adds up to unity. As the number of rows is M, the sum of all entries of matrix (7.47) is M. The same reasoning can also be made when the measurement of B is carried out prior to that of A. In this case the following N-row×M-column matrix is obtained:

$$\mathbf{P}_{BA} = \begin{bmatrix} P(B_1 \rightarrow A_1) & \dots & P(B_1 \rightarrow A_M) \\ \vdots & & \vdots \\ P(B_j \rightarrow A_1) & \dots & P(B_j \rightarrow A_M) \\ \vdots & & \vdots \\ P(B_N \rightarrow A_1) & \dots & P(B_N \rightarrow A_M) \end{bmatrix}, \tag{7.48}$$

with

$$0 \le P(B_j \rightarrow A_i) \le 1, \qquad \sum_{i=1}^{M} P(B_j \rightarrow A_i) = 1. \tag{7.49}$$

As the number of rows in (7.49) is N, the sum of all entries of matrix $\mathbf{P}_{BA}$ is N. It can be proven that it must be

$$P(B_j \rightarrow A_i) = P(A_i \rightarrow B_j) \tag{7.50}$$

for any pair of indices ij. In fact, if (7.50) did not hold, thermodynamic equilibrium would not be possible [69, Ch. V-21]. Equality (7.50) makes $\mathbf{P}_{BA}$ the transpose of $\mathbf{P}_{AB}$. As a consequence, the sum of all entries of the two matrices must be the same, namely, $N = M$. In other terms the outcomes of the measurements have the same multiplicity, and the matrices (7.47), (7.48) are square matrices of order $N = M$. Combining (7.50) with the second of (7.49) yields

$$\sum_{i=1}^{M} P(A_i \rightarrow B_j) = \sum_{i=1}^{M} P(B_j \rightarrow A_i) = 1, \tag{7.51}$$

showing that in the matrices (7.47), (7.48) not only each row, but also each column adds up to unity. A square matrix where all entries are non negative and all rows and columns add up to unity is called *doubly stochastic matrix*. Some properties of this type of matrices are illustrated in [76, Ch. II-1.4] and in Sect. A.11.

Note that (7.50) does not imply any symmetry of $\mathbf{P}_{AB}$. In fact, symmetry would hold if $P(A_j \rightarrow B_i) = P(A_i \rightarrow B_j)$. If the filtered state is A_i and the measurement of the dynamical variable A is repeated, the result is A_i again. In other terms,

$$P(A_i \rightarrow A_i) = 1, \qquad P(A_i \rightarrow A_k) = 0, \quad k \neq i. \tag{7.52}$$

This result can be recast in a more compact form as $\mathbf{P}_{AA} = \mathbf{I}$, with $\mathbf{I}$ the identity matrix.

7.6.2 Massive Bodies

It is useful to consider the special case where the measurement of B does not change the outcome of a previous measurement of A, and viceversa. In other terms, assume that the measurement of A has yielded A_i and the subsequent measurement of B has yielded B_j; then, another measure of A yields A_i again, a later measure of B yields B_j again, and so on. It follows that in $\mathbf{P}'_{AB}$ it is $P'(A_i \rightarrow B_j) = 1$, while all remaining entries in the ith row and jth column are equal to zero. This situation is typical of the bodies that are sufficiently massive, such that the interference suffered during the measurement of a dynamical variable is not detectable. For the sake of clarity an apex is used here to distinguish the probabilities from those of the general case where the body's mass can take any value. Considering a 4 × 4 matrix by way of example, a possible form of the matrix would be

$$\mathbf{P}'_{AB} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{7.53}$$

that is, one of the 4! possible permutation matrices of order 4. Clearly all the other permutation matrices of order 4 different from (7.53) are equally possible. The meaning of a matrix like (7.53) is that the successive measurements of A and B yield either the pair A_1, B_2, or the pair A_2, B_3, or A_3, B_1, or A_4, B_4. Matrix (7.53) may be thought of as a limiting case: starting from a microscopic body described by a 4 × 4, doubly stochastic matrix whose entries are in general different from zero, the size of the body is increased by adding one atom at a time, and the set of measurements of A and B is repeated at each time. As the reasoning that prescribes the doubly-stochastic nature of the matrix holds at each step, the successive matrices must tend to the limit of a permutation matrix. Which of the 4! permutation matrices will be reached by this process depends on the initial preparation of the experiments. One may wonder why a matrix like

$$\mathbf{P}'_{AB} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{7.54}$$

should not be reached. In fact, such a matrix is not acceptable because it is not doubly stochastic: its transpose implies that the outcomes B_1 and B_2 are simultaneously associated to A_2 with certainty, which is obviously impossible not only for a massive body, but for any type of body. This reasoning is associated to another argument, based on the theorem mentioned in Sect. A.11, stating that a doubly-stochastic matrix is a convex combination of permutation matrices. Letting $\theta_1, \dots, \theta_M$ be the combination's coefficients as those used in (A.42), in the process of transforming the microscopic body into a macroscopic one all the coefficients but one vanish, and the non-vanishing one tends to unity. As a consequence, out of the original combination, only one permutation matrix is left.

7.6.3 Need of a Description of Probabilities

The non-negligible influence of the measuring apparatus on the dynamical state of a microscopic body makes it impossible to simultaneously measure the dynamical variables that constitute the initial conditions of the motion. As a consequence, the possibility of using the Hamiltonian theory for describing the dynamics is lost. As outlined in the above sections, the distinctive mark of experiments carried out on microscopic objects is the statistical distribution of the outcomes; thus, a theory that adopts the wave function as the basic tool must identify the connection between the wave function and such a statistical distribution. The theory must also contain the description of the massive bodies as a limiting case.

7.7 Born's Interpretation of the Wave Function

Basing on the optical analogy and the examination of experiments, the *probabilistic interpretation* of the wave function introduced by Born states that the integral

$$\int_\tau |\psi(\mathbf{r}, t)|^2 \, \mathrm{d}^3 r \tag{7.55}$$

is proportional to the probability that a measuring process finds the particle within the volume τ at the time t.[19] Note that the function used in (7.55) is the square modulus of ψ, namely, as noted in Sect. 7.5, the counterpart of the field's intensity in the optical analogy. Also, considering that by definition the integral of (7.55) is dimensionless, the units[20] of ψ are $\mathrm{m}^{-3/2}$.

When $\tau \to \infty$ the integral in (7.55) may, or may not, converge. In the first case, ψ is said to be *normalizable*, and a suitable constant σ can be found such that the integral of $|\sigma\,\psi|^2$ over the entire space equals unity. The new wave function provides a probability proper,

$$\int_\tau |\sigma\,\psi|^2\,\mathrm{d}^3r \le 1, \qquad \sigma^{-2} = \int_\infty |\psi|^2\,\mathrm{d}^3r. \tag{7.56}$$

In the second case ψ is not normalizable:[21] a typical example is the wave function of a free particle, $\psi = A\,\exp[\mathrm{i}\,(\mathbf{k}\cdot\mathbf{r} - \omega t)]$; however, it is still possible to define a probability ratio

$$\int_{\tau_1} |\psi|^2\,\mathrm{d}^3r\,\left(\int_{\tau_2} |\psi|^2\,\mathrm{d}^3r\right)^{-1}, \tag{7.57}$$

where both volumes τ_1 and τ_2 are finite. Relation (7.57) gives the ratio between the probability of finding the particle within τ_1 and that of finding it within τ_2.

Consider a particle whose wave function at time t differs from zero within some volume τ, and assume that a process of measuring the particle's position is initiated at t and completed at some later time t'; let the outcome of the experiment be an improved information about the particle's location, namely, at t' the wave function differs from zero in a smaller volume $\tau' \subset \tau$. This event is also called *contraction* of the wave function.

7.8 Complements

7.8.1 Core Electrons

Throughout this book the term "nucleus" is used to indicate the system made of protons, neutrons, and *core electrons*, namely, those electrons that do not belong to the outer shell of the atom and therefore do not participate in the chemical bonds. In solid-state materials, core electrons are negligibly perturbed by the environment, in contrast to the electrons that belong to the outer shell (*valence electrons*).

---

[19] From this interpretation it follows that $|\psi|^2\,\mathrm{d}^3r$ is proportional to an infinitesimal probability, and $|\psi|^2$ to a probability density.

[20] A more detailed discussion about the units of the wave function is carried out in Sect. 9.7.1.

[21] This issue is further discussed in Sect. 8.2.