# PART I

# FORMALISM AND BASIC PROBLEMS

# Introduction: Historical Background

In this book, we will describe the formalism and applications of quantum mechanics, starting from the first principles and postulates, and expanding them logically into the fully developed system that we currently have. That means that we will start with those principles and postulates, assuming that the reader has had an undergraduate course in modern physics, as well as an undergraduate course in quantum mechanics. This implies a first interaction with the ideas of quantum mechanics, and the historical development and experiments that led to it. However, for the sake of consistency and completeness, in this chapter we will quickly review these historical developments and how we were led to the current formalism for quantum mechanics.

In classical mechanics and optics, there were deterministic laws, involving two types of objects: particles of matter, following well-defined classical paths, with their evolution defined by their Hamiltonian and the initial conditions, and waves for light, defined by wave functions (for the electric and magnetic fields $\vec{E}$ and $\vec{B}$), and the observable intensities $I$ defined by them, showing interference patterns in the sum of waves. But there was no overlap between the two descriptions, in terms of particles and waves. Notably, Newton had a rudimentary particle description for light (in his *Optics*), but with the development of the classical wave picture of Huyghens for optics, that description was forgotten for a long time.

## 0.1 Experiments Point towards Quantum Mechanics

Then, around the turn of the nineteenth century into the twentieth, a string of developments changed the classical picture, allowing for a complementary particle description for waves, and for a complementary wave description for particles. One of the first such developments was the discovery of radioactivity by Henri Becquerel in 1896, who, after Wilhelm Roentgen's discovery of X-rays in 1895, showed that radioactive elements produce rays similar to X-rays in that they can go through matter and then leave a pointlike mark on a photographic plate, thus involving emission of highly energetic particles. Now we know that these emitted particles can be $\alpha$ (nucleons), $\beta$ (electrons and positrons), but also $\gamma$ (high-energy photons, so lightlike).

Next came what is believed to be the start of the modern quantum era, with the first theoretical idea of a quantum, specifically of light, used to describe blackbody radiation. In 1900, Planck managed to explain the blackbody emission spectrum with a simple but revolutionary idea. Assuming that the energy exchange between matter and the emitted radiation occurs not continuously (as assumed in classical physics), but only in given quanta of energy, proportional to the frequency $\nu$ of the radiation,

$$\Delta E = h\nu, \tag{0.1}$$

where the constant $h$ is called the Planck constant, Planck managed to derive a formula for the spectral radiance:

$$B_\nu(\nu, T) = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/k_BT} - 1}, \tag{0.2}$$

which matches the emission spectrum of the radiation of a perfect black body, which was then known experimentally. Note that by integrating, we obtain the power per unit of emission area,

$$P = \int_0^\infty d\nu \int d\Omega B_\nu \cos\theta = \sigma T^4, \tag{0.3}$$

where we have used the differential of solid angle $d\Omega = \sin\theta\, d\theta d\phi$ and have defined the Stefan–Boltzmann constant

$$\sigma = \frac{2k_B^4\pi^5}{15c^2h^3}, \tag{0.4}$$

obtaining the (experimentally discovered) Stefan–Boltzmann law. Planck himself didn't quite believe in the physical reality of the quantum of energy (as existing independently of the phenomenon of emission of radiation), but thought of it as a useful trick that manages to provide the correct result.

It was necessary to wait until Einstein's explanation of the photoelectric effect in 1905 for the true start of the quantum era. Indeed, Einstein took Planck's idea to its logical conclusion, and postulated that there is a quantum of energy, that he called a *photon* and that can interact with matter, such that its energy is $E = h\nu$. Such an energy quantum is *absorbed* in the photoelectric effect by the bound electrons with binding energy $W$, such that they are released with kinetic energy

$$E_{\text{kin}} = \frac{mv^2}{2} = h\nu - W, \tag{0.5}$$

providing an electrical current; see Fig. 0.1a.

Continuing the idea of the interaction of electrons with quanta of light, i.e., photons, we also note the *Compton effect*, observed experimentally by Arthur Compton in 1923, in which a photon scatters off a free electron and changes its wavelength according to the law

$$\Delta\lambda = \frac{2h}{mc}\sin^2\frac{\theta}{2}. \tag{0.6}$$

The law is easily explained by assuming that the photon has a relativistic relation between the energy and momentum, $E = pc$, leading to a momentum

$$p = \frac{h\nu}{c} = \frac{h}{\lambda}. \tag{0.7}$$

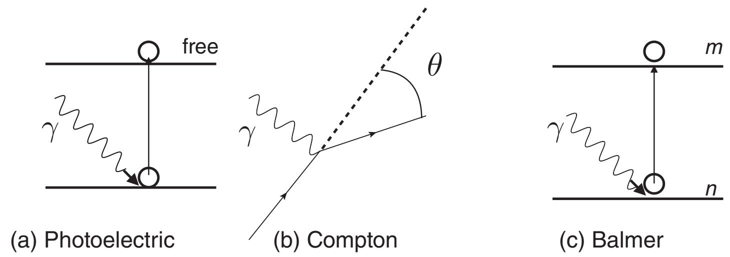


**Figure 0.1** Interactions of photons with electrons: (a) photoelectric effect; (b) Compton effect; (c) Balmer relation (transition to excited state).

Then relativistic energy and momentum conservation for the process $e^- + \gamma \rightarrow e^- + \gamma$ (see Fig. 0.1b), where the initial $e^-$ is (almost) at rest,

$$
\begin{aligned}
\vec{p}_\gamma &= \vec{p}'_\gamma + \vec{p}_e \\
mc^2 + p_\gamma c &= \sqrt{p_e^2 c^2 + m^2 c^4} + p'_\gamma c,
\end{aligned}
\tag{0.8}
$$

leads to the Compton law (0.6).

Next, we can also consider the case where the absorbed or emitted energy quantum leaves the electron bound to an atom, just changing its state inside the atom; see Fig. 0.1c. Since not all possible energies for the photon can be absorbed in this way, this means that the allowed energy states for the electron inside the atom are also quantized, i.e., they can only take well-defined values. In fact, experimentally we have the *Balmer relation*, discovered by Johann Balmer in 1885, for the possible frequencies of the absorbed or emitted radiation for the hydrogen (H) atom, given by

$$
\nu = R\left(\frac{1}{n^2} - \frac{1}{m^2}\right), \tag{0.9}
$$

where $R$ is called the Rydberg constant. Together with Planck's and Einstein's relation $\Delta E = h\nu$, we obtain the law that the possible energies of the electron inside the H atom are given by the *energy levels*

$$
E_n = -\frac{hR}{n^2}, \tag{0.10}
$$

where $n$ is a natural nonzero number, i.e. $n = 1, 2, 3, \ldots$

The fact that the states of the electron in the H atom (and, in fact, of any atoms or molecules, as we now know) are quantized suggests that it is possible to have more general quantization relations for energy states. That this is true, and that we can turn it into a concrete description for states in quantum mechanics, was proved in another important experiment, the *Stern–Gerlach experiment* of 1922 (by Otto Stern and Walter Gerlach). In this experiment, horizontal paramagnetic atomic beams are passed through a magnetic field varying as a function of the vertical position, $B = B(z)$, created by two magnets oriented on a vertical line, as in Fig. 0.2. The electrons are understood as "spinning", like magnets with a magnetic moment $\mu$, so the energy of the interaction of the "spinning" electrons with the magnetic field is given by

$$
\Delta E = -\vec{\mu} \cdot \vec{B}. \tag{0.11}
$$

But the magnetic moment is proportional to the "angular momentum" $\vec{l}$, $\vec{\mu} = m\vec{l}$. Classically, we expect any possible value for $\vec{l}$ and, so any possible value for both the magnetic moment $\vec{\mu}$, and its projection $\mu_z$ on the $z$ direction. Since the force in the $z$ direction is

$$
F_z = -\partial_z(\vec{\mu} \cdot \vec{B}) \simeq \mu_z \partial_z B_z, \tag{0.12}
$$

we expect the arbitrary value of $\mu_z$ to translate into an arbitrary deviation of the beam, experimentally detected on a screen transverse to the original beam. But in fact one observes only two possible deviations, implying only two possible values of $\mu_z$, symmetric about zero, which in fact can only be $+\mu$ and $-\mu$, with a fixed $\mu$. That in turn means that $l_z$ has only the possible values $+l$ and $-l$, with $l$ fixed.

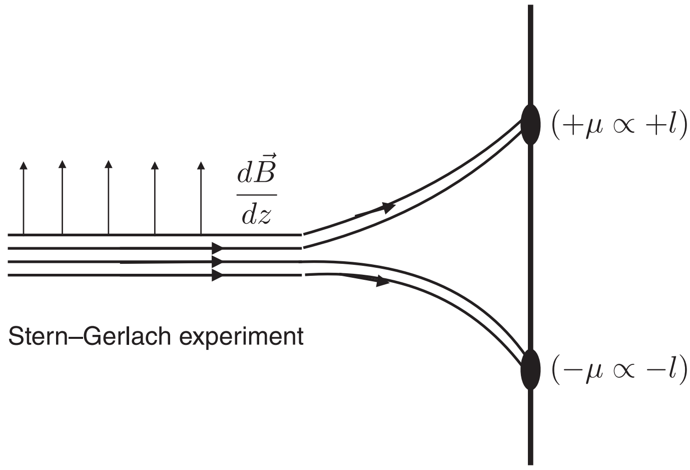


Figure 0.2 The Stern–Gerlach experiment.

## 0.2 Quantized States: Matrix Mechanics, and Waves for Particles: Correspondence Principle

The conclusion is that the "spinning" states of the electron have a projection onto any axis (since $z$ is a randomly chosen direction) that can have only two given values, which is another example of the quantization of states, involving the simplest possibility: two possible choices only, states "1" and "2". In general, we expect more quantized values for the energy, therefore more, but usually countable, states. Observables in these states can be "diagonal", meaning they map the state to itself (for instance, the energy of a state), or "off-diagonal", meaning they map one state to another (like those related to the transition between matter states due to interaction with light). In total, we obtain a "matrix mechanics", for matrix observables $M_{ij}$ acting on states $i$. This was defined by Werner Heisenberg, Max Born, and Pascual Jordan in 1925, in three papers (first by Heisenberg, then generalized and formalized by Born and Jordan, then by all three).

But one still had to understand what is the physical meaning of the states, and obtain an alternative to the classical idea of the path of a matter-particle. This came with the idea of a wave associated with any particle, and conversely, a particle associated with any wave, or *particle–wave complementarity*; see Fig. 0.3a. This is illustrated best in the classic double-slit experiment, which today is a table top experiment covered in undergraduate physics. Consider a plane wave, or a beam of particles, moving perpendicularly towards a screen with two slits in it, and observe the result on a detecting parallel screen behind it, as in Fig. 0.3b.

If we have classical waves, as in classical electromagnetism (the description of the macroscopic quantities of light), a *traveling plane wave* is described by a function $\psi$ depending only on the propagation direction $x$ and the time $t$, as

$$\psi(x,t) = A \exp\left[i\left(\frac{2\pi}{\lambda}x - \frac{2\pi}{T}t\right)\right], \tag{0.13}$$

where $\psi$ is made up from $\vec{E}$ and $\vec{B}$. The general (spherical) form of the traveling wave is

$$\psi(r,t) = \frac{A}{r}\exp\left[i(kr - \omega t)\right], \tag{0.14}$$

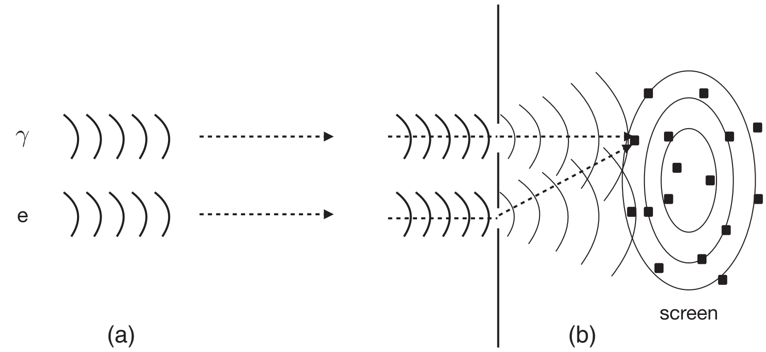


**Figure 0.3** (a) Particle–wave complementarity (correspondence principle). (b) Two-slit interference of waves $\psi_1$, $\psi_2$ giving a screen image made up of individual points (particles), forming an interference pattern at large times.

and is the form that is relevant for the wave that comes out of a slit. One can however measure only the intensity, given by

$$I = |\psi|^2. \tag{0.15}$$

The role of the double slits is to split the plane wave into two waves, 1 and 2, with origin in each slit, but converging on the same point on the detecting screen. Then one detects the total $I = I_{1+2}$, but what adds up is not $I$, but rather $\psi$, so $\psi_{1+2} = \psi_1 + \psi_2$, and

$$I_{1+2} = |\psi_1 + \psi_2|^2. \tag{0.16}$$

This leads to the interference pattern observed on the screen, with many local maxima of decaying magnitude.

On the other hand, for particles, the particle beam divides at the slits in two beams. The “intensity” of the beam is proportional to the number of particles per unit area (with the total number of particles classically a constant, since they cannot be created or destroyed), so for them one finds instead

$$I_{1+2} = I_1 + I_2. \tag{0.17}$$

This would lead to a pattern with a single maximum, situated at the midpoint of the screen, but this is not what it is observed.

In fact, as one knows from the experiment, if one decreases the intensity of the beam, such that, according to the previous description, we have only individual photons (quanta of light) coming through the slits, one can see the individual photons hitting the screen. However, their locations are such that, when sufficiently many of them have hit, they still create the same interference pattern of the waves. That means that, in a sense, a photon can “interfere with itself”.

It also implies a *correspondence principle*, that classical physics corresponds to macroscopic effects, whereas microscopic effects are described by quantum mechanics. As long as something becomes macroscopic (through many iterations of the microscopic, like for instance the large number of photons in the double slit experiment), it becomes classical.

So the question arises, is there a wave associated with the photon itself? And what does it correspond to? And if this is true for a photon, could it be true for any particle, even a particle of matter? In fact, historically, this was first proposed by (the Marquis) Louis de Broglie in 1924, who said that there should be a wave number $k$ associated with any particle of momentum $p$, given by

$$k = \frac{p}{\hbar}, \tag{0.18}$$

where $\hbar = h/(2\pi)$. Then it follows that the de Broglie wavelength of the particle is

$$\lambda = \frac{2\pi}{k} = \frac{h}{p}. \tag{0.19}$$

This gives the correct formula for a photon, since for a photon, $p = E/c$ and $\lambda = c/\nu$. But the formula is supposed to also apply for any matter particle, such as for instance an electron. For electrons, the formula was confirmed experimentally by Davisson and Germer in experiments performed between 1923 and 1927. Today, the principle implied in these experiments is used in the electron microscope, which uses electrons instead of photons in order to "see". This allows it to have a better resolution, since the resolution is of the order of $\lambda$ and the de Broglie $\lambda$ for an electron is much smaller than that for a photon of light (since the momentum $p$ is much larger).

## 0.3 Wave Functions Governing Probability, and the Schrödinger Equation

Because the intensity of a beam of particles is $I = |\psi|^2$, if we allow for the possibility that particles have only *probabilities* of behaving in some way, we arrive at the conclusion that the intensity $I$ must be proportional to this probability, and that as such the probability is given by $|\psi|^2$, where $\psi$ is a *wave function* associated with any particle. This was the interpretation given by Erwin Schrödinger, who then went on to write an equation, now known as *Schrödinger's equation*, for the wave function. The equation for $\psi$ is in general

$$i\hbar\partial_t\psi = \hat{H}\psi, \tag{0.20}$$

where $\hat{H}$ is an operator associated with the classical Hamiltonian $H$, now acting on the wave function $\psi$.

This interpretation of quantum mechanics, of the Schrödinger equation for a wave function associated with probability, was complementary to the matrix mechanics of Heisenberg, Born, and Jordan. The two however are united into a single one by the interpretation of Dirac in terms of bra $\langle\psi|$ and ket $|\psi\rangle$ states, which we will develop. In it, thinking of various states $|\psi_i\rangle$ as the states $i$ of matrix mechanics, operators like $\hat{H}$ appear as matrices $H_{ij}$.

## 0.4 Bohr–Sommerfeld Quantization and the Hydrogen Atom

Quantization of the electrons in the H atom on the other hand is obtained from the Bohr–Sommerfeld quantization rules, which state that the total *action* over the domain of a state, such as a cycle of the motion of an electron around the H atom, is quantized in units of $h$,

$$\oint_{H=E} p \cdot dq = nh, \tag{0.21}$$

where $E$ is the energy of the electron. More generally, this should be true for any variable $q$ and its canonically conjugate momentum $p$, i.e.,

$$\oint p_i dq_i = nh. \tag{0.22}$$

We will study this in more detail later, but in this review we will just recall the principal details.

For the H atom, we have three simple quantizations. For $p_\phi$ (the momentum associated with the angular variable $\phi$),

$$\oint p_\phi d\phi = lh \tag{0.23}$$

leads to quantization of the total angular momentum,

$$L = l\hbar. \tag{0.24}$$

For $p_r$ (the momentum associated with the radial variable $r$),

$$\oint p_r dr = kh \tag{0.25}$$

leads to the relation ($e_0^2 \equiv e^2/(4\pi\epsilon_0)$)

$$\sqrt{\frac{2\pi^2 m e_0^4}{-E}} - 2\pi L = kh. \tag{0.26}$$

The sum of the quantum numbers $k$ and $l$ is called the principal quantum number,

$$n = l + k. \tag{0.27}$$

This then leads to the derivation of the energy levels obtained from the Balmer law,

$$E_n = -\frac{m e_0^4}{2\hbar^2 n^2}. \tag{0.28}$$

Physically, one describes the original Bohr–Sommerfeld quantization as the particle trajectory being an integral number of de Broglie wavelengths, as in Fig. 0.4.

On the other hand, the projection of the angular momentum $L$ on any direction $z$ must also be quantized (as we deduced from the Stern–Gerlach experiment), so

$$\oint L_z d\phi = mh, \tag{0.29}$$

leading to

$$L_z = m\hbar. \tag{0.30}$$

Thus the H atom is described by the three quantum numbers $(n, l, m)$.

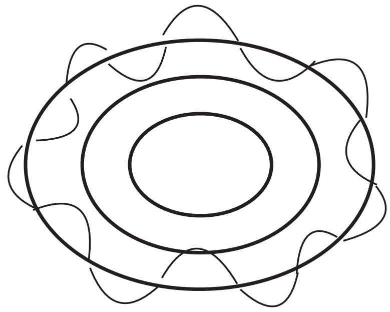

Figure 0.4 Bohr–Sommerfeld quantization, physical interpretation.

## 0.5 Review: States of Quantum Mechanical Systems

To summarize, in the case of quantum mechanical systems we have the following possibilities for the states.

- We can have a finite number of discrete states, as in the case of spin systems. *Quantization* in this case means exactly that: instead of a continuum of possible states (for a classical angular momentum, in this case), now we have a discrete set of states. In particular, for spin $s = 1/2$ (as in the case of electrons, proven by the Stern–Gerlach experiment), we have a two-state system, the simplest possible, which will be analyzed first.
- We can have an infinite number of discrete states, as in the case of the hydrogen atom, considered above. In this case, *quantization* means the same: instead of a continuum of states, we have a discrete set, with an allowed set of energies and states, as well as other observables.
- We can also have a continuum of states, as in the case of free particles. In this case, quantization refers only to the types of states allowed; there is no restriction on the allowed number of energies and states associated with them.

## Important Concepts to Remember

- Some states for some systems are quantized (quantized energy states, for instance), leading to matrix mechanics.
- There is always a wave associated with a particle and a particle with a wave, leading to a complementarity.
- There is always a (complex-valued) wave function $\psi$ describing probabilities for a measurement via $P = |\psi|^2$. It satisfies the Schrödinger equation.
- We can always have a discrete infinite number of states, or even a continuous infinite number, but both are still described by matrix mechanics. In any case, there is a wave function.

### Further Reading

See, for instance, Messiah's book [2]. It has a more thorough treatment of the historical background.

## Exercises

**(1)** Derive the Stefan–Boltzmann law from the Planck law for the spectral radiance.

**(2)** In the photoelectric effect, if the incoming flux of light has frequency $\nu$ and the released electrons get thermalized through interaction, what is the resulting temperature? If $\nu$ is within the visible spectrum and $W$ is comparable with the energy of the photon, what kind of temperature range do you expect?

**(3)** Derive the Compton effect $\Delta\lambda$ from the relativistic collision of the photon from a free electron.