Part I
Introduction

In the first part of this book, we give a general introduction and review all the aspects of classical electrodynamics and quantum mechanics that are needed in later chapters. This brief summary will help us also to agree on the notation used throughout. The reader may very well be familiar with most of the material presented here, in which case it could be skipped and referred back to if needed.

Nonrelativistic Quantum X-Ray Physics, First Edition. Stefan P. Hau-Riege.

1
Introduction

1.1
Motivation

X-ray physics has been essential throughout the last century and continues to be so to this date. It has catalyzed and survived multiple revolutions in physics, and has undergone several renaissances, usually coupled with the advent of new generations of X-ray sources. A couple of events are particularly noteworthy: In 1900, Planck provided an explanation for the spectrum that is emitted by a thermal radiation source by assuming that the radiation is quantized into energy packets of magnitude $\hbar\omega$ per mode [1, 2]. In 1905, Einstein explained the photoelectric effect by assuming that electromagnetic radiation is corpuscular [3]. It has been shown since that a semiclassical theory, which treats the electromagnetic radiation classically and only the matter system quantum mechanically, is actually sufficient to explain this effect. Nevertheless, both Planck's and Einstein's observations suggested that classical electromagnetic field theory needs to be extended to include corpuscular and nondeterministic elements. By combining the special theory of relativity with quantum physics, scientist such as Dirac [4], Feynman [5], Schwinger [6], and Tomonaga [7] developed quantum electrodynamics (QED), the quantum theory of light, which is one of the crown jewels of modern physics. In this book, we develop and apply QED in its nonrelativistic limit, as it is sufficient for many current X-ray applications.

1.2
Comparing X-Rays with Optical Radiation

For *optical* radiation, the invention of the laser [8], an acronym for light amplification by stimulated emission of radiation, led to experiments that could only be understood with a quantum theory of radiation. It thereby greatly accelerated the development of the field of quantum optics in the 1960s. We build on the achievements made in that discipline and describe their extension to the X-ray regime.

Nonrelativistic Quantum X-Ray Physics, First Edition. Stefan P. Hau-Riege.

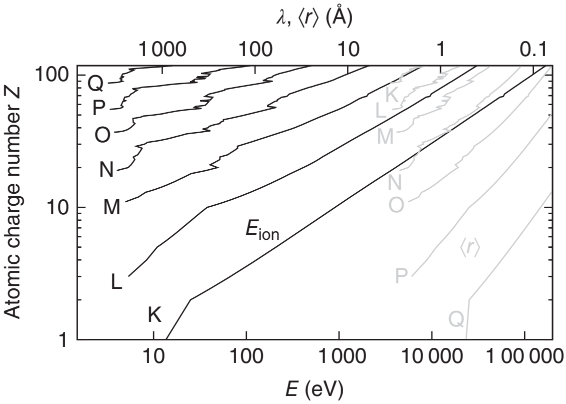

Figure 1.1 Ionization energies E_{ion}, shown as black lines and related to the bottom axis, and the radial expectation values $\langle r\rangle$, shown as grey lines related to the top axis, of neutral atoms as a function of the atomic charge number Z. The top and the bottom axes are aligned so that an energy at the bottom, interpreted as an X-ray energy, corresponds to the wavelength shown at the top.

X-ray and optical radiation fields are very similar in principle, but there are striking differences for practical applications. For example, unlike for optical radiation, X-ray wavelengths are comparable to atomic dimensions and interatomic distances, so they offer the potential to analyze structures with atomic resolution when we use techniques such as elastic X-ray scattering. The photon energy E is related to the wavelength λ and the light frequency ω through

$$E = \hbar\omega = \frac{hc}{\lambda} \quad \Rightarrow \quad E\,[\text{eV}] = \frac{12398}{\lambda\,[\text{Å}]} = 0.6582\,\omega \left[\frac{\text{rad}}{\text{fs}}\right]. \tag{1.1}$$

Here, h is the Planck constant, $\hbar = h/2\pi$ is the reduced Planck constant, and c is the speed of light. Figure 1.1 shows the spatial extent of atoms measured by the radial expectation value $\langle r\rangle$ as a function of the atomic charge number and for different atomic shells. It can be seen that $\langle r\rangle$ is of the same order of magnitude as typical X-ray wavelengths.

In the X-ray regime, *inner-shell* atomic processes tend to dominate. Excited atomic states decay quickly and exhibit an element-specific response associated with the emission of electrons or photons that are characteristic for the participating atomic shells. Also shown in Figure 1.1 are the ionization energies for different principal shells, corresponding to X-ray absorption edges. Both analyzing the relaxation products and tuning the X-ray energy to an atomic resonance and thereby identifying the atoms can be used for the analysis of materials. The details of these resonances depend also on the atomic environment, making it a useful tool to study nearest neighbor interactions. The interaction of X-rays with matter is generally relatively weak, as long as we stay away from atomic resonant energies, so that materials tend to be relatively transparent to X-rays.

1.3
Novel X-Ray Sources

Advances in the development of X-ray sources, such as synchrotrons and, more recently, X-ray free-electron lasers (FELs), continue to excite interest in the X-ray science community. We now discuss the major devices and techniques used to produce X-rays at such facilities, which are mostly based on utilizing the synchrotron radiation emitted by relativistic electron bunches in a magnetic field. If the Lorentz factor $\gamma = E/mc^2$, where E is the kinetic energy of an electron and mc^2 is its rest energy, is large, then the radiation is directed approximately tangentially to the electron orbit. The spectrum of the emitted radiation is continuous. This scenario is realized, for example, in bending magnets in a storage ring.

The characteristics of synchrotron radiation can be greatly enhanced by using insertion devices such as wigglers and undulators, leading to larger intensities and narrower spectral energy distributions of the radiation. These devices consist of periodically arranged, short dipole magnets of alternating polarity, through which electrons move in wavelike trajectories. This motion leads to the emission of nearly monochromatic radiation that is concentrated in a narrow angular cone with an opening angle of about $\pm 1/\gamma$. Calling the magnetic period in the lab frame λ_u, then, owing to relativistic contraction, the electron sees an undulator period of $\lambda_u^* = \lambda_u/\gamma$, so it oscillates at a higher frequency $\omega^* = \gamma 2\pi c/\lambda_u$ and emits dipole radiation accordingly. On Lorentz transformation of this radiation back into the lab frame, we obtain a wavelength $\lambda_{\text{light}} = \lambda_u^*/\gamma = \lambda_u/\gamma^2$. This wavelength can be adjusted by varying the kinetic energy E of the electrons. When all the electrons propagate independently, then the emitted radiation fields add incoherently, and the radiation power is proportional to number of electrons N_e.

An FEL is essentially a very long undulator in which the electrons emit *quasi-coherent*, almost-monochromatic, and well-collimated radiation. Unlike in a laser, the roles of the active laser medium and of the energy pump are both taken over by the relativistic electron bunch. For visible and infrared FELs, optical resonators can be used, and the energy gain per passage of only a few percent is sufficient. These are called *low-gain multi-pass FELs*. For vacuum-ultraviolet and X-ray FELs, such optical resonators are not an option because sufficiently efficient mirrors are not available, so that large gains per single passage are required. The theory for these high-gain, single-pass FELs is covered in References [9–13].

Without a seed laser, the FEL principle is based on self-amplification of spontaneous emission (SASE) and starts from shot noise in the FEL beam which is related to the arrival time of each electron at the undulator. This leads to a statistically fluctuating output that exhibits spikes in the intensity $I(t)$. The width of the spikes is characterized by the coherence time $\tau_c = \sqrt{\pi}/\sigma_\omega$, where σ_ω is the SASE gain bandwidth. The spectral intensity $I(\omega)$ also exhibits spikes, and the width of the spectral spikes is proportional to $1/T_b$, where T_b is the electron bunch length. SASE FEL radiation statistics and coherence are discussed more in Section 7.2.3, where we present the general properties of chaotic light sources.

1.4
Unit Systems

In this book, we use the International System of Units, abbreviated SI units, which stands for *Le Système international d'unités*. The SI system is popular and often used in engineering because it connects seamlessly to the material learned in introductory physics classes and expressions can be readily evaluated. Further, it allows checking of formulas through dimensional analyses. We now discuss how to convert other unit systems, which are also often used in theoretical physics, from and to the SI system. The SI system uses the units meter for length, kilogram for mass, and second for time (MKS system), from which other mechanical units are derived. For instance, the "newton" (N) with $1\ \mathrm{N} = 1\ \mathrm{kg\,m\ s^{-2}}$ is the unit for mechanical force. In electromagnetism, the "ampere" for the electric current was introduced independently of the MKS system, and so an additional proportionality constant is necessary to relate electromagnetic units to kinematic units. This constant is taken as the vacuum permittivity ϵ_0. All other units are derived from these base units. The values of some important physical constants in the SI system are shown in Table 1.1.

The cgs system is a variant of the SI system and measures length in centimeters, mass in grams, and time in seconds. It defines mechanical units through the force unit $1\ \mathrm{dyne} = 1\ \mathrm{g\ cm\ s^{-2}}$ and through the energy unit $1\ \mathrm{erg} = 1\ \mathrm{g\ cm^2 s^{-2}}$. The formulas for mechanical laws are the same as for the SI system. There are multiple options for the unit system of electromagnetic quantities, including the Gaussian, electrostatic, and electromagnetic cgs systems. The cgs system avoids introducing a new base unit for the electric current and derives all electromagnetic units from

Table 1.1 Physical constants in SI units.

Quantity	**Symbol**	**Value**	**Unit**
Speed of light in vacuum	c	2.9979×10^{8}	m/s
Planck constant	h	6.6261×10^{-34}	J s
Reduced Planck constant	$\hbar = h/2\pi$	1.0546×10^{-34}	J s
Fundamental charge	e	1.6022×10^{-19}	C
Vacuum permittivity	ϵ_0	8.8542×10^{-12}	F/m
Magnetic constant	μ_0	1.2566×10^{-6}	$\mathrm{N/A^2}$
Electron mass	m_e	9.1094×10^{-31}	kg
Proton mass	m_p	1.6726×10^{-27}	kg
Neutron mass	m_n	1.6749×10^{-27}	kg
Boltzmann constant	k_B	1.3806×10^{-23}	J/K
		8.6173×10^{-5}	eV/K
Avogadro constant	N_A	6.0221×10^{23}	1/mol
Bohr radius	$a_0 = 4\pi\epsilon_0\hbar^2/m_e e^2$	5.2918×10^{-11}	m
Fine structure constant	$\alpha = e^2/4\pi\epsilon_0\hbar c$	7.2974×10^{-3}	
	$1/\alpha$	137.036	
Classical electron radius	r_e	2.8179×10^{-15}	m

the cgs system based on physical laws that relate electromagnetic phenomena to mechanics. As a result, the formulas describing the laws of electromagnetism are different from those of the SI system, as discussed below.

In yet other treatments, quantities are often measured in units of physical constants. For example, mass may be measured in units of the electron mass m_e, and charge in units of the fundamental charge e. Another way of saying this is that a set of select universal physical constants are normalized to 1, meaning their numerical value in terms of these units is unity. This often greatly simplifies formulas and leads to more elegant mathematical descriptions. It also avoids introducing unit prototypes such as the "International Prototype Metre" which is located in Paris, France, at the expense of clarity, since a greater ambiguity is introduced. For example, when a variable is not clearly defined, a statement such as "$a = 10$" does not make it clear if a is a mass, a charge, or whatever, whereas in SI units, "$a = 10\,\mathrm{kg}$" clearly identifies a as a mass. Of particular importance are atomic units (au), including their two variants Hartree and Rydberg atomic units. The values of fundamental constants within these two units systems, as well as for the cgs and the SI systems, are shown in Table 1.2. Dimensionless physical constants retain their values in any unit system. For example, the value of the fine structure constant, given in the SI system as

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c}, \tag{1.2}$$

is universally

$$\alpha \approx \frac{1}{137.036}. \tag{1.3}$$

In (1.2), e is the fundamental charge and ϵ_0 is the vacuum permittivity.

We now present a general treatment of the electromagnetic units and show how the different unit systems are related. The Coulomb law gives the force on two neighboring point charges and can be written as

$$F = k_C \frac{q_1 q_2}{r^2}, \tag{1.4}$$

Table 1.2 Redefined fundamental constants in different unit systems.

Quantity	**Symbol**	**Hartree amu**	**Rydberg amu**	**Gaussian cgs**	**SI**
Electron mass	m_e	1	1/2	9.1094×10^{-28} g	9.1094×10^{-31} kg
Fundamental charge	e	1	$\sqrt{2}$	4.8032×10^{-10} esu	1.6022×10^{-19} C
Reduced Planck constant	$\hbar$	1	1	1.0546×10^{-27} erg s	1.0546×10^{-34} J s
Coulomb constant	$\frac{1}{4\pi\epsilon_0}$	1	1	1	8.9876×10^{9} F/m

where q_1 and q_2 are the charges separated by a distance r. Given that the units for F and r are already defined, the choice for the constant k_C determines the unit of the charge and with that the unit of the electric current $I = q/t$. The electric field is defined as $E = F/q$ in all unit systems. The Bio–Savart law for the magnetic field can be written as

$$\mathrm{d}\boldsymbol{B} = k_B \frac{I\mathrm{d}\boldsymbol{l} \times \boldsymbol{r}}{r^3}. \tag{1.5}$$

Equation (1.5) gives the magnitude and direction of the static magnetic field induced by an electric current I along a unit vector $\mathrm{d}\boldsymbol{l}$ at position $\boldsymbol{r}$. The choice of the constant k_B determines the unit of the magnetic field $\boldsymbol{B}$. Maxwell's equations in vacuum, independent of the unit system, are

$$\nabla \cdot \boldsymbol{E} = 4\pi k_C \rho, \tag{1.6}$$

$$\nabla \cdot \boldsymbol{B} = 0, \tag{1.7}$$

$$\nabla \times \boldsymbol{E} = -\frac{k_C}{k_B c^2} \frac{\partial \boldsymbol{B}}{\partial t}, \tag{1.8}$$

$$\nabla \times \boldsymbol{B} = 4\pi k_B \boldsymbol{J} + \frac{k_B}{k_C} \frac{\partial \boldsymbol{E}}{\partial t}, \tag{1.9}$$

where ρ is the electric charge density. The values of k_C and k_B for different unit systems are given in Table 1.3. The choice of these units has implications for all other physical laws. For example, the Schrödinger equation for an electron in a hydrogen atom is written in SI units as

$$\left(-\frac{\hbar^2}{2m_e}\nabla^2 - \frac{1}{4\pi\epsilon_0}\frac{e^2}{r}\right)\psi(\boldsymbol{r}, t) = i\hbar\frac{\partial \psi(\boldsymbol{r}, t)}{\partial t}, \tag{1.10}$$

where $\psi(\boldsymbol{r}, t)$ is the wave function and m_e is the electron mass, whereas in Hartree atomic units it simplifies to

$$\left(-\frac{1}{2}\nabla^2 - \frac{1}{r}\right)\psi(\boldsymbol{r}, t) = i\frac{\partial \psi(\boldsymbol{r}, t)}{\partial t}. \tag{1.11}$$

Table 1.4 shows the factors to convert au and cgs-Gaussian units to SI units.

Table 1.3 Electric and magnetic proportionality constants k_C and k_B in different unit systems.

Unit system	k_C	k_B
Electrostatic cgs	1	c^{-2}
Electromagnetic cgs	c^2	1
Gaussian cgs	1	c^{-1}
Hartree amu	1	1
Rydberg amu	1	1
SI	$(4\pi\epsilon_0)^{-1}$	$\mu_0/4\pi$

Table 1.4 Multiplication factors to convert to SI units.

Quantity	Hartree amu	Rydberg amu	Gaussian cgs
Length	5.2918×10^{-11} m	5.2918×10^{-11} m	10^{-2} m/cm
Mass	9.1094×10^{-31} kg	1.8219×10^{-30} kg	10^{-3} kg/g
Time	2.4189×10^{-17} s	4.8378×10^{-17} s	1
Frequency	4.1341×10^{16} Hz	2.0671×10^{16} Hz	1
Speed	2.1877×10^{6} m/s	1.0938×10^{6} m/s	10^{-2} (m/s)/(cm/s)
Momentum	1.9929×10^{-24} kg m/s	1.9929×10^{-24} kg m/s	10^{-5} (kg m/s)/(g cm/s)
Force	8.2387×10^{-8} N	4.1194×10^{-8} N	10^{-5} N/dyne
Power	1.8024×10^{-1} W	4.5059×10^{-2} W	10^{-7} W/(erg/s)
Energy	4.3597×10^{-18} J	2.1799×10^{-18} J	10^{-7} J/erg
Charge	1.6022×10^{-19} C	1.1329×10^{-19} C	3.3356×10^{-10} C/esu
Charge density	1.0812×10^{12} C/m^3	7.6453×10^{11} C/m^3	3.3356×10^{-4} (C/m^3)/(esu/cm^3)
Current	6.6236×10^{-3} A	2.3418×10^{-3} A	3.3356×10^{-10} A/(esu/s)
Current density	2.3653×10^{18} A/m^2	8.3627×10^{17} A/m^2	3.3356×10^{-6} (A/m^2)/(esu/s/cm^2)
Electric field	5.1422×10^{11} N/C	3.6361×10^{11} N/C	2.9979×10^{4} (N/C)/(dyne/esu)
Potential	2.7211×10^{1} V	1.9241×10^{1} V	2.9979×10^{2} (J/C)/(erg/esu)
Magnetic field	2.3505×10^{5} T	3.3241×10^{5} T	10^{-4} T/Gauss
Proton mass	1.6726×10^{-27} kg	3.3452×10^{-27} kg	10^{-3} kg/g
Neutron mass	1.6749×10^{-27} kg	3.3499×10^{-27} kg	10^{-3} kg/g

1.5
Overview of Lagrangian and Hamiltonian Mechanics

The formal transition from a classical to a quantum physical description of matter coupled to an electromagnetic wave field requires the introduction of several mechanical concepts, including conjugate pairs of canonical mechanical variables, the Lagrangian and Hamiltonian functions, and Poisson brackets. We begin by discussing the Lagrangian function to describe the system, and then use the principle of least action to derive the Lagrange equations. From these follow the equations of motion [14].

1.5.1
Lagrangian Mechanics

A system with N degrees of freedom is characterized by N generalized coordinates q_j and N matching generalized velocities $\dot{q}_j$. For brevity, we write $\boldsymbol{q} = (q_1, \ldots, q_N)$. The Lagrangian

$$\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = T(\dot{\boldsymbol{q}}) - V(\boldsymbol{q}, \dot{\boldsymbol{q}}) \tag{1.12}$$

fully describes system. Here, V and T are the potential and kinetic energies, respectively. This form of the Lagrangian is valid only for conservative, meaning energy-conserving, systems. To determine the trajectory, we introduce the *action*, which is a functional of the path $\boldsymbol{q}(t)$, as

$$S[\boldsymbol{q}] = \int_{t_1}^{t_2} \mathcal{L}\left(\boldsymbol{q}, \dot{\boldsymbol{q}}, t\right) \mathrm{d}t. \tag{1.13}$$

The principle of least action states that the actual trajectory $\boldsymbol{q}_{\text{act}}$ with velocity $\dot{\boldsymbol{q}}_{\text{act}}$ makes the action S extremal (stationary):

$$\begin{aligned} 0 &= \delta S[\boldsymbol{q}] && (1.14)\\ &= \int_{t_1}^{t_2} \mathcal{L}\left(\boldsymbol{q}_{\text{act}} + \delta\boldsymbol{q}, \dot{\boldsymbol{q}}_{\text{act}} + \delta\dot{\boldsymbol{q}}, t\right) - \mathcal{L}\left(\boldsymbol{q}_{\text{act}}, \dot{\boldsymbol{q}}_{\text{act}}, t\right) \mathrm{d}t && (1.15)\\ &= \int_{t_1}^{t_2} \sum_j \left(\delta q_j \frac{\partial\mathcal{L}}{\partial q_j} + \delta\dot{q}_j \frac{\partial\mathcal{L}}{\partial\dot{q}_j}\right) \mathrm{d}t && (1.16)\\ &= \int_{t_1}^{t_2} \sum_j \delta q_j \left[\frac{\partial\mathcal{L}}{\partial q_j} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_j}\right)\right] \mathrm{d}t && (1.17) \end{aligned}$$

where the slight variation $\delta\boldsymbol{q}$ can be chosen arbitrarily, except that it has to vanish at the boundaries t_1 and t_2, meaning that $\boldsymbol{q}(t_1)$ and $\boldsymbol{q}(t_2)$ are fixed. In the last step leading to (1.17), we performed an integration by parts. Since $\delta\boldsymbol{q}$ is arbitrary, we obtain the Euler–Lagrange equation

$$\boxed{\frac{\partial\mathcal{L}}{\partial q_j} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_j}\right).} \tag{1.18}$$

There are N Euler–Lagrange equations in (1.18), one for each degree of freedom, from which we can obtain the equations of motion.

1.5.2
Hamiltonian Mechanics

The Lagrangian $\mathcal{L}$ is a function of q_j and $\dot{q}_j$. In the Hamiltonian description, the system is characterized by q_j and the generalized momenta p_j, which are defined as

$$p_j = \frac{\partial\mathcal{L}}{\partial\dot{q}_j}. \tag{1.19}$$

Then it follows from (1.18) that

$$\dot{p}_j = \frac{\partial\mathcal{L}}{\partial q_j}. \tag{1.20}$$

The Hamiltonian of the system is obtained through a Legendre transform of the Lagrangian:

$$\boxed{\mathcal{H}\left(\{q_j\}, \{p_j\}, t\right) = \sum_j \dot{q}_j p_j - \mathcal{L}\left(\{q_j\}, \{\dot{q}_j\}, t\right),} \tag{1.21}$$

where we transition from the set of variables $\{q_j\}$, $\{\dot{q}_j\}$, and t to the variables $\{q_j\}$, $\{p_j\}$, and t. If V is independent of $\{\dot{q}_j\}$, then $\mathcal{H} = T + V$ is the total energy of the system because inserting (1.19) into (1.21) yields

$$\mathcal{H} = \sum_j \dot{q}_j \frac{\partial T}{\partial \dot{q}_j} - T + V = 2T - T + V = T + V \tag{1.22}$$

since T is a function of the squares of the generalized coordinates only. To derive the equations of motion, we take the total differential of the Hamiltonian:

$$\begin{aligned} \mathrm{d}\mathcal{H} &= \sum_j \frac{\partial \mathcal{H}}{\partial q_j}\mathrm{d}q_j + \frac{\partial \mathcal{H}}{\partial p_j}\mathrm{d}p_j + \frac{\partial \mathcal{H}}{\partial t}\mathrm{d}t && (1.23)\\ &= \sum_j \left(\dot{q}_j \mathrm{d}p_j + p_j \mathrm{d}\dot{q}_j - \frac{\partial \mathcal{L}}{\partial q_j}\mathrm{d}q_j - \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\mathrm{d}\dot{q}_j \right) - \frac{\partial \mathcal{L}}{\partial t}\mathrm{d}t && (1.24)\\ &= \sum_j \left(\dot{q}_j \mathrm{d}p_j - \dot{p}_j \mathrm{d}q_j \right) - \frac{\partial \mathcal{L}}{\partial t}\mathrm{d}t. && (1.25) \end{aligned}$$

In (1.24), we used (1.21), and in (1.25) we used (1.19) and (1.20). Comparing the coefficients of (1.23) and (1.25) leads to Hamilton's equations of motion:

$$\boxed{\begin{aligned} \dot{p}_j &= -\frac{\partial \mathcal{H}}{\partial q_j} = \left\{ p_j, \mathcal{H} \right\}, \\ \dot{q}_j &= \frac{\partial \mathcal{H}}{\partial p_j} = \left\{ q_j, \mathcal{H} \right\}, \end{aligned}} \qquad \begin{aligned} (1.26)\\ (1.27) \end{aligned}$$

and

$$\frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t}, \tag{1.28}$$

where the Poisson brackets $\{,\}$ are defined as

$$\{u, v\} = \sum_k \frac{\partial u}{\partial q_k}\frac{\partial v}{\partial p_k} - \frac{\partial u}{\partial p_k}\frac{\partial v}{\partial q_k}. \tag{1.29}$$

In (1.29), u and v are generic functions depending on q_k, p_k, and t. It follows from (1.26) that, if $\mathcal{H}$ is independent of q_j, then p_j is a constant of motion. Equations (1.26) and (1.27) together form a system of $2N$ first-order differential equations, whereas the Lagrangian equations of motion (1.18) are N second-order differential equations. Also,

$$\{q_i, p_j\} = \delta_{ij}, \tag{1.30}$$

$$\{p_i, p_j\} = 0, \tag{1.31}$$

where δ_{ij} is the Kronecker delta, as well as

$$\dot{u} = \sum_k \left(\frac{\partial u}{\partial q_i}\frac{\mathrm{d}q_i}{\mathrm{d}t} + \frac{\partial u}{\partial p_i}\frac{\mathrm{d}p_i}{\mathrm{d}t} \right) + \frac{\partial u}{\partial t} = \{u, \mathcal{H}\} + \frac{\partial u}{\partial t} \tag{1.32}$$

for any phase-space function $u\left(\left\{q_i\right\}, \left\{p_i\right\}, t\right)$. From (1.32), it follows in particular that

$$\dot{\mathcal{H}} = \frac{\partial \mathcal{H}}{\partial t}. \tag{1.33}$$

This means that, if $\mathcal{H}$ is not explicitly time dependent, then $\mathcal{H}$ is constant and so the total energy of the system is conserved.

1.6
Approximations

A full quantum theory for the interaction of matter and radiation, which we develop in the following chapters, is difficult or impossible to solve exactly in many cases, so many simplifications have been proposed, including the semiclassical approximation, the dipole approximation, and the perturbative approach. We now briefly describe the semiclassical and the dipole approximations, both of which are used with caution in this book for the reasons described below. The perturbation approximation, on the other hand, is used extensively and is described in Chapters 10 and 11.

1.6.1
Semiclassical Approximation

In a fully quantum-theoretical treatment, we treat both matter and radiation as quantum systems. Using the Hamiltonian formulation of electrodynamics, the electromagnetic field is quantized similar to matter. We start out with the energy density of the electromagnetic field as the Hamiltonian, and expand it classically in electromagnetic waves. Commonly, these are plane waves. The expansion parameters are the dynamical variables that describe the electromagnetic field. Finally, we transition to quantum mechanics by identifying these expansion parameters as operators that satisfy certain commutation relations.

This procedure is significantly simplified when we make the semiclassical approximation. The term *semiclassical approximation* has different meanings in different contexts. Here, it means that we treat the atomic system quantum mechanically and the radiation classically, so that the electromagnetic field is a definite function of time through Maxwell's equations. Within this approximation, we can predict transfers of discrete amounts of energy, as, for example, in the case of the photoelectric effect. This discrete transfer occurs not because the electromagnetic field is quantized but because the electron transitions involve discrete states due to the quantum properties of the electrons.

As in the semiclassical approximation, in the *neo-classical approximation* we also treat the matter system quantum mechanically and the electromagnetic field classically. In addition, we interpret the matter wave functions Ψ as the source for the electromagnetic field. The product of the electron charge $-e$ and the probability density $|\Psi|^2$ is taken as the charge density $-e\,|\Psi|^2$. Similarly, the product of the electron charge and the probability flux density is taken as a charge current density, $\boldsymbol{J} = -e\,(\Psi^* \nabla \Psi - \Psi \nabla \Psi^*)$. As in the semiclassical approximation, the quantum aspects of the predictions are again due to the quantum properties of the matter system and not due to the radiation.

The classical description of light has been very successful and agrees with a full quantum description in many instances. For example, it has been suggested to describe even spontaneous emission from an atom in the semiclassical approximation by considering radiation reaction, which is the effect of the field that is created by a charge on itself [15]. However, it is usually not clear what the limits of validity of the semiclassical theory are, and so this needs be checked on a case-by-case basis by a more vigorous theory, such as QED. As it turns out, a full quantum description is actually often elegant and conceptually simple once the formalism has been developed.

1.6.2
Dipole Approximation

In the course of this book, we use the normal mode expansion of the electromagnetic field on occasions, as, for example, in plane or spherical waves, which leads to terms of the form e^{ikr}, where $k = 2\pi/\lambda$ is the wavevector, λ is the wavelength, and r is the spatial position. In the optical regime, this expression is often expanded in a Taylor series around $kr = 0$:

$$e^{ikr} = \sum_{n=0}^{N} \frac{f^n(0)}{n!}(kr)^n + R_N(kr) \tag{1.34}$$

$$= 1 + \sum_{n=1}^{N} \frac{(ikr)^n}{n!} + R_N(kr), \tag{1.35}$$

where $f^n(0)$ denotes the nth derivative of $f(kr) = e^{ikr}$ evaluated at $kr = 0$, N is the order of expansion, and $R_N(kr)$ is the remainder term (error). Using the Lagrange form of the remainder term, it can be shown that there is a kr_1 between 0 and kr so that

$$R_N(kr) = \frac{f^{N+1}(kr_1)}{(N+1)!}(kr)^{N+1} \tag{1.36}$$

with

$$\left|R_N(kr)\right| = \frac{\left(kr_1\right)^{N+1}}{(N+1)!} < \frac{(kr)^{N+1}}{(N+1)!}. \tag{1.37}$$

Particularly useful in the optical regime is the *dipole approximation* with $N = 1$, leading in certain cases to matrix elements with the radiative characteristics of a dipole. The dipole approximation often implies strict selection rules for atomic transitions that generally do not hold for higher order approximations, for example, when we include the quadrupole terms with $N = 2$.

For the interaction of X-rays with atoms, this approximation is not always justified. For hydrogen, for example, r is on the order of

$$\langle nl|\hat{r}|nl\rangle = \frac{a_0}{2Z}\left[3n^2 - 3l(l+1)\right], \tag{1.38}$$

where $|nl\rangle$ is a hydrogen eigenfunction, n is the principal quantum number, and l is the angular quantum number. In particular for the 1s electron, $\langle nl|\, r\, |nl\rangle = a_0 \approx 0.529$ Å. For optical light with $\lambda \approx 6200$ Å, we obtain $k \approx 0.001$ Å^{-1}, so that $kr \approx 0.0005$, and $e^{ikr} \approx 1$ is an excellent approximation. For hard X-rays, on the other hand, $\lambda \approx 1$Å, therefore $k \approx 6.3$ Å^{-1}, so that $kr \approx 3.3$, and $e^{ikr} \approx 1$ is actually a poor approximation. We tend to avoid the dipole approximation in this book because it is not clear *a priori* what effect this approximation has on the validity of the solution. The shortcomings of the dipole approximation become apparent, for example, when we consider the angular distribution of the photoelectrons that are emitted when X-rays interact with matter through single-photon ionization. This is one of the dominant interaction processes when a photon collides with an atom or a molecule. The deviations from the dipole approximation are noticeable even well below 1 keV and are greatly enhanced just below the ionization threshold, as discussed in more detail in Reference [16].