1

Scope, motivation, and orientation

If one accepts gravitational forces on the Newtonian level of precision and ignores nuclear fission and fusion, then most physical phenomena on the scale of the Earth are accounted for by electrons, nuclei, and photons. Here photons play a double role: they mediate the interaction between charges, and appear freely propagating in the form of electromagnetic radiation. In their first role it often suffices to ignore all dynamical aspects and replace the photons by the effective electrostatic Coulomb interaction. Conversely, in the study of radiation phenomena, matter in the form of nuclei and electrons can mostly be replaced by prescribed macroscopic quantities like charge, current, and polarization densities. In our treatise we plan to dwell on the border area, where the interaction between photons and electrons, respectively nuclei, must be fully retained. Our goal is to discuss the dynamics of the coupled system, charges and their radiation field.

Although such a description might give the impression that we will deal with relativistic quantum electrodynamics (QED), in fact we will not even touch upon it. This theory has been devised for predicting a few very specific effects, like the anomalous g-factor of the electron, and it does so with astounding precision. Relativistic QED is, however, not well adapted to discuss, say, the fluorescence of the hydrogen atom. Thus the subject to be covered is what is commonly known as nonrelativistic quantum electrodynamics. In fact our enterprise also has a classical part. Just as in studying quantum mechanics a good grasp of classical mechanics is most useful, we believe that an understanding of classical electron theory, i.e. classical charges in interaction with the Maxwell field, serves as a solid basis for taking up the corresponding quantum theory. The classical models discussed will be semirelativistic with one exception, namely a fully relativistic theory of extended classical charges.

Classical electron theory was at the forefront of research in the early 1900s when the development of a dynamical theory of the then newly discovered electron was attempted. The basic prediction was an energy–momentum relation for

the electron (compare with chapter 4), which, however, depended on the details of the particular electron model adopted. This enterprise came to a standstill because of the advent of the theory of special relativity, which, advancing with a totally different set of arguments, required a relativistically covariant link between energy and momentum for massive particles. Classical electron theory further deteriorated simply because it had become evident that for the investigation of radiation from atoms the newly born quantum mechanics had to be used. A brief revival occurred in the struggle to formulate a consistent relativistic quantum theory for the electron–positron field coupled to the photons. The hope was that a refined understanding of the classical theory should give a hint on how to quantize and how to handle correctly the ultraviolet infinities. But as the proper quantum field theory surfaced, classical considerations faded away. In fact the theory emerged in a worse state than before as summarized in the 1963 opinion of R. Feynman: "The classical theory of electromagnetism is an unsatisfactory theory all by itself. The electromagnetic theory predicts the existence of an electromagnetic mass, but it also falls on its face in doing so, because it does not produce a consistent theory."

Because of its peculiar history, classical electron theory never had any share in the good fortune of being rewritten, modernized, and rewritten again, as can be seen from a rapid sample of standard textbooks on electrodynamics. While the conventional chapters essentially follow the same intrinsic pattern, obviously with a lot of variations on details, once it comes to the chapter on radiation reaction, Pandora's box opens. As a student I was rather dissatisfied with such a state of affairs and promised myself to come back to it at some point. The first few chapters of this treatise are my own rewriting of the classical theory. It is based on two cornerstones:

- a well-defined dynamical theory of extended charges in interaction with the electromagnetic field;
- a study of the effective dynamics of charges under the condition that they are far apart and the external potentials vary slowly on the scale given by the size of the charge distribution. This is the *adiabatic limit*.

Our approach reflects the great progress which has taken place in the theory of dynamical systems. After all, charges coupled to their radiation field can be considered as one particular case, but with some rather special features. Perhaps the most unusual one is the appearance of a center manifold in the effective dynamics, in case friction through radiation is included.

For nonrelativistic QED the situation could hardly be more different. Through the efforts made in atomic physics and quantum optics a structured theory emerged which is well covered in textbooks and reviews. It would make little sense in trying to compete with them. However, almost exclusively this theory is based either on

such drastic simplifications that an exact solution becomes possible or on second-order time-dependent perturbation theory. In recent years there has been substantial progress, mostly within the quarters of mathematical physicists, in gaining an understanding of *nonperturbative* properties of the full basic Hamiltonian, among others the structure of resonances the relaxation to the ground state through emission of photons, the nonperturbative derivation of the g-factor of the electron, and the stability of matter when the quantized radiation field is included. These and other topics will be covered in the second half of the book. Readers less interested in the classical theory may jump ahead to chapter 12, where the conclusions of chapters 2–11 are summarized and the contents of the quantum part outlined.

A few words on the style are in order. First of all, I systematically develop the theory and discuss some of the most prominent applications. No review is intended. For a subject with a long history, such an attitude looks questionable. After all, what did the many physicists working in that area contribute? To compensate, I include one historical chapter, which as very often in physics is the history as viewed from our present understanding. Since there are excellent historical studies, I hope to be excused. Further, at the end of each chapter I add *Notes and References* intended as a guide to all the material which has been left out. The level of the book is perhaps best characterized as being an advanced textbook. I assume a basic knowledge of Maxwell's theory of electromagnetism and of nonrelativistic quantum mechanics. On the other hand, the central topics are explained in detail and, for the reader to follow the discussion, there is no need of further outside sources. This brings me to the issue of mathematical rigor. In the case of classical electron theory, many claims of uncertain status are in the literature, hardly any numerical work is available, and there are no quantitative experimental verifications, as yet, with the exception of the lifetime of an electron captured in a Penning trap. More than in other fields one has to rely on fixed points in the form of mathematical theorems, which seems to be the only way to disentangle hard facts from "truths" handed down by tradition. For the quantum theory we venture into the nonperturbative regime which by definition requires a certain mathematical sophistication. In a few cases I decided to provide the full proof of the mathematical theorem. Otherwise I usually indicate its basic idea to proceed then with the formal computation. To give always full details would overload the text on an unacceptable scale and, in addition, would be duplication, since mostly the complete argument can be found elsewhere in the literature. Of course, there are stretches, possibly even long stretches, where such a firm foundation is not available and one has to proceed on the basis of limited evidence.

Our introduction might give the impression that all basic problems are resolved, nonrelativistic quantum electrodynamics is in good shape, and one only has to

turn to exciting applications. This would be a far too simplistic reading. What I hope is to bring the dynamics of charges and their radiation field properly into focus. Once this point is reached, there are many loose ends. On the theoretical side, to mention only a few of them: on the classical level, the comparison between the true microscopic and approximate particle dynamics could be more precise; a similar program for the relativistic theory of an extended charge is hardly tackled; in the quantum theory the removal of the ultraviolet cutoff at the expense of energy and mass renormalization is still not understood; and the dynamics of many charges remains largely unexplored. Also quantitative experimental confirmation of the effective dynamics of an electron, as given through the Lorentz–Dirac equation on its center manifold, remains on the agenda. The greatest reward would be if my notes encourage further research.

Part I

Classical theory

2

A charge coupled to its electromagnetic field

We plan to study the dynamics of a well-localized charge, like an electron or a proton, when coupled to its own electromagnetic field. The case of several particles is reserved for chapter 11. In a first attempt, one models the particle as a point charge with a definite mass. If its world line is prescribed, then the fields are determined through the inhomogeneous Lorentz–Maxwell equations. On the other hand, if the electromagnetic fields are given, then the motion of the point charge is governed by Newton's equation of motion with the Lorentz force as force law. While it then seems obvious how to marry the two equations, such as to have a coupled dynamics for the charge and its electromagnetic field, ambiguities and inconsistencies arise due to the infinite electrostatic energy of the Coulomb field of the point charge. Thus one is forced to introduce a slightly smeared charge distribution, i.e. an extended charge model. Mathematically this means that the interaction between particle and field is cut off or regularized at short distances, which seems to leave a lot of arbitrariness. There are also strong constraints, however. In particular, local charge conservation must be satisfied, the theory should be of Lagrangian form, and it should reproduce the two limiting cases mentioned already. In addition, as expected from any decent physical model, the theory should be well defined and empirically accurate within its domain of validity. In fact, up to the present time only two models have been worked out in some detail: (i) the semirelativistic Abraham model of a rigid charge distribution; and (ii) the Lorentz model of a relativistically covariant extended charge distribution. The aim of this chapter is to introduce and explain both models at some length. On the way we recall a few properties of the inhomogeneous Lorentz–Maxwell equations for later use.

A short preamble on units and other conventions is in order. We use the Heaviside–Lorentz units. In particular, the Coulomb potential is simply the inverse of the Laplacian with no extra factor. The vacuum susceptibilities are $\varepsilon_0 = 1 = \mu_0$, which fixes the unit of charge. c is the speed of light. Mostly we will set $c = 1$ for convenience, thereby linking the units of space and time. If needed, one can easily

retrieve these natural constants in the conventional way. At some parts below we will do this without notice, so as to have the dimensions right and to keep better track of the order of magnitudes. In the nonrelativistic setting we use $\nabla\times$ for rotation, but switch to the more proper exterior derivative, $\nabla_{\mathrm{g}}\wedge$, with $\mathbf{g}$ the metric tensor, in the relativistic context. We will use standard notation as often as possible. Since a fairly broad spectrum of material is covered, double meaning cannot be avoided entirely. At the risk of some repetition we strive for minimal ambiguity within a given chapter. In the classical part of the book we use boldface italic letters, $\boldsymbol{x}$, for three-vectors and boldface roman letters, $\mathbf{x}$, for four-vectors. In the quantum section such a notation tends to be cumbersome and we use lightface letters, x, throughout.

2.1 The inhomogeneous Maxwell–Lorentz equations

We prescribe a charge density, $\rho(\boldsymbol{x},t)$, and an associated current, $\boldsymbol{j}(\boldsymbol{x},t)$, linked through the law of charge conservation

$$\partial_t\,\rho(\boldsymbol{x},t)+\nabla\cdot\boldsymbol{j}(\boldsymbol{x},t)=0\,. \tag{2.1}$$

Of course, $\boldsymbol{x}\in\mathbb{R}^3$ and $t\in\mathbb{R}$, where we use $\mathbb{R}^3$ to describe physical space and $\mathbb{R}$ as the time axis. The Maxwell equations for the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{B}$ consist of the two evolution equations

$$\begin{aligned} c^{-1}\partial_t\boldsymbol{B}(\boldsymbol{x},t)&=-\nabla\times\boldsymbol{E}(\boldsymbol{x},t)\,,\\ c^{-1}\partial_t\boldsymbol{E}(\boldsymbol{x},t)&=\nabla\times\boldsymbol{B}(\boldsymbol{x},t)-c^{-1}\boldsymbol{j}(\boldsymbol{x},t) \end{aligned} \tag{2.2}$$

and the two constraints

$$\nabla\cdot\boldsymbol{E}(\boldsymbol{x},t)=\rho(\boldsymbol{x},t)\,,\quad \nabla\cdot\boldsymbol{B}(\boldsymbol{x},t)=0\,. \tag{2.3}$$

$\diamond$ *How are the Maxwell equations written and named?* According to my survey, there seems to be no universally accepted standard. As indicated by the name "electromagnetic", the order $\boldsymbol{E},\boldsymbol{B}$ is very common and also adopted here. In the Lagrangian version $\boldsymbol{B}$ is position-like and $-\boldsymbol{E}$ is velocity-like, which would suggest the opposite order, namely $(\boldsymbol{B},-\boldsymbol{E})$. In the nineteenth century the time-derivative was written at the right side of the equation. By present standards, in evolution equations like the Boltzmann, Navier–Stokes, and Schrödinger equation, the time-derivative is always at the left, which is also our convention here.

The common practice is to call the first equation of (2.2) together with the second equation of (2.3) the "homogeneous Maxwell equations" and the remaining

Paragraphs indicated by $\diamond$ give explanations of notation and names.

pair the "inhomogeneous Maxwell equations". We follow here the convention used in the context of wave equations and call (2.2) with $\boldsymbol{j} = 0$ the "homogeneous Maxwell–Lorentz equations" and (2.2) with $\boldsymbol{j} \neq 0$ the "inhomogeneous Maxwell–Lorentz equations". The constraints (2.3) are always understood. "Maxwell–Lorentz equations" and "Maxwell equations" are used synonymously. ◇

We solve the Maxwell equations as a Cauchy problem, i.e. by prescribing the fields at time $t = 0$. If the constraints (2.3) are satisfied at $t = 0$, then by the continuity equation (2.1) they are satisfied at all times. Thus the initial data are

$$\boldsymbol{E}(\boldsymbol{x}, 0)\,, \quad \boldsymbol{B}(\boldsymbol{x}, 0) \tag{2.4}$$

together with the constraints

$$\nabla \cdot \boldsymbol{E}(\boldsymbol{x}, 0) = \rho(\boldsymbol{x}, 0)\,, \quad \nabla \cdot \boldsymbol{B}(\boldsymbol{x}, 0) = 0\,. \tag{2.5}$$

The choice $t = 0$ is merely a convention. In some cases it is preferable to prescribe the fields either in the remote past or the distant future. We will only consider physical situations where the fields decay at spatial infinity and thus have the finite energy

$$\mathcal{E} = \frac{1}{2} \int \mathrm{d}^3x \left(\boldsymbol{E}(\boldsymbol{x}, t)^2 + \boldsymbol{B}(\boldsymbol{x}, t)^2 \right) < \infty\,. \tag{2.6}$$

In a thermal state at nonzero temperature, typical fields fluctuate without decay and one would be forced to consider infinite-energy solutions.

The Maxwell equations (2.2), (2.3) are inhomogeneous wave equations and are thus easy to solve. This will be done in Fourier space first, where the Fourier transform is denoted by $\widehat{\ }$ and defined through

$$\widehat{f}(\boldsymbol{k}) = (2\pi)^{-n/2} \int \mathrm{d}^n x\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} f(\boldsymbol{x})\,. \tag{2.7}$$

Then, setting $c = 1$, (2.2) becomes

$$\begin{aligned} \partial_t \widehat{\boldsymbol{B}}(\boldsymbol{k}, t) &= -\mathrm{i}\boldsymbol{k} \times \widehat{\boldsymbol{E}}(\boldsymbol{k}, t)\,, \\ \partial_t \widehat{\boldsymbol{E}}(\boldsymbol{k}, t) &= \mathrm{i}\boldsymbol{k} \times \widehat{\boldsymbol{B}}(\boldsymbol{k}, t) - \widehat{\boldsymbol{j}}(\boldsymbol{k}, t) \end{aligned} \tag{2.8}$$

with the constraints

$$\mathrm{i}\boldsymbol{k} \cdot \widehat{\boldsymbol{E}}(\boldsymbol{k}, t) = \widehat{\rho}(\boldsymbol{k}, t)\,, \quad \mathrm{i}\boldsymbol{k} \cdot \widehat{\boldsymbol{B}}(\boldsymbol{k}, t) = 0 \tag{2.9}$$

and the conservation law

$$\partial_t \widehat{\rho}(\boldsymbol{k}, t) + \mathrm{i}\boldsymbol{k} \cdot \widehat{\boldsymbol{j}}(\boldsymbol{k}, t) = 0\,. \tag{2.10}$$

To solve the inhomogeneous equations (2.8), we rely, as usual, on the solution of the homogeneous equations,

$$\widehat{\boldsymbol{E}}_0(\boldsymbol{k},t) = \big(\cos|\boldsymbol{k}|t + (1-\cos|\boldsymbol{k}|t)\widehat{\boldsymbol{k}}\otimes\widehat{\boldsymbol{k}}\big)\widehat{\boldsymbol{E}}(\boldsymbol{k},0) + \Big(\frac{1}{|\boldsymbol{k}|}\sin|\boldsymbol{k}|t\Big)\mathrm{i}\boldsymbol{k}\times\widehat{\boldsymbol{B}}(\boldsymbol{k},0)\,,$$
$$\widehat{\boldsymbol{B}}_0(\boldsymbol{k},t) = \big(\cos|\boldsymbol{k}|t + (1-\cos|\boldsymbol{k}|t)\widehat{\boldsymbol{k}}\otimes\widehat{\boldsymbol{k}}\big)\widehat{\boldsymbol{B}}(\boldsymbol{k},0) - \Big(\frac{1}{|\boldsymbol{k}|}\sin|\boldsymbol{k}|t\Big)\mathrm{i}\boldsymbol{k}\times\widehat{\boldsymbol{E}}(\boldsymbol{k},0)\,. \tag{2.11}$$

Here $\widehat{\boldsymbol{k}} = \boldsymbol{k}/|\boldsymbol{k}|$ is the unit vector along $\boldsymbol{k}$ and for any pair of vectors $\boldsymbol{a},\boldsymbol{b}$, $\boldsymbol{a}\otimes\boldsymbol{b}$ is the tensor of rank 2 defined through $(\boldsymbol{a}\otimes\boldsymbol{b})\boldsymbol{c} = \boldsymbol{a}(\boldsymbol{b}\cdot\boldsymbol{c})$ as acting on the vector $\boldsymbol{c}$.

We insert (2.11) in the time-integrated version of (2.8). Taking account of the constraints, making a partial integration, and using charge conservation, we arrive at

$$\begin{aligned}\widehat{\boldsymbol{E}}(\boldsymbol{k},t) &= (\cos|\boldsymbol{k}|t)\widehat{\boldsymbol{E}}(\boldsymbol{k},0) + (|\boldsymbol{k}|^{-1}\sin|\boldsymbol{k}|t)\mathrm{i}\boldsymbol{k}\times\widehat{\boldsymbol{B}}(\boldsymbol{k},0)\\ &\quad + \int_0^t \mathrm{d}s\Big(-(|\boldsymbol{k}|^{-1}\sin|\boldsymbol{k}|(t-s))\mathrm{i}\boldsymbol{k}\hat{\rho}(\boldsymbol{k},s) - (\cos|\boldsymbol{k}|(t-s))\widehat{\boldsymbol{j}}(\boldsymbol{k},s)\Big)\\ &= \widehat{\boldsymbol{E}}_{\mathrm{ini}}(\boldsymbol{k},t) + \widehat{\boldsymbol{E}}_{\mathrm{ret}}(\boldsymbol{k},t)\,, \qquad (2.12)\\ \widehat{\boldsymbol{B}}(\boldsymbol{k},t) &= (\cos|\boldsymbol{k}|t)\widehat{\boldsymbol{B}}(\boldsymbol{k},0) - (|\boldsymbol{k}|^{-1}\sin|\boldsymbol{k}|t)\mathrm{i}\boldsymbol{k}\times\widehat{\boldsymbol{E}}(\boldsymbol{k},0)\\ &\quad + \int_0^t \mathrm{d}s(|\boldsymbol{k}|^{-1}\sin|\boldsymbol{k}|(t-s))\mathrm{i}\boldsymbol{k}\times\widehat{\boldsymbol{j}}(\boldsymbol{k},s)\\ &= \widehat{\boldsymbol{B}}_{\mathrm{ini}}(\boldsymbol{k},t) + \widehat{\boldsymbol{B}}_{\mathrm{ret}}(\boldsymbol{k},t)\,. \qquad (2.13)\end{aligned}$$

The first terms are the initial fields propagated up to time t, while the second terms are the retarded fields. If one wanted to solve the Maxwell equations run into the past, then the retarded fields should be replaced by the advanced fields.

Next, let us introduce the fundamental propagator, $G_t(\boldsymbol{x})$, of the wave equation which is defined as the Fourier transform of $(2\pi)^{-3/2}\,|\boldsymbol{k}|^{-1}\sin|\boldsymbol{k}|t$ and satisfies

$$\partial_t^2 G - \Delta G = \delta(\boldsymbol{x})\delta(t)\,. \tag{2.14}$$

This means $G_t(\boldsymbol{x}) = (2\pi)^{-1}\,\delta(|\boldsymbol{x}|^2 - t^2)$ and in particular for $t > 0$

$$G_t(\boldsymbol{x}) = \frac{1}{4\pi t}\delta(|\boldsymbol{x}| - t)\,. \tag{2.15}$$

Then in physical space the solution (2.12), (2.13) of the inhomogeneous Maxwell–Lorentz equations reads as

$$\begin{aligned}\boldsymbol{E}(t) &= \partial_t G_t * \boldsymbol{E}(0) + \nabla \times G_t * \boldsymbol{B}(0) - \int_0^t \mathrm{d}s \big(\nabla G_{t-s} * \rho(s) + \partial_t G_{t-s} * \boldsymbol{j}(s)\big) \\ &= \boldsymbol{E}_{\mathrm{ini}}(t) + \boldsymbol{E}_{\mathrm{ret}}(t)\,, \end{aligned} \tag{2.16}$$

$$\begin{aligned}\boldsymbol{B}(t) &= \partial_t G_t * \boldsymbol{B}(0) - \nabla \times G_t * \boldsymbol{E}(0) + \int_0^t \mathrm{d}s \nabla \times G_{t-s} * \boldsymbol{j}(s) \\ &= \boldsymbol{B}_{\mathrm{ini}}(t) + \boldsymbol{B}_{\mathrm{ret}}(t)\,. \end{aligned} \tag{2.17}$$

Here $*$ denotes convolution, i.e. $f_1 * f_2(\boldsymbol{x}) = \int \mathrm{d}^n y f_1(\boldsymbol{x}-\boldsymbol{y}) f_2(\boldsymbol{y})$.

For later purposes it will be convenient to have a more concise notation. In matrix form, the solution of the homogeneous Maxwell–Lorentz equations can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\boldsymbol{E}(t)\\ \boldsymbol{B}(t)\end{pmatrix} = \begin{pmatrix}0 & \nabla\times\\ -\nabla\times & 0\end{pmatrix}\begin{pmatrix}\boldsymbol{E}(t)\\ \boldsymbol{B}(t)\end{pmatrix}, \quad \frac{\mathrm{d}}{\mathrm{d}t}F(t) = AF(t) \tag{2.18}$$

with the column vector $F = (\boldsymbol{E}, \boldsymbol{B})$. They have the solution

$$F(t) = U(t)F(0)\,, \quad U(t) = \mathrm{e}^{At} \tag{2.19}$$

with $U(t)$ given explicitly by the terms with subscripts 'ini' in (2.17), (2.16). If we set $g(t) = (\boldsymbol{j}(t), 0)$ as a column vector, then

$$\frac{\mathrm{d}}{\mathrm{d}t}F(t) = AF(t) - g(t)\,, \quad F(t) = U(t)F(0) - \int_0^t \mathrm{d}s\, U(t-s)g(s)\,. \tag{2.20}$$

The expressions (2.16), (2.17) remain meaningful even in case $\rho, \boldsymbol{j}$ are generated by the motion of a single point charge. Let us denote by $\boldsymbol{q}(t)$ the position and by $\boldsymbol{v}(t) = \dot{\boldsymbol{q}}(t)$ the velocity of the particle carrying charge e. Then

$$\rho(\boldsymbol{x}, t) = e\delta(\boldsymbol{x} - \boldsymbol{q}(t))\,, \quad \boldsymbol{j}(\boldsymbol{x}, t) = e\delta(\boldsymbol{x} - \boldsymbol{q}(t))\boldsymbol{v}(t)\,. \tag{2.21}$$

Upon inserting this in (2.16), (2.17) one arrives at the Liénard–Wiechert fields. Since their derivation is presented in most textbooks, we do not repeat the computation here and only discuss the result. We take the world line, $t \mapsto \boldsymbol{q}(t)$, of the particle to be given for all times. Since the particle is assumed to have a relativistic kinetic energy, $|\dot{\boldsymbol{q}}(t)| < 1$. Next we prescribe the initial data for the fields at time $t = t_0$ and take the limit $t_0 \to -\infty$ in (2.16), (2.17). Then, at a fixed space-time point $(\boldsymbol{x}, t)$, the contribution from the initial fields vanishes and the retarded fields become the Liénard–Wiechert fields. To describe them we introduce the retarded