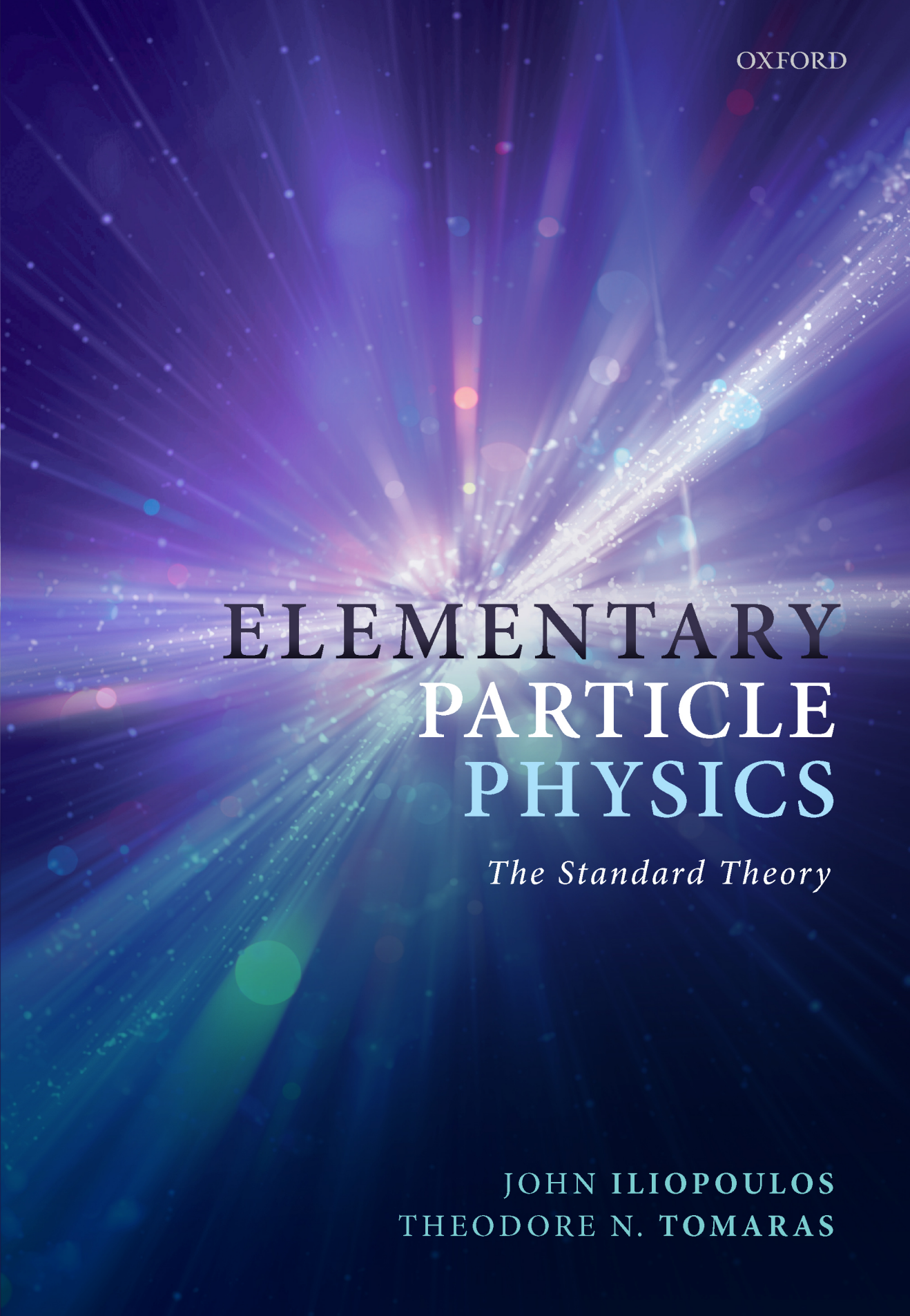
OXFORD
ELEMENTARY PARTICLE PHYSICS
The Standard Theory
JOHN ILIOPOULOS
THEODORE N. TOMARAS

Elementary Particle Physics

Elementary Particle Physics

The Standard Theory

J. Iliopoulos and T.N. Tomaras

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2021936417

ISBN 978–0–19–284420–0 (hbk.)
ISBN 978–0–19–284421–7 (pbk.)

DOI: 10.1093/oso/9780192844200.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

To Alexander, Avra and Jiarui

Contents

1
Introduction

The question of the microscopic composition of matter has preoccupied natural philosophers since at least as far back as the 5th century BC. The *atomic hypothesis*, i.e. the existence of "elementary" constituents, is attributed to Democritus of Abdera (c.460 – c.370 BC). Determining the nature of these constituents and understanding the interactions among them is the subject of a branch of fundamental physics called *the physics of elementary particles*. It is the subject of this book. In the past few decades this field has gone through a "phase transition". At first sight the critical time was somewhere between the late 1960s and the early 1970s when the theory, which became known as *the Standard Model*, was fully formulated. However, this is only part of the story and the term "phase transition" may be misleading. It took several decades of intense effort by experimentalists and theorists, both before and after the "critical point", for the various ingredients of this Model to be discovered and for its main predictions to be computed theoretically and verified experimentally. The agreement has been impressive to the extent that we should no more talk about *the Standard Model*, but rather *the Standard Theory.*

This transition has brought profound changes in our way of thinking and understanding the nature of the fundamental forces. The changes are subtle and a casual observer may miss them. Superficially the Standard Model is not fundamentally different from other models that people have considered for many years. They are all based on relativistic quantum mechanics or, as it is usually called, quantum field theory. This has been the language of elementary particle physics since the early 1930s, when Enrico Fermi introduced the notion of a quantum field associated with every elementary particle. The main new element brought by the Standard Model concerns the nature of the interactions between these various quantum fields. In the old days the interactions were chosen on purely phenomenological grounds, like the various potential functions $V(\boldsymbol{x})$ used in non-relativistic quantum mechanics. The new vision brought by the Standard Model is based on geometry: the interactions are required to satisfy a certain geometrical principle. In the physicists' jargon this principle is called *gauge invariance*; in mathematics it is a branch of differential geometry. This "geometrisation" of physics is the main legacy of the Standard Model.[1]

It is the purpose of this book to present and explain this modern viewpoint to a readership of well-motivated undergraduate students. It is our impression that, although the Standard Model is well established in elementary particle physics and is

[1] Above the entrance of Plato's Academy there was the inscription: "Μηδείς αγεωμέτρητος εισίτω μοι τη θύρα", i.e. "Let no one ignorant of geometry enter my door".

Elementary Particle Physics. John Iliopoulos and Theodore N. Tomaras, Oxford University Press.
 DOI: 10.1093/oso/9780192844200.003.0001

widely used, its underlying principles are not easily found in books that undergraduate students usually read. Our ambition is to show that this theory is more than an efficient way to compute scattering amplitudes at lowest order in the perturbation expansion. The subjects we cover and the way we choose to present them are dictated by this goal. We believe that it is time to introduce undergraduate students to these new concepts and methods. And we mean physics concepts. Mathematics will be introduced only when it is absolutely needed. The emphasis will be on the theoretical aspects, and this choice is partly due to our own competence and partly to limitations of space. A good exposition of experimental techniques would take a second volume.

The plan of the book is as follows: We start with a presentation of Dirac's theory of spontaneous emission in atomic physics. This makes it possible to introduce concepts such as the quantised radiation field, canonical commutation relations, creation and annihilation operators, gauge invariance and gauge fixing, but also transition probabilities and Fermi's golden rule, and all these in a concrete and well-defined physical context. Indeed the problem we were facing was that presenting the Standard Model required lengthy chapters of pure formalism before real physics questions could be addressed.

In the following two chapters we recall some results from classical field theory and we introduce the scattering formalism in non-relativistic quantum mechanics. Chapter 5 presents an elementary introduction to the theory of Lie groups and Lie algebras, including Lorentz and Poincaré. The student who has attended a course on group theory may go very fast through it.

Chapter 6 constitutes the introduction to the physics of elementary particles. It is phenomenological and follows, to a certain extent, the historical evolution of the field during the twentieth century. History is not our primary goal, but we believe that it helps in understanding the birth and development of new ideas. The following two chapters present, in a systematic way, the classical relativistic wave equations for fields of low spin and the attempts to use them in order to build a relativistic one-particle quantum mechanics. We derive the well known result that all these attempts point unambiguously to a system with an infinite number of degrees of freedom, i.e. to quantum field theory.

Going from a classical field theory to its quantum descendant is the object of Chapters 9–12. We decided to do it using Feynman's path integral method. Several reasons made us choose this approach, although it is somewhat unorthodox for an undergraduate textbook. First, we believe that the sum over histories, with its relations to stochastic processes, offers a more profound vision of the quantum world. Second, it is by far the most practical way to obtain the quantum theory of a non-linear constrained system, such as a Yang–Mills theory. Third, and very important, it offers the only quantisation method that is not restricted to the perturbation expansion. The path integral formulation does not assume that the coupling is weak. In fact, appropriately truncated on a space-time lattice, it becomes suitable for numerical simulations in the strong coupling regime. Non-perturbative results from lattice simulations have already reached a remarkable precision, and their agreement with the observed hadronic spectrum is impressive. Furthermore, in the coming years, with the continuing rise in computing power, the importance of these calculations is predicted to increase accordingly.

A systematic introduction to gauge invariant theories and the phenomenon of spontaneous symmetry breaking are the subjects of Chapters 13 and 14. We want to emphasise the conceptual step involved in the introduction of gauge invariant theories. What we physicists call "gauge fields" are not like any other field. The correct mathematical description is given by differential geometry, but we present a simplified version based on a formulation on a space-time lattice. Lattice field theory is poor man's differential geometry.

A book on particle physics today cannot be limited to the calculation of tree-level Feynman diagrams. The precision of the experiments is such that a meaningful comparison requires us to take into account the effects of higher orders. This cannot be done consistently without some notions from the theory of renormalisation and the renormalisation group. They can be found in Chapter 15. A simple treatment of the infrared divergences associated with massless particles is given in Chapter 17.

Chapters 16 to 21 present the Standard Model. They contain a one-loop calculation of the electron gyromagnetic ratio in quantum electrodynamics, the phenomenology of the weak interactions, the $SU(2) \times U(1)$ electroweak gauge theory and quantum chromodynamics. A discussion of neutrino physics, with the poorly understood phenomenon of neutrino oscillations, can be found in Chapter 20

We end with Chapter 22, which offers a panorama of the comparison of the Standard Model theoretical predictions with experimental measurements, including the most recent results. When the Large Hadron Collider (LHC) started operating in 2008 we were all expecting new physics to be around the corner. Today, more than a decade later, we must admit that no corner has been found. The reasons why we still believe that there must be physics beyond the Standard Model are briefly exposed in the last chapter. Finally, in an appendix we explain the notation we are using and present a collection of some useful formulae.

One of us (J.I.) has recently co-authored a book on quantum field theory.[2] Although the scope and level of this present book are different, there is some overlap in chapters that are common to both.

[2] "From Classical to Quantum Fields", by Laurent Baulieu, J.I. and Roland Sénéor, Oxford University Press 2017.

2

Quantisation of the Electromagnetic Field and Spontaneous Photon Emission

2.1 Introduction

The first great success of quantum mechanics was the accurate description of atomic spectra. However, this very success also showed its limitations. Indeed, by solving the Schrödinger equation, we find the eigenstates of the Hamiltonian which correspond to the stationary states. It follows that all levels should describe stable states of the atom. On the other hand, we know experimentally, that only the ground state is stable. All excited states decay by the emission of one or more quanta of radiation – photons. Here we shall analyse the simplest case

$$A^{(n)} \quad \rightarrow \quad A^{(0)} + \gamma \tag{2.1}$$

in which the transition to the ground state is a single-step process accompanied by the emission of one photon. $A^{(n)}$ represents the atom in the n-th excited state and $A^{(0)}$ the same atom in the ground state. The phenomenon is known as *spontaneous emission of radiation*, and it is not described by the Schrödinger equation. It was the need to compute the rate of such decays that prompted Dirac in 1927 to develop and use the quantum description of the electromagnetic field.

In quantum mechanics the time evolution of a physical system is given by the operator $U(t, t_0) = \mathrm{e}^{-\mathrm{i}H(t-t_0)}$, where H is the Hamiltonian. So, we must first find the Hamiltonian which, when applied to the state $A^{(n)}$, can yield the atom in its ground state *and* a photon. In other words, this Hamiltonian should have terms with non-zero matrix elements between states containing different numbers of particles. Since this will turn out to be a central theme in our efforts to describe the phenomena we observe in particle physics experiments, it will be useful, as a warm-up exercise, to start with this problem of atomic physics. It will allow us to introduce, in a well-defined physical context, several concepts that will be essential later.

2.2 The Principle of Canonical Quantisation

In the following we shall follow Dirac and build up a formalism, which will make it possible for particles to be created, or absorbed, as a result of the interaction. This formalism will turn out to be that of a quantised field theory and we will study first the quantum theory of the electromagnetic field.

Elementary Particle Physics. John Iliopoulos and Theodore N. Tomaras, Oxford University Press.
 DOI: 10.1093/oso/9780192844200.003.0002

Let us start with a brief reminder of the principle of canonical quantisation. It is based on the knowledge of the physical system at the classical level. Let us consider, as an example, a system with one degree of freedom. In classical mechanics it is described by a generalised coordinate $q(t)$ and its canonical conjugate momentum $p(t)$. The so-called "canonical quantisation" of this system is given, by definition, by the prescription according to which q and p are promoted to operators, acting in a certain Hilbert space, and satisfying the equal time commutation rule[1]

$$[\hat{q}(t), \hat{p}(t)] = \mathrm{i}\hbar \tag{2.2}$$

where we have used the notation $\hat{A}$ to denote the operator corresponding to the classical quantity A.[2] We recover the classical theory in the limit $\hbar \to 0$, in which the operators become commuting variables.

This prescription to pass from a classical to a quantum system generalises easily to n degrees of freedom. The commutation relation (2.2) becomes

$$[\hat{q}_I(t), \hat{p}_J(t)] = \mathrm{i}\hbar\delta_{IJ}\,, \quad [\hat{q}_I(t), \hat{q}_J(t)] = [\hat{p}_I(t), \hat{p}_J(t)] = 0\,, \quad I, J = 1, \ldots, n \tag{2.3}$$

A special case of a system given by (2.3) consists of $\alpha = 1, 2, \ldots, s$ degrees of freedom living on each site of a space lattice with N points and lattice spacing a. In that case the index I is naturally denoted as $I = \{i, \alpha\}$ with $i = 1, 2, \ldots, N$ labelling the lattice points and $n = sN$. For $s = 1$, in particular, the indices i and j denote both the lattice points and the variable at each point. We have studied such systems in statistical mechanics where we often considered the limit $N \to \infty$. We can also consider an appropriate double limit $N \to \infty$ and $a \to 0$, in which i becomes the label $\boldsymbol{x}$ of points of the spatial continuum and $q_I(t) \to q_{\boldsymbol{x},\alpha}(t)$, which we shall write conveniently as $q_\alpha(t, \boldsymbol{x})$. We thus obtain a system with s continuous infinities of degrees of freedom. In classical physics such a system is called a *classical field* and the best known example is the electromagnetic field.

If the classical theory is well defined and the canonical variables $q_\alpha(t, \boldsymbol{x})$ and $p_\beta(t, \boldsymbol{x})$ correctly identified, the quantisation of such a system is, in principle, straightforward: The commutation relation (2.3) is replaced by

$$[\hat{q}_\alpha(t, \boldsymbol{x}), \hat{p}_\beta(t, \boldsymbol{y})] = \mathrm{i}\hbar\delta_{\alpha\beta}\delta^3(\boldsymbol{x} - \boldsymbol{y}) \tag{2.4}$$

with the remaining equal time commutators of two q's, as well as of two p's equal to zero. We notice that the right-hand side of the relations (2.4) is proportional to the Dirac δ-function, which is not a "function" in the usual sense of the word. In particular, the square, or any power of it, cannot be defined. In mathematics, such generalised functions are called *distributions*. It follows that the operators $\hat{q}$ and $\hat{p}$ must also be

[1] In the usual formulation of non-relativistic quantum mechanics, the so-called *Schrödinger representation*, states $|\Psi(t)\rangle$ depend on time, and evolve according to $|\Psi(t)\rangle = U(t, 0)\,|\Psi(0)\rangle = \exp(-\mathrm{i}Ht)\,|\Psi(0)\rangle$, while observables A without explicit time dependence are represented by time-independent operators A_S. In the commutation relation (2.2) we used the *Heisenberg representation*, in which the states are fixed $|\Psi(0)\rangle = U^{-1}(t, 0)\,|\Psi(t)\rangle$ and the observables $A_H(t) = U^{-1}(t, 0)A_S U(t, 0)$ are time-dependent. At $t = 0$ the two coincide: $A_H(0) = A_S$.

[2] In order to simplify the formulae we will often drop the "hat" if there is no danger of confusion.

represented by distributions, and we expect their powers to be ill defined. We say that $\hat{q}$ and $\hat{p}$ are not *operator valued functions* of the space point, but become instead *operator valued distributions*. This leads to some mathematical difficulties, which are common to all quantum field theories and which are only partially mastered. We shall come back to this point quite often in this book.

As promised, we next apply this programme to the electromagnetic field and obtain the quantum descendant of Maxwell's theory. It will be our first example of a quantum field theory and the only one which has a well known classical limit.

2.3 The Quantum Theory of Radiation

2.3.1 Maxwell's theory as a classical field theory

The simplest version of Maxwell's equations takes the form[3]

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

$$\boldsymbol{\nabla}\wedge\boldsymbol{B}-\frac{\partial\boldsymbol{E}}{\partial t}=\boldsymbol{j}\ ,\quad \boldsymbol{\nabla}\wedge\boldsymbol{E}+\frac{\partial\boldsymbol{B}}{\partial t}=0 \tag{2.6}$$

where $\boldsymbol{E}(t,\boldsymbol{x})$ and $\boldsymbol{B}(t,\boldsymbol{x})$ are the electric and magnetic fields, respectively, while $\rho(t,\boldsymbol{x})$ and $\boldsymbol{j}(t,\boldsymbol{x})$ are the external electric charge and current densities. Consistency of (2.5), (2.6), requires that ρ and $\boldsymbol{j}$ satisfy the continuity condition: $\partial\rho/\partial t+\boldsymbol{\nabla}\cdot\boldsymbol{j}=0$. According to our recipe, if we want to describe this system as a classical field theory, we must first identify a set of independent variables $q_\alpha(t,\boldsymbol{x})$. These cannot be the six components of $\boldsymbol{E}(t,\boldsymbol{x})$ and $\boldsymbol{B}(t,\boldsymbol{x})$ because they are constrained by the first two equations (2.5). We should solve these constraints and eliminate the redundant variables.

It seems that a first step in this direction was taken by Gauss in 1835, long before Maxwell wrote his equations. It consists of introducing the vector and scalar potentials $\boldsymbol{A}(t,\boldsymbol{x})$ and $\phi(t,\boldsymbol{x})$. It will be convenient to use a compact relativistic notation in which $x=(t,\boldsymbol{x})$ and introduce the four-vectors $j^\mu(x)=(\rho,\boldsymbol{j})$ and $A^\mu(x)=(\phi,\boldsymbol{A})$. We then construct the two-index antisymmetric tensor

$$F_{\mu\nu}(x)=\frac{\partial A_\nu}{\partial x^\mu}-\frac{\partial A_\mu}{\partial x^\nu} \tag{2.7}$$

Since the derivative operator $\partial/\partial x_\mu$ will appear very often, we introduce a short-hand notation for it: $\partial/\partial x_\mu=\partial^\mu$. Similarly, $\partial/\partial x^\mu=\partial_\mu$. The electric and magnetic fields are given in terms of $F_{\mu\nu}$ by

$$F_{0i}=-\partial_0 A^i-\partial_i A^0=\left(-\frac{\partial\boldsymbol{A}}{\partial t}-\boldsymbol{\nabla}A^0\right)^i=E^i\ ,\quad B^i=\frac{1}{2}\epsilon_{ijk}F_{jk} \tag{2.8}$$

[3]We use the symbol $\wedge$ to denote the vector product of two three-dimensional vectors : $(\boldsymbol{a}\wedge\boldsymbol{b})_i=\epsilon_{ijk}\boldsymbol{a}^j\boldsymbol{b}^k$.

where $\epsilon^{ijk}(=-\epsilon_{ijk})$ is the three-index completely antisymmetric tensor, equal to $+1$ if $\{ijk\}$ form an even permutation of $\{123\}$.[4] It is now easy to check that the two inhomogeneous equations (2.6) combine to

$$\partial^\mu F_{\mu\nu}(x) = j_\nu(x) \tag{2.9}$$

while the two homogeneous ones (2.5) are automatically satisfied. The equation (2.9) follows from a variational principle applied to the action

$$S = \int d^4x\, \mathcal{L} = \int d^4x \left(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + A_\mu j^\mu\right) \tag{2.10}$$

Can we choose the four components of A_μ as independent dynamical variables? The answer is no, because the Lagrangian (2.10) does not contain the time derivative of A_0; in other words, the canonical conjugate momentum of A_0 would be identically zero. We know that this problem is related to the fact that the Lagrangian density in (2.10) is invariant under the transformation $A_\mu(x) \to A_\mu(x) + \partial_\mu\theta(x)$ with $\theta(x)$ an arbitrary function of x. In classical electrodynamics we call this invariance "gauge invariance" and we can use the freedom of choosing a particular function θ to reduce the number of independent variables. In Chapter 13 we will study this problem in a more general context, but here we just recall that, experimentally, an electromagnetic wave in empty space has only transverse degrees of polarisation. Therefore, we can impose the transversality condition $\boldsymbol{\nabla}\cdot\boldsymbol{A}(x) = 0$ (the so-called "Coulomb gauge condition") under which the zero component of the vector potential satisfies[5]

$$\Delta A_0(x) + \rho(x) = 0 \tag{2.11}$$

where Δ is the Laplacian. This implies the Coulomb law (hence the name of this condition)

$$A_0(t,\boldsymbol{x}) = \frac{1}{4\pi}\int \frac{\mathrm{d}^3x'}{|\boldsymbol{x}-\boldsymbol{x}'|}\rho(t,\boldsymbol{x}') \tag{2.12}$$

which shows that A_0 is entirely given by the external source and it is not an independent dynamical degree of freedom.[6] We are left with the spatial components, which are constrained by the Coulomb condition. The simplest way to solve it and obtain an unconstrained system is to take the three-dimensional Fourier transform

$$\boldsymbol{A}(t,\boldsymbol{x}) = \frac{1}{(2\pi)^3}\int \mathrm{d}^3k \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\widetilde{\boldsymbol{A}}(t,\boldsymbol{k}) \tag{2.13}$$

in terms of which the constraint becomes

[4] In this formula, raising and lowering the indices of three-dimensional vectors is performed using the Minkowski metric, as we explain in Appendix A.

[5] Another choice is the Lorentz covariant gauge condition of the form $\partial_\mu A^\mu(x) = 0$, the so-called "Lorenz gauge condition", first introduced by the Dane Ludvig Lorenz in 1867.

[6] Unless noted otherwise, we will assume that both the sources and the dynamical fields vanish at infinity.

$$\boldsymbol{k}\cdot\widetilde{\boldsymbol{A}}(t,\boldsymbol{k})=0 \tag{2.14}$$

which suggests to choose an orthonormal system of unit vectors $\boldsymbol{\epsilon}^{(3)}(\boldsymbol{k})=\boldsymbol{k}/|\boldsymbol{k}|$ and $\boldsymbol{\epsilon}^{(\lambda)}(\boldsymbol{k})$, $\lambda=1,2$, satisfying

$$\boldsymbol{k}\cdot\boldsymbol{\epsilon}^{(\lambda)}(\boldsymbol{k})=0\ ,\quad \boldsymbol{\epsilon}^{(\lambda)}(\boldsymbol{k})\cdot\boldsymbol{\epsilon}^{(\lambda')}(\boldsymbol{k})=\delta_{\lambda\lambda'}\ ,\quad \boldsymbol{\epsilon}^{(\lambda)}(-\boldsymbol{k})=(-)^{\lambda}\boldsymbol{\epsilon}^{(\lambda)}(\boldsymbol{k}) \tag{2.15}$$

i.e. $\boldsymbol{\epsilon}^{(3)}(\boldsymbol{k})$ is parallel to the wave vector and the other two are transverse to it. Because of the gauge condition, in this frame the vector potential has only transverse components

$$\boldsymbol{A}(t,\boldsymbol{x})=\sum_{\lambda=1}^{2}\frac{1}{(2\pi)^3}\int \mathrm{d}^3k\ \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\widetilde{A}^{(\lambda)}(t,\boldsymbol{k})\boldsymbol{\epsilon}^{(\lambda)}(\boldsymbol{k}) \tag{2.16}$$

in agreement with the experimental fact that the electromagnetic waves in empty space are transverse.

Let us summarise: Formulating the theory in terms of the vector potential A_μ, and making full use of its gauge invariance, allowed us to show that, out of the six components of $\boldsymbol{E}$ and $\boldsymbol{B}$, only two are really independent, as expected from the known properties of electromagnetic radiation. This result, which, as we shall prove later, follows from general invariance principles of the theory, made it possible to formulate electromagnetism as a dynamical system. For each $\{\boldsymbol{k},\lambda\}$, $\widetilde{A}^{(\lambda)}(t,\boldsymbol{k})$ is an independent variable. The associated canonical momentum can be computed from the Lagrangian (2.10). We find

$$q_I\to\widetilde{A}^{(\lambda)}(t,\boldsymbol{k})\ ,\quad p_I\to\partial\mathcal{L}/\partial\dot{\widetilde{A}}^{(\lambda)}(t,\boldsymbol{k})=\dot{\widetilde{A}}^{(\lambda)*}(t,\boldsymbol{k}) \tag{2.17}$$

where we have used the standard notation of mechanics in which "dot" means derivative with respect to time.

In the absence of external sources, the vector potential $\boldsymbol{A}(x)$ in the Coulomb gauge satisfies the wave equation

$$\ddot{\boldsymbol{A}}(x)-\Delta\boldsymbol{A}(x)=0 \tag{2.18}$$

If $\widetilde{\boldsymbol{A}}(k)$ is the four-dimensional Fourier transform of $\boldsymbol{A}(x)$, the wave equation (2.18) becomes an algebraic equation: $k^2\widetilde{\boldsymbol{A}}(k)=0$, which implies that $\widetilde{\boldsymbol{A}}(k)=\boldsymbol{F}(k^2)\,\delta(k^2)$ with $\boldsymbol{F}(k^2)$ arbitrary functions of k^2, provided they are regular at $k^2=0$.

The general solution of (2.18) can be expanded in plane waves, i.e. functions of the form $\mathrm{e}^{-\mathrm{i}k\cdot x}$ with $k\cdot x=k_0t-\boldsymbol{k}\cdot\boldsymbol{x}$, and with the four vector k^μ satisfying $k^2\equiv k_0^2-\boldsymbol{k}^2=0$. It will be convenient to introduce the notation

$$\mathrm{d}\Omega_m(k)=\frac{\mathrm{d}^3k}{(2\pi)^3 2E_k}=\frac{\mathrm{d}^4k}{(2\pi)^4}(2\pi)\delta(k^2-m^2)\theta(k_0)\ ,\quad E_k=\sqrt{\boldsymbol{k}^2+m^2} \tag{2.19}$$

which is the Lorentz invariant measure on the positive energy branch of the mass hyperboloid given by $k^2=m^2$. We thus write

$$\boldsymbol{A}(t,\boldsymbol{x}) = \int \mathrm{d}\Omega_0(k) \sum_{\lambda=1}^{2} \boldsymbol{\epsilon}^{(\lambda)}(\boldsymbol{k}) \left[a^{(\lambda)}(\boldsymbol{k}) \mathrm{e}^{-\mathrm{i}k\cdot x} + a^{(\lambda)*}(\boldsymbol{k}) \mathrm{e}^{\mathrm{i}k\cdot x} \right] \tag{2.20}$$

where $\mathrm{d}\Omega_0(k)$ is the value for $m = 0$ of the expression (2.19). It follows that the integration in (2.20) is over the boundary of the positive energy $k^2 = 0$ cone.

We can invert (2.20) to express the coefficient functions $a^{(\lambda)}(\boldsymbol{k})$ as

$$a^{(\lambda)}(\boldsymbol{k}) = \int \mathrm{d}^3x \ \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \boldsymbol{\epsilon}^{(\lambda)}(\boldsymbol{k}) \cdot \left[E_k \boldsymbol{A}(t,\boldsymbol{x}) + \mathrm{i}\dot{\boldsymbol{A}}(t,\boldsymbol{x}) \right] \tag{2.21}$$

In Problem 2.1 we ask the reader to verify that the energy H and momentum $\boldsymbol{P}$ of the field in terms of the coefficient functions are

$$H = \int \mathrm{d}\Omega_0(k) \ E_k \sum_{\lambda=1}^{2} \left[a^{(\lambda)*}(\boldsymbol{k}) a^{(\lambda)}(\boldsymbol{k}) \right], \quad \boldsymbol{P} = \int \mathrm{d}\Omega_0(k) \ \boldsymbol{k} \sum_{\lambda=1}^{2} \left[a^{(\lambda)*}(\boldsymbol{k}) a^{(\lambda)}(\boldsymbol{k}) \right] \tag{2.22}$$

2.3.2 Quantum theory of the free electromagnetic field–photons

We are now ready to apply our general quantisation prescription and obtain the corresponding quantum theory. The canonical variables (2.17) are promoted to operators satisfying the canonical commutation relations (2.4).[7]

$$[\widetilde{A}^{(\lambda)}(t,\boldsymbol{k}), \dot{\widetilde{A}}^{(\lambda')\dagger}(t,\boldsymbol{k'})] = \mathrm{i}\delta^{\lambda\lambda'}(2\pi)^3\delta^3(\boldsymbol{k}-\boldsymbol{k'})$$

$$[\widetilde{A}^{(\lambda)}(t,\boldsymbol{k}), \widetilde{A}^{(\lambda')}(t,\boldsymbol{k'})] = 0 \ , \quad [\dot{\widetilde{A}}^{(\lambda)\dagger}(t,\boldsymbol{k}), \dot{\widetilde{A}}^{(\lambda')\dagger}(t,\boldsymbol{k'})] = 0 \tag{2.23}$$

where $\dagger$ denotes the Hermitian adjoint of the operator. The presence of the factor $(2\pi)^3$ is due to our convention on the Fourier transform. Here, and throughout this book, we have adopted the system of units we introduce in Appendix A in which $c = \hbar = 1$. Physical units will be restored only when it is necessary.

The relations (2.23) imply for the operators corresponding to the coefficients of the plane wave expansion (2.20) the commutation relations

$$[a^{(\lambda)}(\boldsymbol{k}), a^{(\lambda')\dagger}(\boldsymbol{k'})] = \delta^{\lambda\lambda'} 2E_k (2\pi)^3 \delta^3(\boldsymbol{k}-\boldsymbol{k}')$$

$$[a^{(\lambda)}(\boldsymbol{k}), a^{(\lambda')}(\boldsymbol{k'})] = 0 \ , \quad [a^{(\lambda)\dagger}(\boldsymbol{k}), a^{(\lambda')\dagger}(\boldsymbol{k'})] = 0 \tag{2.24}$$

These relations are still formal because we have not yet defined the space in which these operators act. In order to do it we remark that, for every fixed value of $\{\boldsymbol{k}, \lambda\}$, the operators $a^{(\lambda)}(\boldsymbol{k})$ and $a^{(\lambda)\dagger}(\boldsymbol{k})$, appropriately rescaled, satisfy the commutation relations of annihilation and creation operators for a harmonic oscillator with frequency $E_k = |\boldsymbol{k}|$. In other words, we can interpret the electromagnetic field in the vacuum as a doubly (one for each value of λ) continuous infinite set of independent harmonic

[7] We omit the symbol $\wedge$ on the operators.

oscillators. Following the example of the harmonic oscillator, we define the space of states as follows:

1. First we assume the existence of a state which is annihilated by all annihilation operators $a^{(\lambda)}(\boldsymbol{k})$, for both values of λ and all $\boldsymbol{k}$. We assume this state to be unique and normalised to one. We denote it by $|0\rangle$ and we call it *the vacuum state.*

$$a^{(\lambda)}(\boldsymbol{k})\,|0\rangle = 0\ , \quad \langle 0|0\rangle = 1 \tag{2.25}$$

2. Starting from this state we build excited states by applying the creation operators on it. For example, the first excited states are given by

$$|\boldsymbol{k},\lambda\rangle = \ a^{(\lambda)\dagger}(\boldsymbol{k})\,|0\rangle \tag{2.26}$$

Using the commutation relations (2.24) and the equation (2.25), which defined the vacuum state, we find

$$\left\langle \boldsymbol{k}',\lambda' \middle| \boldsymbol{k},\lambda \right\rangle \ = \ \langle 0|\, a^{(\lambda')}(\boldsymbol{k}')a^{(\lambda)\dagger}(\boldsymbol{k})\,|0\rangle = \delta^{\lambda\lambda'} 2E_k(2\pi)^3\delta^3(\boldsymbol{k}-\boldsymbol{k'}) \tag{2.27}$$

Of course, we expect the state $|\boldsymbol{k},\lambda\rangle$, since it is a state with fixed momentum, to correspond to a wave function described by a plane wave and, therefore, to be non-normalisable. This is the meaning of the δ function on the r.h.s. of equation (2.27). As we did in quantum mechanics, we can build normalisable states using wave packets. Given a function $\Phi(\boldsymbol{k})$ satisfying

$$\int \mathrm{d}\Omega_0(k)|\Phi(\boldsymbol{k})|^2 = 1 \tag{2.28}$$

we define the wave packet

$$|\Phi,\lambda\rangle = \int \mathrm{d}\Omega_0(k)\Phi(\boldsymbol{k})\,|\boldsymbol{k},\lambda\rangle \tag{2.29}$$

which is normalised to one.

In the terminology of the harmonic oscillator, $|\boldsymbol{k},\lambda\rangle$ is the state of one excitation of type $\{\boldsymbol{k},\lambda\}$. We see that, for each value of λ, the space of the one-excitation states is our familiar Hilbert space of square-integrable complex-valued functions, which we denote by $\mathcal{H}_1$.[8]

In a similar way we build "multi-excitation" states by acting on the vacuum with products of creation operators

$$|\boldsymbol{k}_1,\lambda_1;\boldsymbol{k}_2,\lambda_2;\ldots;\boldsymbol{k}_n,\lambda_n\rangle = \ a^{(\lambda_n)\dagger}(\boldsymbol{k}_n)\ldots a^{(\lambda_2)\dagger}(\boldsymbol{k}_2)a^{(\lambda_1)\dagger}(\boldsymbol{k}_1)\,|0\rangle \tag{2.30}$$

Again, for each set of values $\{\lambda_i, i=1,...,n\}$, we denote by $\mathcal{H}_{n,\lambda}$ the Hilbert space of states in the direct product of $\mathcal{H}_1$ with itself n times. The entire space of states is

[8] More precisely, the space we use in quantum mechanics is a ray-space. Let $|\psi\rangle$ be a vector in $\mathcal{H}_1$ normalised to 1 $\langle\psi|\psi\rangle = 1$ and $C \in \mathbb{C}$ $C \neq 0$. A unit ray associated to $|\psi\rangle$ is the set of all vectors of the form $C\,|\psi\rangle$. In quantum mechanics we have learned that all vectors in this set represent the same physical state because the wave functions $\Psi(x)$ and $C\Psi(x)$ are identified.

the direct sum of all $\mathcal{H}_n$'s, for all n and all sets $\{\lambda\}$. We call this space *the Fock space of states*

$$\mathcal{F} = \sum_{n=0}^{\infty} \sum_{\{\lambda\}} \oplus \mathcal{H}_{n,\lambda} \tag{2.31}$$

where, for notational simplicity, we have defined $\mathcal{H}_0$ as the one-dimensional space spanned by the vacuum state, i.e. the ray-space of complex numbers.

The physical meaning of the states in the Fock space becomes more transparent if we express the Hamiltonian and the momentum operators of the system in terms of a and $a^\dagger$, to obtain the quantum versions of the relations (2.22). It is easy to see that, since creation and annihilation operators do not commute, we end-up with

$$H = \frac{1}{2} \int \mathrm{d}\Omega_0(k)\ E_k \sum_{\lambda=1}^{2} \left[a^{(\lambda)\dagger}(\boldsymbol{k}) a^{(\lambda)}(\boldsymbol{k}) + a^{(\lambda)}(\boldsymbol{k}) a^{(\lambda)\dagger}(\boldsymbol{k}) \right] \tag{2.32}$$

and

$$\boldsymbol{P} = \frac{1}{2} \int \mathrm{d}\Omega_0(k)\ \boldsymbol{k} \sum_{\lambda=1}^{2} \left[a^{(\lambda)\dagger}(\boldsymbol{k}) a^{(\lambda)}(\boldsymbol{k}) + a^{(\lambda)}(\boldsymbol{k}) a^{(\lambda)\dagger}(\boldsymbol{k}) \right] \tag{2.33}$$

We next use (2.32) and (2.33) to compute the energy and momentum of each state in $\mathcal{F}$. But here we face a subtle problem: Let us start with the energy of the vacuum state. As usual, it is given by the expectation value of the Hamiltonian $\langle 0| H |0\rangle$. Using the commutation relations (2.24), we obtain

$$H = \int \mathrm{d}\Omega_0(k)\ E_k \sum_{\lambda=1}^{2} \left(a^{(\lambda)\dagger}(\boldsymbol{k}) a^{(\lambda)}(\boldsymbol{k}) + \frac{1}{2} \left[a^{(\lambda)}(\boldsymbol{k})\ ,\ a^{(\lambda)\dagger}(\boldsymbol{k}) \right] \right) \tag{2.34}$$

The trouble comes from the commutator which, formally, is proportional to $\delta^3(0)$ and is, therefore, meaningless. It is easy to understand the origin of this problem, both physically and mathematically. First with the physics: We recall that the energy of a single harmonic oscillator with frequency ω is expressed as $H = \omega(a^\dagger a + 1/2)$ and its mean value in the ground state is just $\omega/2$, *the zero-point energy* of the harmonic oscillator. It is a quantum mechanical phenomenon and, as such, in physical units it is proportional to $\hbar$. We have noted already that the quantum electromagnetic field in the absence of sources is equivalent to an infinite set of harmonic oscillators satisfying the relativistic dispersion law $\omega = |\boldsymbol{k}|$. So, it is not surprising that the ground state energy is infinite; it is the infinite sum of the zero-point energies. Now with the mathematics: A divergent expression often results from a mathematical mistake; some quantity is ill defined. Indeed, in writing the Lagrangian density, or the Hamiltonian, of the classical electromagnetic field, we used expressions such as those in equation (2.10), which contain the square of the field $F_{\mu\nu}(x)$. This is fine for the classical theory. In the quantum theory, however, we have noted already that the field variables are distributions, and their square, or any higher power, is not well defined. At this stage the problem is a nuisance rather than a catastrophe. Since we understand its origin we

can easily guess its solution. In the absence of a gravitational field we never measure the absolute value of the energy of any state. Only energy differences are measurable. In other words, the expression for the energy of a system is always defined up to an arbitrary additive constant. So, the equation (2.34) should be corrected to $H \to H+C$, with C an arbitrary constant. We shall use this freedom to redefine the energy of the electromagnetic field in the absence of sources so that the vacuum energy vanishes. We define a *renormalised* energy operator by imposing the condition

$$\langle 0|\, H_{\rm ren}\, |0\rangle = 0 \tag{2.35}$$

We see that this amounts to dropping the constant divergent term in (2.34) and we find

$$H_{\rm ren} = \int \mathrm{d}\Omega_0(k)\; E_k \sum_{\lambda=1}^{2} \left(a^{(\lambda)\dagger}(\boldsymbol{k}) a^{(\lambda)}(\boldsymbol{k}) \right) \tag{2.36}$$

This is an example of a general procedure, called *renormalisation*, which allows us to attribute precise values to mathematically ill defined expressions involving products of distributions. It is based on physical requirements, such as the vanishing of the vacuum energy and it has been at the basis of all the spectacular success of quantum field theory.

This prescription can also be expressed as an ordering prescription for a product of operators. Consider an operator $\mathcal{O}$, which is the product of n creation and m annihilation operators. We shall call *normal-ordered* $\mathcal{O}$, denoted by $:\mathcal{O}:$, the expression in which we write first the n creation operators and then the m annihilation ones:

$$:\mathcal{O}: \equiv a^{(\lambda_1)\dagger}(\boldsymbol{k}_1) a^{(\lambda_2)\dagger}(\boldsymbol{k}_2) .... a^{(\lambda_n)\dagger}(\boldsymbol{k}_n) a^{(\lambda'_1)}(\boldsymbol{k}'_1) a^{(\lambda'_2)}(\boldsymbol{k}'_2) .... a^{(\lambda'_m)}(\boldsymbol{k}'_m) \tag{2.37}$$

Obviously, the definition extends to sums and products of such operators. In any product, all creation operators will be written on the left of all annihilation ones. As a result, the vacuum expectation value of a normal-ordered operator vanishes identically. We see also that for the Hamiltonian operator in particular $:H: = H_{\rm ren}$. Similarly, we define the renormalised version of the momentum operator (2.33) with $:\boldsymbol{P}: = \boldsymbol{P}_{\rm ren}$. Now it is straightforward to verify that a state in the Fock space of the form given in (2.30) is an eigenstate of the renormalised energy and momentum operators with eigenvalues $E_1+E_2+.....+E_n$ and $\boldsymbol{k}_1+\boldsymbol{k}_2+....+\boldsymbol{k}_n$, respectively. We remind ourselves also that the momenta $k_a^\mu = (E_a, \mathbf{k}_a)$ satisfy $|\boldsymbol{k}_a| = E_a$, i.e. $k_a^2 = E_a^2 - \boldsymbol{k}_a^2 = 0$. These properties allow us to give a physical interpretation to the multi-excitation states (2.30).

The state $|\boldsymbol{k}, \lambda\rangle$ has momentum $\boldsymbol{k}$ and energy $E = |\boldsymbol{k}|$, thus its 4-momentum satisfies $k^2 = 0$. Therefore, we can interpret it as describing a massless particle which we shall call a *photon*.[9] Similarly, the states (2.30) are energy and momentum eigenstates with eigenvalues corresponding to any number of non-interacting photons. Therefore, the quantised free electromagnetic Maxwell field theory is the theory of free photons.

[9]Looking at the transformation properties of the state under Lorentz transformations, we can prove that λ, defined in (2.15), corresponds to the polarisation of the one-photon state, see section 5.6.2.

Two remarks before closing this section: First, we note that, since the creation operators for any values of momentum and polarisation commute, we can apply them in any order in the formula (2.30). We conclude that the multiphoton state is totally symmetric under the interchange of any pair of photons; in other words, photons obey the Bose–Einstein statistics. The second is more general: Following this quantisation procedure, we expressed a quantum field, here the electromagnetic field $A_\mu(x)$, which satisfies a linear differential equation, as an infinite superposition of creation and annihilation operators acting in a certain Fock space. The states in this space are eigenstates of the operators of energy and momentum. This property makes it possible to interpret these states as describing non-interacting particles. This field–particle duality will be generalised and used extensively in this book. It will become clear that at a fundamental level any types of elementary particles are the quanta of a corresponding relativistic field.

2.4 Interaction of Atoms and Radiation

2.4.1 The Hamiltonian

We are now in a position to attack the problem of spontaneous emission, equation (2.1). In order to simplify the discussion we shall make some assumptions and approximations. First, we take the atom to be infinitely heavy and neglect the recoil resulting from the photon emission. Second, we choose a hydrogen-like atom with only one electron in the outer shell. The approximation consists of treating this electron as moving in an effective Coulomb potential and neglect the multibody interactions with the individual electrons of the inner shells. Third, we shall neglect all relativistic corrections for the electron.

The second assumption implies that the Hamiltonian of the atom alone is given by the usual expression

$$H_a = \frac{1}{2m}\boldsymbol{p}^2 + eU_C \tag{2.38}$$

where m is the effective mass of the electron, e its electric charge and U_C the Coulomb potential. To this expression we must add the Hamiltonian H_r of the free electromagnetic field, equation (2.36).

Finally, we must include the term that describes the interaction between the electron in the atom and the electromagnetic field. In classical electrodynamics we have found a prescription, called *minimal substitution*, which amounts to replacing everywhere the momentum p_μ by $p_\mu - eA_\mu$, with A_μ the electromagnetic vector potential. In Chapter 13 we shall give a justification of this prescription based on fundamental symmetry principles, but here we apply it starting from the Hamiltonian (2.38) with the result

$$H_I = -\frac{e}{2m}[\boldsymbol{p}\cdot\boldsymbol{A} + \boldsymbol{A}\cdot\ \boldsymbol{p}] + \frac{e^2}{2m}\boldsymbol{A}^2 \tag{2.39}$$

So, the total Hamiltonian of the system is

$$H = H_a + H_r + H_I \tag{2.40}$$

where H_a is the hydrogen-like atom Hamiltonian (2.38) with $U_C = A_0$, H_r that of the electromagnetic field in the vacuum (2.36) and H_I is the interaction Hamiltonian given by (2.39).[10]

The quantisation of this system is, in principle, straightforward. The canonical variables are the position of the electron $\boldsymbol{x}$ and its conjugate momentum $\boldsymbol{p}$, as well as the infinite set of annihilation and creation operators $a^{(\lambda)}(\boldsymbol{k})$ and $a^{(\lambda)\dagger}(\boldsymbol{k})$ of the electromagnetic field. In the Coulomb gauge, $\boldsymbol{p}$ and $\boldsymbol{A}$ commute. However, the interaction term H_I couples the electronic and the photonic degrees of freedom and, as a result, we are unable to diagonalise H exactly and find its eigenvalues and eigenstates. In fact, there is a very small number of physical problems for which we have exact solutions and for the vast majority, we have to use some approximation scheme. The most common is a perturbation expansion, which we have studied already in quantum mechanics. We recall the main features of the method here.

2.4.2 Elements of perturbation theory

In quantum mechanics the probability amplitude $\mathcal{A}_{fi}$ for the transition of a system from an initial state $|\Psi_i(t_i)\rangle$ to a final state $|\Psi_f(t_f)\rangle$ is given by

$$\mathcal{A}_{fi} = \langle \Psi_f(t_f)| \, U(t_f, t_i) \, |\Psi_i(t_i)\rangle \tag{2.41}$$

where $U(t_f, t_i) = \exp[-\mathrm{i}H(t_f - t_i)]$ is the evolution operator and H the Hamiltonian of the system. For most interesting cases this expression is only formal, because we do not know how to diagonalise the Hamiltonian and, consequently, we do not know how to compute the evolution operator U.

To proceed, let us start by addressing a general problem in which the Hamiltonian H, is the sum

$$H = H_0 + H_P \tag{2.42}$$

of a part H_0, whose spectrum is supposed to be known

$$H_0 \, |n\rangle_{(0)} = E_n^{(0)} \, |n\rangle_{(0)} \tag{2.43}$$

and a perturbation H_P. In order to simplify the discussion let us first assume that H_0 has a discrete, non-degenerate spectrum with normalisable eigenstates. We shall deal shortly with the case of the spontaneous emission, in which the spectrum of the Hamiltonian for the free electromagnetic field is continuous. Furthermore, we shall assume that the set of vectors $|n\rangle_{(0)}$ form a basis in the entire Hilbert space, so that any state can be written as a linear superposition of them. If H_0 and H_P were not operators but c-number functions, the evolution operator could be written as

[10] We see here the physical interpretation of the constant e: it is the electric charge of the particle, which characterises the strength of its coupling with the electromagnetic field. We shall call it *the coupling constant*, a terminology we shall use more generally, when we want to speak about the strength of an interaction.

$U(t) = U_0(t)U_P(t) = \exp[-\mathrm{i}H_0 t]\exp[-\mathrm{i}H_P t]$. However, for operators this is not true because $\exp[A]\exp[B] \neq \exp[A+B]$. Nevertheless, let us define the unitary operator

$$\widetilde{U}_P(t) = U_0^{-1}(t)U(t) \tag{2.44}$$

Both $U(t)$ and $U_0(t)$ satisfy the Schrödinger equation with Hamiltonian H and H_0, respectively, and it is straightforward to derive the equation satisfied by $\widetilde{U}_P(t)$

$$\mathrm{i}\frac{\partial \widetilde{U}_P}{\partial t} = -U_0^{-1}\mathrm{i}\frac{\partial U_0}{\partial t}U_0^{-1}U + U_0^{-1}(H_0+H_P)U_0\widetilde{U}_P = \widetilde{H}_P\widetilde{U}_P \tag{2.45}$$

where

$$\widetilde{H}_P(t) = U_0^{-1}(t)H_P U_0(t) \tag{2.46}$$

Equation (2.45), together with the initial condition $\widetilde{U}_P(0) = \mathbf{1}$, can be written as the integral equation

$$\widetilde{U}_P(t) = \mathbf{1} - \mathrm{i}\int_0^t \widetilde{H}_P(t')\widetilde{U}_P(t')\mathrm{d}t' \tag{2.47}$$

It is this equation that we want to solve perturbatively in powers of H_P. Once $\widetilde{U}_P(t)$ is known we can reconstruct the total evolution operator from the definition $U(t) = U_0(t)\widetilde{U}_P(t)$. We find

$$\widetilde{U}_P(t) = \mathbf{1} - \mathrm{i}\int_0^t \widetilde{H}_P(t_1)\mathrm{d}t_1 - \int_0^t\int_0^{t_1} \widetilde{H}_P(t_1)\widetilde{H}_P(t_2)\mathrm{d}t_1\mathrm{d}t_2 + \ldots \tag{2.48}$$

which gives, for $U(t)$,

$$U(t) = \mathrm{e}^{-\mathrm{i}H_0 t}\left(\mathbf{1} - \mathrm{i}\int_0^t \mathrm{e}^{\mathrm{i}H_0 t_1}H_P\,\mathrm{e}^{-\mathrm{i}H_0 t_1}\mathrm{d}t_1 + \ldots\right) \tag{2.49}$$

In this formula H_P is the perturbation Hamiltonian in the Schrödinger representation.

2.4.3 The transition probability

Let us assume that at $t = t_i$ the system is in an eigenstate $|i\rangle$ of H_0. Under the influence of the perturbation, at time $t = t_f$ the system will be in some state $|f\rangle$, which, under our assumptions, we can write as a linear superposition of the eigenstates of H_0. We would like to compute the probability to find the system in a particular eigenstate of H_0. It will be given by the square of the amplitude (2.41):

$$\mathcal{P}_{fi} = |\,\langle f|\,U(t_f,t_i)\,|i\rangle\,|^2 \tag{2.50}$$

We wish to compute this matrix element in perturbation theory using the expansion (2.49). In all interesting cases, the initial and the final states are different and

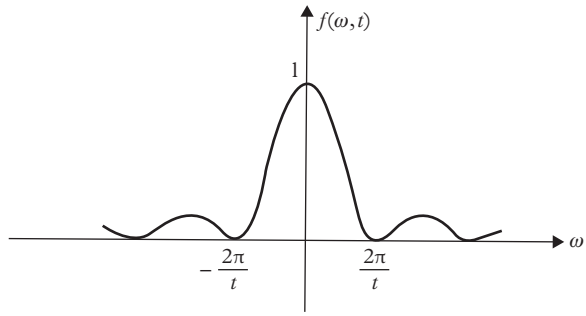

Fig. 2.1 The function $f(\omega, t)$.

orthogonal to each other, so the zeroth order term, the unit operator, drops out. At first order we obtain

$$\mathcal{P}_{fi} = \left| \int_0^t \mathrm{e}^{\mathrm{i}(E_i - E_f)t_1} \, \langle f | \, H_P \, | i \rangle \, \mathrm{d}t_1 \right|^2 \tag{2.51}$$

where we have set $t_i = 0$ and $t_f = t$.

In our problem of spontaneous emission, as well as in most physically interesting problems, the perturbation Hamiltonian is time-independent. In this case the integration over t_1 is trivial and gives

$$\mathcal{P}_{fi} = | \, \langle f | \, H_P \, | i \rangle \, |^2 f(\omega, t) t^2 \tag{2.52}$$

where $\omega = E_i - E_f$ and

$$f(\omega, t) = \left(\frac{\sin\gamma}{\gamma} \right)^2 \tag{2.53}$$

with $\gamma = \omega t/2$. This is shown graphically in Figure 2.1. The maximum around $\omega = 0$ gets more and more pronounced as t increases. In fact, we can show that, as a distribution, f satisfies

$$\lim_{t \to \infty} t f(\omega, t) = 2\pi\delta(\omega) \tag{2.54}$$

Contrary to what we have assumed up to now, the unperturbed Hamiltonian $H_0 = H_a + H_r$ of our problem, describing the atom and the free electromagnetic radiation, has a continuous spectrum. We go around this problem by imagining that we quantise the electromagnetic field in a large box of size L with periodic boundary conditions. The spectrum is discrete but, if L is large enough, the energy difference between adjacent levels can be taken to be arbitrarily small, much smaller than the experimental resolution. So, the sensible quantity to compute is the transition probability of the system from the state $|i\rangle$ at $t = t_i$ to any state belonging to a set F of states with energies in the interval $(E_f, E_f + \Delta E)$ at a much later time $t = t_f$, i.e.

$$\mathcal{P}_{i \to F} = \sum_{f \in F} \mathcal{P}_{fi} \tag{2.55}$$

It will prove convenient to replace the sum in (2.55) by an integral, but, in order to do that, we must assume that the matrix element $\langle f|\, H_P\, |i\rangle$ for all $f \in F$, depends only on E_f. Let $\rho_F(E_f)dE_f$ be the number of states $f \in F$ with energies between E_f and $E_f + dE_f$. Then the probability (2.55) can be written as

$$\mathcal{P}_{i\to F} = \int_{E_f}^{E_f+\Delta E} \mathrm{d}E_f \rho_F(E_f) |\, \langle f|\, H_P\, |i\rangle\, |^2 f(\omega, t) t^2 \tag{2.56}$$

Because of the property (2.54), when t tends to infinity, the transition strictly conserves energy, but in such a case both the physical and the mathematical problems are not well defined. Indeed, this calculation shows that the excited state of the atom has a finite lifetime, so we cannot perform infinite time experiments in such a system. On the other hand, equation (2.56) shows that the perturbation expansion will break down for very large times. For a finite time, let us call it Δt, we expect to have an uncertainty in energy of the order of Δt^{-1}. This is a result of classical wave mechanics supplemented with the quantum mechanical condition that relates energy to frequency. Notice also that this result is, to first order in perturbation, independent of the strength of the perturbation.

Thus, let us assume that t is large enough for (2.54) to be approximately valid. A non-zero result for $\mathcal{P}_{i\to F}$ requires that E_i belongs to the interval $(E_f, E_f + \Delta E)$. Finally, dividing by t we obtain for the transition probability per unit time

$$\lambda_{i\to F} = \frac{\mathcal{P}_{i\to F}}{t} = 2\pi \Big[\rho_F(E_f) |\, \langle f|\, H_P\, |i\rangle\, |^2\Big]_{E_f=E_i} \tag{2.57}$$

This formula is called *Fermi's golden rule*. The transition probability per unit time is constant, independent of t, with dimensions of energy. The sum of $\lambda_{i\to F}$ over all possible final states F gives what is called the "total decay rate" Γ of the unstable system in the state $|i\rangle$, which in turn leads to the well known "Law of Radioactivity". Indeed, if at time t we have a number $N(t)$ of unstable objects (excited atoms, nuclei, or anything else) with decay probability per unit time given by a constant Γ, in the time interval $(t, t + dt)$ its population will change by $dN(t) = -\Gamma N(t) dt$.[11] Upon integration with initial condition $N(0) = N_0$ we obtain

$$N(t) = N_0\, \mathrm{e}^{-\Gamma t} \tag{2.58}$$

In the special case at hand, we see that Γ is proportional to the square of the matrix element of the perturbation Hamiltonian, i.e. the square of the coupling constant e, which characterises the strength of the interaction. It is a straightforward exercise in probability theory to show that $\tau = \Gamma^{-1}$, is *the mean lifetime*, or simply *the lifetime* of the excited state, a concept first introduced by Rutherford in 1900.

In non-relativistic quantum mechanics we have established the result that the time evolution of a stationary state with energy E_i is given by $\Psi(t, \boldsymbol{x}) = \Psi(0, \boldsymbol{x}) \exp(-\mathrm{i} E_i t)$.

[11] If $dN(t)$ objects "succeeded" to decay in the time interval dt out of the total $N(t)$ which "tried" to decay, the probability of success, i.e. decay, per unit time is $dN(t)/(N(t)dt)$.

We can incorporate the result we just obtained by saying that, at first order in perturbation theory, the time evolution is corrected to become

$$\Psi(t, \boldsymbol{x}) = \Psi(0, \boldsymbol{x}) \mathrm{e}^{-\mathrm{i}(E_i - \mathrm{i}\Gamma/2)t} \tag{2.59}$$

It is in this sense that we often say that excited states correspond to complex values of the energy. We shall give a more rigorous justification of this terminology when we discuss resonance phenomena in scattering experiments.

Before closing this general discussion, let us come back to the conditions for the validity of (2.57) and (2.58). We have noted already that, for the first order perturbation theory to be valid, t must be smaller than $\tau = \Gamma^{-1}$. On the other hand, since Γ is given in terms of the matrix elements of the perturbation Hamiltonian, we must have $E_i \gg \Gamma$. Finally, for the delta-function approximation we used in deriving the equation (2.57) to be valid, t must be much larger than $\omega = E_i - E_f$, which is bounded by the experimental resolution ΔE. We conclude that Fermi's golden rule is valid if the following inequalities are satisfied

$$E_i \gg \Gamma \gg \Delta E \tag{2.60}$$

which seem to be intuitively obvious. We often say that they are "obviously" satisfied in real experiments, but we shall see that, at least in one case, namely in the detection of the so-called J/ψ resonance, it was crucial to make sure that they were indeed satisfied.

2.4.4 Application to the problem of spontaneous emission

After all this general discussion, let us return to our problem of spontaneous emission. We start by splitting the Hamiltonian (2.40) into an unperturbed part H_0 and a perturbation: $H = H_0 + H_P$. H_0 must be exactly solvable; in other words we must be able to find its eigenvalues and eigenvectors exactly, so we choose

$$H_0 = H_a + H_r \ , \quad H_P = H_I \tag{2.61}$$

in the notation of the equations (2.40) and (2.39). We denote the eigenvectors of H_0 by $|\mathcal{E}_m, \{n(\boldsymbol{k}, \lambda)\}\rangle$ in an obvious notation in which $\mathcal{E}_m$ is the energy of the m-th level of the atom and $n(\boldsymbol{k}, \lambda)$ the number of photons with momentum $\boldsymbol{k}$ and polarisation λ. In other words, the space of states of H_0 is the direct product of the Hilbert space of the atomic eigenfunctions and the Fock space of the photons.

$$\begin{aligned} H_a \, |\mathcal{E}_m, \{n(\boldsymbol{k}, \lambda)\}\rangle &= \mathcal{E}_m \, |\mathcal{E}_m, \{n(\boldsymbol{k}, \lambda)\}\rangle \\ H_r \, |\mathcal{E}_m, \{n(\boldsymbol{k}, \lambda)\}\rangle &= \sum_{\lambda, \boldsymbol{k}} n(\boldsymbol{k}, \lambda) |\boldsymbol{k}| \, |\mathcal{E}_m, \{n(\boldsymbol{k}, \lambda)\}\rangle \end{aligned} \tag{2.62}$$

Two remarks:

1) When we set up the quantisation prescription for the operator $\boldsymbol{A}$ of the electromagnetic field, we used the Heisenberg representation in which the operator is time

dependent: $\boldsymbol{A}(t, \boldsymbol{x})$. However, the expressions that appear in the Hamiltonian (2.39) are in the Schrödinger representation. Therefore we must use $\boldsymbol{A}(0, \boldsymbol{x})$.

2) We shall use the parameter e, the electric charge of the electron, to organise the terms in the perturbation expansion. Terms multiplied by e^1, e^2, ... are respectively terms of first order, second order, and so on. It follows that, at first order, we can keep in (2.39) only the term that is linear in $\boldsymbol{A}$.

We must now compute the matrix element of H_I between the initial and the final states

$$\langle f|\, H_I\, |i\rangle = \frac{e}{2m_e}\, \langle f|\, \boldsymbol{p}\cdot\boldsymbol{A} + \boldsymbol{A}\cdot\boldsymbol{p}\, |i\rangle = \int \mathrm{d}^3x\, \langle \{n_f\}|\, \boldsymbol{A}(0,\boldsymbol{x})\, |\{n_i\}\rangle \cdot \boldsymbol{j}_{fi}(\boldsymbol{x}) \tag{2.63}$$

where $|\{n_i\}\rangle$ and $|\{n_f\}\rangle$ are the initial and final states of the Fock space of free photons and $\boldsymbol{j}_{fi}(\boldsymbol{x})$ is given by the matrix element of the current operator $\boldsymbol{j}$ between the initial and final atomic wave functions, i.e.

$$\boldsymbol{j}_{fi}(\boldsymbol{x}) = \frac{e}{2\,\mathrm{i}\,m_e}\left[\Psi_f^*\boldsymbol{\nabla}\Psi_i - (\boldsymbol{\nabla}\Psi_f^*)\Psi_i\right] \tag{2.64}$$

The expansion (2.20) of the electromagnetic potential $\boldsymbol{A}(0, \boldsymbol{x})$ in terms of creation and annihilation operators shows that the matrix element (2.63) vanishes unless the number of photons in the final and the initial states differs by ± 1. Therefore this expression contributes only to the processes of emission or absorption of a single photon.

In our case of spontaneous emission we have $n_i = 0$ and $n_f = 1$. The calculation of the matrix element is now straightforward, but for the fact that, in the continuum spectrum, the normalisation of the one photon state contains the factor $[2\omega_k(2\pi)^3\delta^3(0)]^{-1/2}$; see equation (2.27). It is simpler if we quantise the theory in a cubic box of size L, with arbitrarily large but finite L with periodic boundary conditions. In this case the factor $(2\pi)^3\delta^3(0)$ is replaced by L^3. The final result is

$$\langle f|\, H_I\, |i\rangle = \int \frac{\mathrm{d}^3x}{\sqrt{2\omega_q L^3}} \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{x}} \boldsymbol{\epsilon}^{(\lambda)}(\boldsymbol{q}) \cdot \boldsymbol{j}_{fi}(\boldsymbol{x}) \tag{2.65}$$

where $\boldsymbol{q}$ and λ are the momentum and polarisation of the emitted photon. In principle we have completed the computation because the wave functions of the atom in the initial and the final states are supposed to be known. However, using some reasonable approximations, we can further simplify this expression and make the comparison with experiment easier. First, let us note that the current matrix $\boldsymbol{j}_{fi}(\boldsymbol{x})$ vanishes exponentially fast outside a sphere of roughly the Bohr radius, so in the exponential of (2.65), $\boldsymbol{x}$ is on the order of the Bohr radius which is $1/m_e\alpha$, where $\alpha \equiv e^2/4\pi \simeq 1/137$. Furthermore, $|\mathbf{q}| \simeq \mathcal{E}_i - \mathcal{E}_f$, which is of order $m_e\alpha^2$. So, $\boldsymbol{q}\cdot\boldsymbol{x}$ is of order α and so, to leading order in our computation, we should approximate the exponential by unity. Second, let us replace the current operator (2.64) by the electron electric dipole moment operator $\boldsymbol{d} = e\,\boldsymbol{x}$. This can be done by using the commutation relation

$$[H_a, \boldsymbol{x}] = \left[\frac{\boldsymbol{p}^2}{2m_e}, \boldsymbol{x}\right] = -\frac{\mathrm{i}}{m_e}\boldsymbol{p} \tag{2.66}$$

so that the spatial integral of the matrix element of the current takes the form

$$\int \mathrm{d}^3x\, \boldsymbol{j}_{fi}(\boldsymbol{x}) = \frac{e}{m_e}\langle\Psi_f|\,\boldsymbol{p}\,|\Psi_i\rangle = \mathrm{i}\,\langle\Psi_f|\,[H_a, \boldsymbol{d}]\,|\Psi_i\rangle = \mathrm{i}(\mathcal{E}_f - \mathcal{E}_i)\boldsymbol{d}_{fi} \tag{2.67}$$

Let us now compute the density of states ρ_F. Again, it is simpler if we quantise the electromagnetic field in a box. Every component of $\boldsymbol{q}$ takes the values $q_i = (2\pi n_i)L^{-1}$, with integer $n_i = 1, 2, 3, \ldots$. Therefore, the volume in momentum space corresponding to each photon state is $(2\pi)^3/L^3$, and the number of photon states within the volume element $q^2\Delta E_f \Delta\Omega$ with $q = |\boldsymbol{q}| \simeq \mathcal{E}_i - \mathcal{E}_f$ is

$$\Delta N_F = \frac{L^3 q^2\, \Delta E_f\, \Delta\Omega}{(2\pi)^3} \tag{2.68}$$

The density of states, up to terms of order $\Delta E_f/q$, is

$$\rho_F(E_f) = \frac{\Delta N_f}{\Delta E_f} = \frac{L^3 q^2 \Delta\Omega}{(2\pi)^3} \tag{2.69}$$

Putting everything together, we obtain that the transition probability per unit time of an atom in an excited state Ψ_i to the same atom in a state Ψ_f with the simultaneous emission of a photon inside a solid angle $\Delta\Omega_{\boldsymbol{q}}$ and with polarisation vector $\boldsymbol{\epsilon}^{(\lambda)}(\boldsymbol{q})$ is

$$\Delta\lambda_{i\to F} = \frac{\omega_q^3}{8\pi^2}\left|\boldsymbol{\epsilon}^{(\lambda)}(\boldsymbol{q})\cdot\boldsymbol{d}_{fi}\right|^2 \Delta\Omega_{\boldsymbol{q}} \tag{2.70}$$

In order to compute the *intensity* ΔI of the emitted radiation, i.e. the energy emitted per unit time in the solid angle $\Delta\Omega_{\boldsymbol{q}}$, we multiply (2.70) by ω_q and sum over the two polarisations making use of the relation $\sum_\lambda \epsilon_k^{(\lambda)}(\boldsymbol{q})\,\epsilon_l^{(\lambda)*}(\boldsymbol{q}) = \delta_{kl} - q_k q_l/q^2$.[12] We find

$$\Delta I = \frac{\omega_q^4}{8\pi^2}\sin^2\theta\,|\boldsymbol{d}_{fi}|^2\Delta\Omega_{\boldsymbol{q}} \tag{2.71}$$

where θ is the angle between $\boldsymbol{d}_{fi}$ and $\boldsymbol{q}$. In physical units this expression becomes

$$\Delta I = (\omega_q^4/8\pi^2c^3)\sin^2\theta|\boldsymbol{d}_{fi}|^2\,\Delta\Omega_{\boldsymbol{q}} \tag{2.72}$$

We observe that ΔI is independent of $\hbar$. Furthermore, it is identical to the expression for the intensity emitted by an oscillating classical dipole. This explains the success of the old Thomson model in describing the phenomenon of atomic radiation.

Formally, we can continue the calculation and compute the higher order terms in the perturbation expansion, but we will not do it here; first, because they present some technical difficulties which we will study in Chapter 15 and second because they are smaller numerically than the relativistic corrections which we have not included. For a qualitative discussion, see Problem 2.4.

[12] The left-hand side is a transverse two-index symmetric tensor under rotations, which depends only on the vector $\boldsymbol{q}$ and has trace equal to 2. The right-hand side is the unique expression with these properties.

2.5 Problems

Problem 2.1 Compute the total energy and momentum of the electromagnetic field in terms of the coefficient functions $a^{(\lambda)}(\boldsymbol{k})$ and derive the expressions (2.22).

Problem 2.2 *Emission and absorption of radiation.* In subsection 2.4.4 we computed the probability per unit time for the transition of an atom from the initial state Ψ_i to the final Ψ_f, accompanied by the emission of a photon with polarisation λ and momentum in the solid angle $\Delta\Omega_{\boldsymbol{q}}$ around $\boldsymbol{q}$ with $|\boldsymbol{q}| = \mathcal{E}_i - \mathcal{E}_f$.

1. Compute the probability per unit time $\Delta\lambda^{\rm em}$ for the same atomic transition, assuming that there are $N_{\boldsymbol{q}\lambda}$ photons with the above characteristics present in the initial state. Show, in particular, that $\Delta\lambda^{\rm em}$ is proportional to $N_{\boldsymbol{q}\lambda} + 1$. The 1 corresponds to the spontaneous emission, and the rest defines the *induced* emission. So, $\Delta\lambda^{\rm em} = \Delta\lambda^{\rm ind} + \Delta\lambda^{\rm spont}$.

2. Compute the probability per unit time $\Delta\lambda^{\rm abs}$ of the inverse atomic process (absorption), i.e. the transition of the atom from Ψ_f to Ψ_i with the absorption of one of the $N_{\boldsymbol{q}\lambda}$ present initially. Show that they satisfy the relations (*A. Einstein, 1916*)

$$\Delta\lambda^{\rm abs} = \Delta\lambda^{\rm ind} = N_{\boldsymbol{q}\lambda}\,\Delta\lambda^{\rm spont}\,, \qquad \frac{\Delta\lambda^{\rm em}}{\Delta\lambda^{\rm abs}} = \frac{N_{\boldsymbol{q}\lambda}+1}{N_{\boldsymbol{q}\lambda}}$$

Problem 2.3 The purpose of this problem is to study the properties of the electromagnetic field in a cavity with the help of the formalism of coherent states.

I. Introduction of the set of coherent states. Consider a one-dimensional harmonic oscillator with frequency ω and let a and $a^\dagger$ be the annihilation and creation operators, respectively, and $|n\rangle$, $n = 0, 1, \ldots$ the eigenvectors of the operator $N = a^\dagger a : N\,|n\rangle = n\,|n\rangle$.

1. Show that to every complex number z, there corresponds a normalised state $|z\rangle$, eigenstate of the operator a with eigenvalue z. Construct explicitly $|z\rangle$ as superposition of the states $|n\rangle$. We shall call the states $|z\rangle$ *coherent states*. Find the coherent state which corresponds to the complex number $z = 0$.

2. Show that $|z\rangle$ can be written as : $|z\rangle = \exp[-\frac{|z|^2}{2}]\,\exp(za^\dagger)\,|0\rangle$.

3. If $|z_1\rangle$ and $|z_2\rangle$ are two coherent states, find the value of the scalar product $\langle z_1|z_2\rangle$.

4. Give the physical meaning of the real number $|z|^2$.

5. Consider the operator $D(z) = \exp[za^\dagger - z^* a]$, where z^* is the complex conjugate of z. Show that D is unitary and satisfies the relations:
$D^{-1}(z)aD(z) = a + z\mathbb{1}$ and $D(z_1)\,|z_2\rangle = |z_1 + z_2\rangle$, where $\mathbb{1}$ is the unit operator. Because of these properties we call D *the translation operator*.

II. Physical meaning of the coherent states.

1. Assuming that at $t = 0$ the oscillator is in the coherent state
$|\Psi(t = 0)\rangle = |z\rangle$, find its state at a later time t.

2. Study the evolution of a classical harmonic oscillator which, at $t = 0$, is characterised by $q(0)$ and $p(0)$. Find $q(t)$ and $p(t)$ and compare with the evolution of the coherent state.

3. Consider the harmonic oscillator subject to a time-dependent external force. We write the Hamiltonian as:

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 - qF(t) = H_0 + H_1$$

where $H_0 = \omega(a^\dagger a + 1/2)$ and $H_1 = -(2m\omega)^{-1/2}(a^\dagger + a)F(t)$.

Show that the evolution operator of the system in the interaction representation, $U_I(t)$ equals, up to a phase, the translation operator $D(z(t))$ and compute $z(t)$. Assuming that at $t = 0$ the system is in the fundamental state $|0\rangle$, find its state $|\Psi(t)\rangle$ at a later time t.

III. Coherent states of the electromagnetic field. Consider the electromagnetic field quantised in a cubic box of dimension L.

1. Define the coherent states a such an electromagnetic field.

2. If $|\Psi(0)\rangle$ is the coherent state describing the state of the field at $t = 0$, find its state $|\Psi(t)\rangle$ at a later time t.

Problem 2.4 *The $\mathcal{O}(e^2)$ terms in the perturbation expansion (2.49).* In the notation of section 2.4 we consider an initial state $|\mathcal{E}_i, n_i = 0\rangle$ and a final state $|\mathcal{E}_f, n_f\rangle$.

1. Write the expression of the matrix element.

2. Show that the possible values of n_f are 0 or 2 and give the physical meaning for each value.

3. Are there any measurable effects corresponding to transitions with $n_f = 0$?

Hint: Think of the case in which the unperturbed Hamiltonian has a degenerate energy level.

Problem 2.5 *"Atom" and photons in a cavity.*

1. Consider a free two-state "atom" at rest: a system with just two states, the ground state $|g\rangle$ with energy set to zero and an excited state $|e\rangle$ with energy E.

We define the operators b and $b^\dagger$ which act in the two-dimensional space spanned by the vectors $|g\rangle$ and $|e\rangle$ as follows:

$$b\,|g\rangle = 0\ ,\quad b\,|e\rangle = |g\rangle\ ,\quad b^\dagger\,|g\rangle = |e\rangle\ ,\quad b^\dagger\,|e\rangle = 0$$

Prove that b and $b^\dagger$ satisfy the anti-commutation relation $\{b, b^\dagger\} \equiv bb^\dagger + b^\dagger b = 1$ and express the Hamiltonian which describes this atom in terms of them.

2. Consider a "cavity" with electromagnetic radiation, photons, which can only have one value of energy, the energy E.

Using the corresponding creation and annihilation operators $a^\dagger$ and a, write the Hamiltonian for these free photons.

3. Describe the energy eigenstates and the corresponding spectrum of the whole system of the "atom", together with the radiation, ignoring their interaction. Show that with the exception of the ground state, all energy eigenstates are doubly degenerate.

4. Allow the atom and the photons to interact, i.e. introduce an interaction term in the Hamiltonian, which describes the fundamental processes

$$|e; n\rangle \to |g; n+1\rangle\ ,\quad |g; n+1\rangle \to |e; n\rangle\ ,\quad \text{for any non-negative } n$$

where the second entry n is the number of photons in the state. *Assume that the interaction can be treated as a small perturbation.* Show that the simplest interaction term which describes the above processes is

$$H_I = \lambda a^\dagger b + \lambda^* a b^\dagger$$

with λ a constant which is assumed to satisfy $|\lambda| \ll E$.

5. Compute to leading order in the perturbation H_I the probabilities of the transitions:

$$|e; n\rangle \to |g; n+1\rangle$$

of induced ($n \neq 0$) or spontaneous ($n = 0$) photon emission by the excited atom.

6. Use the standard *perturbation theory of degenerate states* to compute the energies of the first two excited states, after the degeneracy is lifted by the perturbation. Show that the two new states have energies $E \pm |\lambda|$. What are the corresponding eigenstates?

7. *Rabi oscillations.* Assume that the cavity at time $t = 0$ is in the state $\psi(0) = |e; 0\rangle$. Take for simplicity $\lambda \in R$ and compute the probability that the cavity will be, at time t, in the state $|g; 1\rangle$. Show that the cavity oscillates between the states $|e; 0\rangle$ and $|g; 1\rangle$ and compute the period of oscillation.

3
Elements of Classical Field Theory

3.1 Introduction

Although the natural framework to describe the interactions among elementary particles is the quantum theory, we will start here by recalling some elements of the classical theory of fields. There is a good reason for that. As we alluded to in the first chapter, in order to obtain a quantum theory we start from the corresponding classical theory to which we apply the quantisation prescription. This applies to any physical system, no matter whether it has a finite or an infinite number of degrees of freedom. It follows that the knowledge of the classical system is essential in the formulation of the corresponding quantum system. Since the quantum theory of fields will be the language of elementary particle physics, it is essential to understand the corresponding classical field theory.

3.2 Lagrangian and Hamiltonian Mechanics

The high level of conceptualisation of classical mechanics has, since the 19th century, played an essential part in the development of physical theories. We shall give here a very brief review of the main results with no proofs, essentially in order to fix terminology and notations.

The Lagrangian. Let us consider a system with N degrees of freedom and let $q_a(t)$, $a = 1, \ldots, N$, denote the corresponding generalised coordinates. We will assume that they determine a point $\boldsymbol{q}$ in an N-dimensional differentiable manifold $\mathcal{M}$, for example the N-dimensional real space $\mathbb{R}^N$.[1] We shall call $\mathcal{M}$ *the configuration space* of the system. Since $\mathcal{M}$ is differentiable, we can consider at every point $\boldsymbol{q}$ the set of N tangent vectors $\dot{q}_a(t) = \mathrm{d}q_a(t)/\mathrm{d}t$ of curves passing through $\boldsymbol{q}$. Together with q_a they span a $2N$-dimensional space, which we shall call $\mathcal{T}(\mathcal{M})$.[2]

[1] In the simple case of N unconstrained real variables we just write the corresponding dynamical equations, for example Newton's equations. However, during the 18th and 19th centuries people realised that there were problems for which this simple formulation is not straightforward. A typical example is a problem with constraints, such as a particle subject to an external force but constrained to move on a given surface. The more abstract Lagrangian and Hamiltonian formulations of classical mechanics were developed to make possible also the description of such dynamical systems.

[2] If the configuration space is the N-dimensional real space $\mathbb{R}^N$, the space of the tangent vectors is again $\mathbb{R}^N$ and $\mathcal{T}(\mathcal{M}) = \mathbb{R}^N \times \mathbb{R}^N$. However, for a general differentiable manifold $\mathcal{M}$, this construction cannot be done globally and we must consider the tangent space $\mathcal{T}_{\boldsymbol{q}}(\mathcal{M})$ built above the point $\boldsymbol{q}$. We define $\mathcal{T}(\mathcal{M})$ as the union over all $\boldsymbol{q}$ of these tangent spaces and we call it *the tangent space* of the manifold $\mathcal{M}$. It is obvious that $\mathcal{T}(\mathcal{M})$ is a $2N$-dimensional vector space because, given any two points $\boldsymbol{q_1}$ and $\boldsymbol{q_2}$, in order to expand a vector of $\mathcal{T}_{\boldsymbol{q_1}}(\mathcal{M})$ in terms of the basic vectors of $\mathcal{T}_{\boldsymbol{q_2}}(\mathcal{M})$ we need

Elementary Particle Physics. John Iliopoulos and Theodore N. Tomaras, Oxford University Press.
 DOI: 10.1093/oso/9780192844200.003.0003

A *Lagrangian* L is a real function of the $2N$ variables q_a and $\dot{q}_a$ and, possibly, the time t, i.e. $L(q_a, \dot{q}_a, t) : \mathcal{T}(\mathcal{M}) \times \mathbb{R} \to \mathbb{R}$. An important mathematical tool, which was developed for functional analysis problems, of the kind we shall deal with in this book, is *the calculus of variations*. For the simple case of $\mathcal{M} = \mathbb{R}^N$ it derives the following well-known theorem:

• Consider $\boldsymbol{q} \in \mathbb{R}^N$ and let $\gamma = \{t, \boldsymbol{q} \mid \boldsymbol{q} = \boldsymbol{q}(t), t_0 \leq t \leq t_1\}$ be a curve in $\mathbb{R}^N \times \mathbb{R}$ such that $\boldsymbol{q}(t_0) = \boldsymbol{q}_0$ and $\boldsymbol{q}(t_1) = \boldsymbol{q}_1$, and let the Lagrangian $L : \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}$ be a sufficiently regular function of $2N + 1$ variables. We can prove that the curve γ is extremal for the *action* functional defined by $S[\gamma] = \int_{t_0}^{t_1} L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)\mathrm{d}t$ in the space of the curves joining $(t_0, \boldsymbol{q}_0)$ to $(t_1, \boldsymbol{q}_1)$ if and only if the Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\boldsymbol{q}}}\right) - \frac{\partial L}{\partial \boldsymbol{q}} = 0 \tag{3.1}$$

are satisfied along γ.

The Least Action Principle. We can now formulate the *Principle of Least Action* which links Newton's equations to the Euler–Lagrange equations:

• A dynamical system is called *natural* if $L = T - V$, with T the kinetic energy and V the potential energy of the system. It is easy to prove that, for a natural system, the extrema of the functional $S[\gamma]$ are given by the solutions of Newton's equations.

The Hamiltonian. The connection between the Lagrangian and the Hamiltonian formulations is given by a *Legendre transformation*. We define the N quantities $p_a(t)$, usually called *conjugate momenta* to $q_a(t)$, and the *Hamiltonian* $H(\boldsymbol{q}, \boldsymbol{p}, t)$ of the system by

$$p_a(t) = \frac{\partial L}{\partial \dot{q}_a} \quad \text{and} \quad H(\boldsymbol{q}, \boldsymbol{p}, t) = \boldsymbol{p}\,\dot{\boldsymbol{q}} - L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) \tag{3.2}$$

respectively. In writing equation (3.2) we assume that we can invert the definition of $\boldsymbol{p}$ and express $\dot{\boldsymbol{q}}$ in terms of $\boldsymbol{p}$ and $\boldsymbol{q}$. Then we can prove the following theorem:

• The Euler–Lagrange system (3.1) of N second order differential equations is equivalent to the Hamilton system of $2N$ first order equations

$$\dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{q}} \quad , \quad \dot{\boldsymbol{q}} = \frac{\partial H}{\partial \boldsymbol{p}} \tag{3.3}$$

As a result of these equations the time evolution of any quantity $f(\boldsymbol{p}, \boldsymbol{q}, t)$ is obtained from its equation of motion

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \boldsymbol{q}}\frac{\partial H}{\partial \boldsymbol{p}} + \frac{\partial f}{\partial \boldsymbol{p}}\left(-\frac{\partial H}{\partial \boldsymbol{q}}\right) = \frac{\partial f}{\partial t} + \{H, f\} \tag{3.4}$$

where, by definition, the *Poisson bracket* of two functions f and g is given by

$$\{f, g\} = \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q_a} - \frac{\partial f}{\partial q_a}\frac{\partial g}{\partial p_a} \tag{3.5}$$

also the basic vectors of $\mathcal{M}$. In addition, we can endow $\mathcal{T}(\mathcal{M})$ with the structure of a fibre bundle, *the tangent bundle*, but we will not use it in this book.

• It follows that if f does not depend explicitly on time, the statement "f is a constant of the motion" is equivalent to $\{H, f\} = 0$. An important corollary of this property is the fact that a Hamiltonian, which does not depend explicitly on time, is conserved, i.e.

$$\frac{\partial H}{\partial t} = 0 \quad \text{implies that} \quad \frac{\mathrm{d}H}{\mathrm{d}t} = 0 \tag{3.6}$$

• Clearly q_a and p_a span a $2N$-dimensional space which we will denote by $\mathcal{T}^*(\mathcal{M})$. It is the *cotangent space* of $\mathcal{M}$.[3] In physics texts it is usually called *the phase space* of the system. So, the Hamiltonian is a real function $H : \mathcal{T}^*(\mathcal{M}) \times \mathbb{R} \to \mathbb{R}$.

Noether's theorem. Between 1915 and 1918 A.E. Noether proved a mathematical theorem which profoundly influenced all aspects of the physical sciences. In rather loose terms it established a connection between continuous symmetries and conservation laws. We will give here a more precise formulation, still avoiding most technical details.

• To each one parameter group of diffeomorphisms of the configuration manifold, which preserves the Lagrangian, there corresponds a prime integral of the equations of motion, i.e. a conserved quantity.

Let us consider a differentiable map $h : \mathcal{M} \to \mathcal{M}$ and let $T(h) : \mathcal{T}(\mathcal{M}) \to \mathcal{T}(\mathcal{M})$ be the induced map on the tangent space. A Lagrangian system $(\mathcal{M}, L)$ is invariant under h, if its Lagrangian function remains unchanged under the action of h on its variables, i.e. if $\forall\, v \in \mathcal{T}(\mathcal{M})$

$$L(T(h)v) = L(v) \tag{3.7}$$

Noether's theorem asserts that:

Theorem: If the Lagrangian system $(\mathcal{M}, L)$ is invariant under the one parameter group of diffeomorphisms $h_s : \mathcal{M} \to \mathcal{M}, s \in \mathbb{R}$,[4] then the system of the Euler–Lagrange equations admits a prime integral $I : \mathcal{T}(\mathcal{M}) \to \mathbb{R}$, which, furthermore, in a local coordinate system is given by

$$I(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \frac{\partial L}{\partial \dot{\boldsymbol{q}}} \frac{\mathrm{d}h_s(\boldsymbol{q})}{\mathrm{d}s}\bigg|_{s=0} \tag{3.8}$$

Applications of Noether's theorem. We present some important consequences of this theorem. See also Problem 3.1. Let us consider a system of N particles[5] with Lagrangian

$$L = \sum_{a=1}^{N} \frac{1}{2} m_a \dot{\boldsymbol{q}}_a^2 - V(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N), \ \boldsymbol{q}_a \in \mathbb{R}^3, \ a = 1, \ldots, N$$

• We assume that the Lagrangian is invariant under translations

[3] We can again endow it with the structure of a fibre bundle, *the cotangent bundle*.

[4] In fact, we can prove Noether's theorem without assuming the invariance of the Lagrangian expressed by equation (3.7). A more global version, in which only the action S is invariant, suffices. The difference is important because it allows to consider transformations under which the Lagrangian changes by the time derivative of a function. We will need this more general version later in this chapter.

[5] In this case the variables $\boldsymbol{q}_a$ are just the position vectors $\boldsymbol{r}_a$ of the particles.

$$h_s : \boldsymbol{q}_a \to \boldsymbol{q}_a + s\,\boldsymbol{c}, \quad \boldsymbol{c} \in \mathbb{R}^3, \quad a = 1, \ldots, N \tag{3.9}$$

The corresponding conserved quantity is the *total momentum* of the system.

$$0 = \frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t} \cdot \boldsymbol{c} \quad \text{from which} \quad \frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t} = 0 \tag{3.10}$$

We see that the invariance of the Lagrangian under translations implies the *conservation of the total momentum*.

- The Lagrangian is invariant under rotations

$$\boldsymbol{q}_a \to R(\boldsymbol{n}, \theta)\boldsymbol{q}_a, \quad a = 1, \ldots, N \tag{3.11}$$

where $R(\boldsymbol{n}, \theta)$ is a rotation with axis of rotation and direction determined by $\boldsymbol{n}$ and angle θ. In this case, what follows is the *conservation of the total angular momentum*

$$\frac{\mathrm{d}\boldsymbol{L}}{\mathrm{d}t} = 0 \ \text{ with } \ \boldsymbol{L} = \sum_{a=1}^{N} \boldsymbol{q}_a \wedge \boldsymbol{p}_a \tag{3.12}$$

- The conservation of the Hamiltonian expressed in equation (3.6) is also a consequence of Noether's theorem. It is valid whenever the Lagrangian does not depend explicitly on time, in which case it is invariant under time translations.

3.3 Classical Field Theory

Although in the discussion of the previous section we used mainly the example of a system of point particles moving in space, nothing in the formalism depends on this. We can consider a general dynamical system with $q_a(t)$ $a = 1, \ldots, N$, the corresponding generalised coordinates. In Chapter 2 we indicated a formal way to take the large N limit and obtain a classical field theory starting from a system defined on the points of a spatial lattice. In this book we shall be interested in Lorentz invariant field theories, so we are led to formulate a classical field theory based on a set of postulates, which extend those we used in the formulation of classical mechanics.

The Minkowski space. The base space of the theory is the four-dimensional Minkowski space $\mathbb{M}^4$ with the metric $\eta_{\mu\nu} = \eta^{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$, which is left invariant by the ten parameter group of the Poincaré transformations. A point in this space will be denoted by $x = (x^0, \boldsymbol{x})$.

The fields. The dynamical variables are real or complex valued functions of x, i.e. $\phi : \mathbb{M}^4 \to \mathbb{R}$ or $\mathbb{C}$. Unless otherwise stated, they are taken to be C^∞ and to vanish at infinity. In addition, we will assume that the fields $\phi(x)$ may transform non-trivially under a group of transformations, for example Lorentz transformations, whose properties we will study in Chapter 5.

The set of fields forms an infinite-dimensional functional space $\mathcal{M}$, which is the configuration space of our dynamical system. It is a differentiable manifold, so at every point ϕ we can consider the tangent vectors $\partial_\mu \phi$. This way we can build the tangent space at each point ϕ. The union of the tangent spaces for all ϕ forms the tangent space of $\mathcal{M}$ and, by analogy with what we did in classical mechanics, we build the space $\mathcal{T}(\mathcal{M})$ as the functional space build out of ϕ and $\partial_\mu \phi$.

Obviously, we can generalise this construction by considering a family of fields $\{\phi_a(x)\}_{a\in I}$ indexed by I.

The Lagrangian density. The Lagrangian density is a real function of ϕ, $\partial_\mu\phi$ and, possibly, the space-time point x, that is $\mathcal{L} : \mathcal{T}(\mathcal{M}) \times \mathbb{M}^4 \to \mathbb{R}$. We will always choose $\mathcal{L}$ to be a Lorentz scalar. The action S is a functional of the fields given by

$$S[\phi] = \int_\Omega \mathcal{L}(\phi(x), \partial_\mu\phi(x), x)\, \mathrm{d}^4x. \tag{3.13}$$

where Ω is a regular region of space-time over which the Lagrangian density is being integrated. We will usually take Ω to be the entire Minkowski space $\mathbb{M}^4$.

The principle of least action and the Euler–Lagrange equations. Only very few changes are necessary in order to adapt the variational calculus of mechanics to the case of a field theory. We define the *functional derivative*, $\delta F[\phi]/\delta\phi(x)$, of a functional $F[\phi]$ by the limit, when it exists, of

$$\frac{\delta F[\phi]}{\delta\phi(x)} = \lim_{\varepsilon\to 0} \frac{F[\phi + \varepsilon\delta_x] - F[\phi]}{\varepsilon} \tag{3.14}$$

where δ_x is the Dirac measure at the point x, $\delta_x(y) = \delta(y - x)$. As an example, for $F = \int \phi^n(y)\, \mathrm{d}^4y$, we obtain

$$\begin{aligned}\frac{\delta F[\phi]}{\delta\phi(x)} &= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \left(\int \mathrm{d}^4y\, [(\phi(y) + \varepsilon\delta_x(y))^n - \phi^n(y)] \right) \\ &= n \int \mathrm{d}^4y\, \phi^{n-1}(y)\delta(y - x) = n\, \phi^{n-1}(x)\end{aligned} \tag{3.15}$$

On the other hand, the change in the Lagrangian, induced by an infinitesimal shift $\delta\phi_a(x)$ of the fields, $\phi_a(x) \to \phi_a(x) + \delta\phi_a(x)$, is given by

$$\begin{aligned}\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi_a}\delta\phi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\partial_\mu\phi_a \\ &= \frac{\partial\mathcal{L}}{\partial\phi_a}\delta\phi_a + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\delta\phi_a\right) - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)\delta\phi_a \\ &= \left[\frac{\partial\mathcal{L}}{\partial\phi_a} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\right)\right]\delta\phi_a + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}\,\delta\phi_a\right)\end{aligned} \tag{3.16}$$

This relation is very important and can be used in various ways. Let us first consider the induced variation of the action as defined in equation (3.13). The last term in (3.16) is a 4-divergence. It gives no contribution upon integration in Ω, because of the vanishing on the boundary $\partial\Omega$ of the fields and of their infinitesimal variations. Thus, the first result derived from (3.16) is that the requirement that the action $S[\phi]$ be stationary under an arbitrary variation $\phi_a(x) \to \phi_a(x) + \delta\phi_a(x)$ with ϕ_a and $\delta\phi_a$ vanishing on $\partial\Omega$, implies the equations of motion for the fields ϕ_a:

$$\frac{\partial\mathcal{L}}{\partial\phi_a(x)} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a(x))}\right) = 0 \tag{3.17}$$

These are the Euler–Lagrange equations of motion of the classical fields.

It should be noted that, under the conditions we imposed for the derivation of the Euler–Lagrange equations and, in particular, the vanishing of the field variables on the boundary Ω, the action is not related in a unique way to the Lagrangian density. We can in fact add to $\mathcal{L}$ a term of the form $\partial_\mu R^\mu$, with R^μ any sufficiently regular function of the fields and their derivatives, since by Gauss's theorem its integral over Ω vanishes.

The Hamiltonian density. With the Lagrangian density being a Lorentz scalar, the entire formulation we presented so far is explicitly Lorentz covariant. The Hamiltonian on the other hand is supposed to transform like the zero component of a four vector, therefore we must single out the time direction in x: $x = (t, \boldsymbol{x})$ and perform a Legendre transformation with respect to $\dot{\phi}(x) = \partial_0 \phi(x)$. Let us define

$$\pi_a(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_a(x)} \tag{3.18}$$

where we have assumed that the fields $\{\phi_a(x)\}_{a \in I}$ form a set of independent dynamical variables. We have seen already in Chapter 2 that this is not always true, for example when the index "a" denotes the components of a vector field. We will address this question in more detail in Chapter 7.

The Hamiltonian density is defined by the analogue of equation (3.2):

$$\mathcal{H} = \sum_{a \in I} \pi_a \dot{\phi}_a - \mathcal{L}\,, \tag{3.19}$$

and the Hamiltonian $H[\phi]$ by

$$H[\phi_a, \pi_a, t] = \int \mathrm{d}^3 x\, \mathcal{H}\left(\phi_a(x), \partial_\mu \phi_a(x), x\right) \tag{3.20}$$

from which Hamilton's equations of motion

$$\dot{\phi}_a(x) = \frac{\delta H}{\delta \pi_a(x)} \quad \text{and} \quad \dot{\pi}_a(x) = -\frac{\delta H}{\delta \phi_a(x)} \tag{3.21}$$

can be obtained.

Symmetry transformations. Following the formalism we developed for the case of classical mechanics with a finite number of degrees of freedom, we shall call "symmetry" a transformation that leaves invariant the equations of motion. Strictly speaking, this implies that only the action should be left invariant, but unless stated otherwise we will make here a stronger assumption, namely to assume that the transformations leave invariant the Lagrangian density. As we will see shortly, this assumption, apart from simplifying the formalism, will also have some important physical consequences. For practical and conceptual reasons, we will distinguish two types of transformations, *space-time transformations* and *internal symmetry transformations.*

• *The space-time transformations* are of the general form $x \to x' = f(x)$, so they are automorphisms of the Minkowski space $\mathbb{M}^4$. In this book we will consider only the ten-parameter group of Poincaré transformations which we will study briefly in

Chapter 5. They consist of the space-time translations and the Lorentz transformations. A translation is of the form $x'_\mu = x_\mu + \epsilon_\mu$ which, for ϵ infinitesimal, induces a transformation of the fields of the form $\phi_a(x) \to \phi'_a(x) = \phi(x-\epsilon) \simeq \phi_a(x) - \epsilon^\mu \partial_\mu \phi_a(x)$. A Lorentz transformation is of the form $x'^\mu = \Lambda^\mu{}_\nu\, x^\nu$ and we assume that the fields transform linearly. To these continuous transformations we should add the two discrete ones, the *space inversion*, or *parity* ($\mathcal{P}$): $t' = t$, $\boldsymbol{x}' = -\boldsymbol{x}$, and the *time reversal* ($\mathcal{T}$): $t' = -t$, $\boldsymbol{x}' = \boldsymbol{x}$.

• *Internal symmetry transformations.* These do not affect the space-time point x but transform the fields $\phi_a(x)$ among themselves. They play a very important role in particle physics and we will see many examples in this book. We will assume that all internal transformations form a compact Lie group[6] G under which the fields transform linearly. The action of an infinitesimal such transformation with parameters $\delta\theta_n$ on the fields $\phi_a(x)$ has the general form

$$\phi_a(x) \to \phi'_a(x) = \phi_a(x) + \delta\phi_a(x)\,, \qquad \delta\phi_a(x) = \mathrm{i}\,\delta\theta_n\,(T_n)_{ab}\,\phi_b(x) \tag{3.22}$$

The number N of parameters θ_n depends on the group G, while the form of the matrices T_n depends on the group and the transformation properties of the fields ϕ_a.

• *Global and local transformations.* Most transformations we are considering, space-time or internal symmetry ones, depend continuously on a set of parameters, such as ϵ^μ and θ_n in the above examples. We will call a transformation *global* if the corresponding parameters are constants, independent of the space-time point x. In the opposite case, if the parameters are arbitrary functions of x, we will call the transformation *local*. Local transformations are often also called *gauge transformations*. In Chapter 13 we shall study the physical consequences of local symmetries in more detail, but here let us just point out an obvious difference between the two cases. Let us consider the example of the internal transformation of equation (3.22). If the transformation is global, the derivative of the field $\phi_a(x)$ transforms the same way as the field itself. If, however, the parameters become functions $\delta\theta_n(x)$, the derivative picks up a second term proportional to $\partial_\mu \delta\theta_n(x)$. This will make the transformed theory look more complicated, but, as we will see, it will eventually result in a mathematically richer structure with amazing physical consequences.

Noether's theorem in classical field theory. We have seen in section 3.2 that to every continuous transformation that leaves the Lagrangian invariant there corresponds a conserved quantity. We will now prove that to every continuous change of the fields leaving invariant the action $S[\phi]$ there corresponds a conserved current.

We assume that the infinitesimal change of the fields, $\phi_a(x) \to \phi_a(x) + \delta\phi_a(x)$, leaves the action $S[\phi]$ invariant. The variation of the Lagrangian due to this change is still given by (3.16), but now the change $\delta\phi$ is not arbitrary. By assumption, it leaves the action invariant, which means that the change of the Lagrangian density has to be the 4-divergence of some quantity, i.e.

$$\delta\mathcal{L} = \partial_\mu R^\mu \tag{3.23}$$

[6]Some notions of group theory will be necessary in analysing the symmetry properties of a system and we will present a brief review in Chapter 5.

and this is true for arbitrary configurations of the fields ϕ_a. This is often called an "off-shell equality", because the fields do not need to satisfy the Euler–Lagrange equations.

Now suppose that the fields ϕ_a satisfy the Euler–Lagrange equations of motion. In this case, comparison of (3.23) and (3.16) leads to the relation

$$\delta\mathcal{L} = \partial_\mu R^\mu = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \delta\phi_a \right) \tag{3.24}$$

from which it follows that

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \delta\phi_a - R^\mu \tag{3.25}$$

is conserved (it is an "on-shell property"):

$$\partial_\mu J^\mu(x) = 0\,. \tag{3.26}$$

The Noether current J^μ is conserved and this conservation is a consequence of the symmetry of the theory. The associated "charge" $Q(V)$ defined by the spatial integral of $J^0(x)$ over a volume V bounded by the surface $S(V)$ satisfies the relation

$$\frac{\mathrm{d}Q(V)}{\mathrm{d}t} = \int_V \mathrm{d}^3x \frac{\partial J^0}{\partial t} = -\int_V \mathrm{d}^3x \nabla \cdot \boldsymbol{J} = -\int_{S(V)} \mathrm{d}\boldsymbol{S} \cdot \boldsymbol{J} \tag{3.27}$$

expressing the fact that the amount of charge $Q(V)$ inside any volume changes per unit time by exactly the amount that flows through its boundary. Under the assumption that nothing flows to infinity, the total charge Q in the entire space is conserved, since it satisfies

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = 0\,. \tag{3.28}$$

There is a certain freedom in the definition of the Noether current because we can add to it and/or multiply it by any constant without changing the fact that its 4-divergence vanishes.

Examples of Noether currents. We shall give examples of space-time and internal symmetries.

• *Space-time symmetries: The energy-momentum tensor.* In classical mechanics, invariance under spatial translations gives rise to the conservation of the linear momentum while invariance under time translation gives rise to the conservation of energy. We will now derive the analogue of this in field theories. Specifically, we will show directly that the fact that the Lagrangian does not explicitly depend on x makes it possible to define, using the Euler–Lagrange equations, a conserved tensor. We choose Ω to be the entire four-dimensional space-time (a domain of integration which is invariant under translations) and assume that the fields decrease fast enough at infinity to ignore all surface terms.[7]

[7]The invariance under translations is the only case we will consider in this book in which the action is invariant but not the Lagrangian density. Indeed, $\mathcal{L}(\phi(x), \partial_\mu\phi(x))$ cannot be invariant under $x_\mu \to x_\mu + \epsilon_\mu$, unless it is a constant. However, under the assumption that the fields vanish sufficiently fast at infinity, its integral over the entire space-time, i.e. the action, is invariant.

Let us write in two different ways the total derivative of $\mathcal{L}$ with respect to x^μ. Since $\mathcal{L}$ depends on x^μ only through the dependence of the fields and their derivatives, we have

$$\begin{aligned}\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}x^\mu} &= \frac{\partial\mathcal{L}}{\partial\phi_a(x)}\partial_\mu\phi_a(x) + \frac{\partial\mathcal{L}}{\partial[\partial_\nu\phi_a(x)]}\partial_\mu\partial_\nu\phi_a(x)\\ &= \partial_\nu\left(\frac{\partial\mathcal{L}}{\partial[\partial_\nu\phi_a(x)]}\right)\partial_\mu\phi_a(x) + \frac{\partial\mathcal{L}}{\partial[\partial_\nu\phi_a(x)]}\partial_\mu\partial_\nu\phi_a(x)\\ &= \partial_\nu\left(\frac{\partial\mathcal{L}}{\partial[\partial_\nu\phi_a(x)]}\partial_\mu\phi_a(x)\right)\end{aligned}$$

But, for any field configuration $\phi(x)$ the Lagrangian density is a function of x and its derivative with respect to x^μ is then just

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}x^\mu} = \partial_\mu\mathcal{L} = \delta^\nu_\mu\,\partial_\nu\mathcal{L}$$

Equating the above two expressions for $\mathrm{d}\mathcal{L}/\mathrm{d}x^\mu$ we find

$$\partial_\nu\widetilde{T}^\nu_\mu = 0 \tag{3.29}$$

where

$$\widetilde{T}^\nu_\mu(x) = \frac{\partial\mathcal{L}}{\partial[\partial_\nu\phi_a(x)]}\partial_\mu\phi_a(x) - \mathcal{L}(x)\,\delta^\nu_\mu \tag{3.30}$$

is called the *energy-momentum* tensor of the theory with Lagrangian density $\mathcal{L}$.

We deduce from it that the four-vector P^μ given by

$$P^\mu = \int \mathrm{d}^3x\,\widetilde{T}^{0\mu}(t,\boldsymbol{x}) \tag{3.31}$$

is time independent since

$$\dot{P}^\mu = \int \mathrm{d}^3x\,\partial_t\,\widetilde{T}^{0\mu}(t,\boldsymbol{x}) = -\int \mathrm{d}^3x\,\partial_i\widetilde{T}^{i\mu}(t,\boldsymbol{x}) = 0 \tag{3.32}$$

if the fields vanish fast enough at infinity, as it has been assumed.

The fact that P^μ is a four-vector can become manifest by changing the spatial integration in their definition from an integration over a surface of constant t to one over a space-like surface with d^3x being replaced by the associated covariant surface element $\mathrm{d}\sigma^\nu$ normal of the surface

$$P^\mu = \int \mathrm{d}\sigma_\nu\widetilde{T}^{\nu\mu}(\boldsymbol{x},t) \tag{3.33}$$

For reasons which will become clear later on, it is useful to replace this energy-momentum tensor by another conserved tensor $T^{\mu\nu}$ symmetric in its two indices, obtained from $\widetilde{T}^{\mu\nu}$ by adding to it the 4-divergence of a three-index tensor. Alternatively,

it can be shown that $T^{\mu\nu}$ can be obtained by varying with respect to the space-time metric $g_{\mu\nu}$ the action obtained when the "matter" theory with the given $\mathcal{L}$ is coupled to gravity.

• *Internal symmetry.* Let us consider a global internal symmetry of the Lagrangian density with the fields transforming as in equation (3.22). From the definition of the Noether current (3.25), it follows

$$J^\mu = \frac{\partial \mathcal{L}(x)}{\partial(\partial_\mu \phi_a(x))} \delta\phi_a(x) = \mathrm{i}\, \delta\theta_n \frac{\partial \mathcal{L}(x)}{\partial(\partial_\mu \phi_a(x))} (T_n)_{ab}\, \phi_b(x) \tag{3.34}$$

thus showing the existence of N conserved currents

$$J^\mu_n(x) = \mathrm{i}\, \frac{\partial \mathcal{L}(x)}{\partial(\partial^\mu \phi_a(x))} (T_n)_{ab}\, \phi_b(x)\,, \quad n = 1, \ldots, N \tag{3.35}$$

where we made use of the previous remark, that we can add or multiply a current by any constant.

3.4 Problems

Problem 3.1 Consider a system of N particles with Lagrangian:

$$L = \sum_{a=1}^{N} \frac{1}{2} m_a \dot{\boldsymbol{q}}_a^2 - V(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N),\ \boldsymbol{q}_a \in \mathbb{R}^3,\ a = 1, \ldots, N$$

Using Noether's theorem prove that:

1. Invariance of L under translations implies the conservation of the total momentum, eq. (3.10).

2. Invariance of L under rotations implies the conservation of the total angular momentum, eq. (3.12).

Note that either one of these assumptions concerns in fact the potential V because the kinetic energy part, by construction, is invariant under both translations and rotations.

Problem 3.2 Find the equations of motion for the field $\phi(x)$ resulting from the following Lagrangian densities:

1. $\mathcal{L}_1 = \frac{1}{2}\, \partial^\mu \phi(x)\, \partial_\mu \phi(x) - \frac{1}{2} m^2 \phi^2(x) - \lambda \phi^3(x)$
2. $\mathcal{L}_2 = -\frac{1}{2} \phi(x) \Box \phi(x) - \frac{1}{2} m^2 \phi^2(x) - \lambda \phi^3(x)$
3. $\mathcal{L}_3 = \frac{1}{2}\, \partial^\mu \phi(x)\, \partial_\mu \phi(x) - \frac{1}{2} m^2 \phi^2(x) - \lambda \phi^3(x)$
$+ 2g\, \phi(x)\, \partial^\mu \phi(x)\, \partial_\mu \phi(x) + g \phi^2(x) \Box \phi(x)$

Comment on the results.

Problem 3.3 Let $\phi_a(x)$, $a = 1, 2, \cdots, N$ denote N real scalar fields. Find the group of internal transformations which leave invariant the Lagrangian density:

$\mathcal{L} = \frac{1}{2}\sum_{a=1}^{N}\left[\partial^\mu\phi_a(x)\,\partial_\mu\phi_a(x) - m^2\phi_a(x)\phi_a(x)\right] - \lambda\left[\sum_{a=1}^{N}\phi_a(x)\phi_a(x)\right]^2$
where λ is a constant. Write the corresponding conserved currents.

Problem 3.4 *The conserved charges generate the corresponding symmetry transformations.* Consider a Lagrangian density $\mathcal{L}$ invariant under the global transformations (3.22) leading to the conserved currents (3.35). Show that the Poisson brackets of the corresponding conserved charges Q_n with the fields $\phi_a(x)$ generate the infinitesimal transformations (3.22).

4

Scattering in Classical and Quantum Physics

4.1 Introduction

Almost every act of observation involves a scattering experiment. In everyday life it is usually the scattering of visible light by the object under study and its resolution is limited by the light's wavelength. During the last century the quest for higher resolution forced us to abandon light as a probe and, following Rutherford's pioneering experiment, to use more and more energetic particle beams. Today all information we have about the structure of matter at the deepest accessed level comes from high energy scattering experiments. The probes are particles we can accelerate, which are essentially protons (or ions) and electrons. A particle accelerator is, in fact, a "microscope" whose spatial resolution is determined by its maximum energy. The CERN Large Hadron Collider (LHC), with proton beams up to 7 TeV, has a resolution reaching 10^{-19} m,[1] and is today – in 2021 – the most powerful microscope man has ever built. In this chapter we shall introduce the basic concepts necessary to describe and understand the results of scattering experiments in particle physics. Only the main ideas will be presented, with no detailed proofs. Some of the proofs are proposed as exercises at the end of the chapter.

4.2 The Scattering Cross Section

We start by considering the case of two colliding particles. The *laboratory frame* is defined to be the reference frame in which one of these particles, called *the target*, is at rest. The other is the projectile. By contrast, in the *centre of mass reference frame* nothing distinguishes the target from the beam.[2]

In the simplest case a projectile, idealised as a hard sphere of radius r in straight motion, will hit the target, a sphere of radius R, if the trajectory of the centre of the projectile intersects the disk of radius $r + R$ perpendicular to it and centred at the

[1] This is only a rough order-of-magnitude estimation of an "ideal resolution" δ obtained by setting $\delta \sim \lambda \sim h/E$. The actual resolution in a particular measurement will depend on the experimental conditions.

[2] The terminology has a purely historical origin. In the early experiments a beam of accelerated particles was hitting a target, which was fixed in the laboratory, hence the name of the reference frame. Today, however, most accelerators are colliders, in which two beams of particles are accelerated in opposite directions and are brought into a head-on collision. In these cases the "laboratory" frame is in fact the centre-of-mass one.

Elementary Particle Physics. John Iliopoulos and Theodore N. Tomaras, Oxford University Press.
 DOI: 10.1093/oso/9780192844200.003.0004

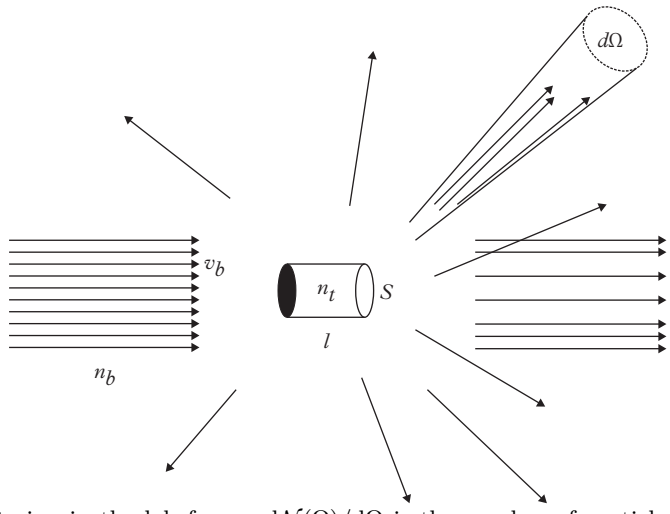

Fig. 4.1 Scattering in the lab frame. $\mathrm{d}\mathcal{N}(\Omega)/\mathrm{d}\Omega$ is the number of particles per unit time scattered inside $\mathrm{d}\Omega$ in the direction $\Omega(\theta, \phi)$.

centre of the target. The surface area of this disk, $\sigma_{\mathrm{tot}} = \pi(r + R)^2$, is called *the total cross section* of this collision process.

In actual experiments the situation is more complex. First, we are not interested only in collisions of hard spheres. Even in classical physics we may want to compute the results of scattering among particles interacting through a potential, for example two electrically charged particles. Second, we do not usually consider the collision of just one particle against another. We send, instead, a beam of particles against a target containing many particles and place a detector in the direction $\Omega(\theta, \phi)$ with an acceptance of $\mathrm{d}\Omega$, which counts the number of particles $\mathrm{d}\mathcal{N}(\Omega)/\mathrm{d}\Omega$ going through it per unit solid angle and per unit time. We want to extract out of such a measurement a quantity, like the cross section we introduced previously, which refers to the collision of one particle in the beam and one particle in the target.

Let us consider a unidirectional monoenergetic beam of particles scattered off a target containing n_t centres of collision per unit volume. Suppose, for concreteness, that this target is a cylinder of length l and base area S perpendicular to the direction of the beam (see Figure 4.1). If the target is thin enough, the density n_t not too high and the number of incoming particles per unit time large enough but not too large,[3] then this experiment yields information about all processes involving one beam particle and one particle in the target. The reason is that the number $\mathrm{d}\mathcal{N}(\Omega)/\mathrm{d}\Omega$ measured in our detector is proportional to the number $n_t S\, l$ of scattering centres in the target and to the relative flux Φ_b of the incoming particles in the beam, i.e. the number of particles per unit transverse surface area, which reach the target per unit time. Thus, we write

[3] All these restrictions amount to assuming that the probability of multiple scattering, i.e. the same incident particle hitting several particles in the target, is small and can be neglected.

$$\frac{\mathrm{d}\mathcal{N}(\Omega)}{\mathrm{d}\Omega} = n_t S l \Phi_b \frac{\mathrm{d}\sigma(\Omega)}{\mathrm{d}\Omega} \tag{4.1}$$

We shall call the proportionality factor $\mathrm{d}\sigma(\Omega)/\mathrm{d}\Omega$ *differential cross section* of the scattering process involving one beam particle and one target particle. If the density of particles in the beam is n_b and they all move with speed v_b with respect to the target, the incident flux is given by $\Phi_b = n_b v_b$. It is clear that by moving the detector around we can determine the function $\mathcal{N}(\Omega)$ and, assuming it is integrable, we can integrate over all angles and obtain *the integrated cross section* σ. Note that the differential cross section in the forward direction $\theta = 0$, is not directly measurable because a detector placed in the direction of the beam will count all particles in this direction, irrespective of whether or not they had a collision with a particle in the target. It follows that the forward cross section can be determined only by continuity.

From (4.1) we conclude that the total number $\mathrm{d}N$ of collisions that take place inside the volume element $\mathrm{d}V$ of the target in the time interval $\mathrm{d}t$ is given by

$$\mathrm{d}N = \sigma\, v_b\, n_b\, n_t\, \mathrm{d}V \mathrm{d}t \tag{4.2}$$

The differential cross section $\mathrm{d}\sigma/\mathrm{d}\Omega$ and the integrated cross section σ are defined in the lab frame, i.e. with the target at rest. The same number $\mathrm{d}N$ will be measured by any other observer, with respect to whom the colliding beams will have e.g. velocities $\boldsymbol{v}_A$ and $\boldsymbol{v}_B$, and particle densities n_A and n_B, respectively. In Problem 4.3 we invite the reader to prove that, in terms of quantities measured in that frame, $\mathrm{d}N$ is given by

$$\mathrm{d}N = \sigma \sqrt{(\boldsymbol{v}_A - \boldsymbol{v}_B)^2 - (\boldsymbol{v}_A \wedge \boldsymbol{v}_B)^2}\; n_A\, n_B\, \mathrm{d}V \mathrm{d}t \tag{4.3}$$

For collinear $\boldsymbol{v}_A$ and $\boldsymbol{v}_B$, e.g. in the centre of mass frame, $\mathrm{d}N = \sigma |\boldsymbol{v}_A - \boldsymbol{v}_B|\; n_A\, n_B\, \mathrm{d}V \mathrm{d}t$.

So far we have considered the collision of two classical particles but we can generalise the picture. Let us call A the particle in the beam and B the one in the target. The initial state, i.e. the state of the system before the collision, is characterised by the momenta p_A and p_B of the two particles (and possibly other quantities like spin orientations, which we will not discuss in this chapter). Assuming relativistic kinematics, they satisfy the on-shell condition, namely $p_A^2 = m_A^2$ and $p_B^2 = m_B^2$. We shall call *elastic scattering* the process described by $A(p_A) + B(p_B) \to A(p'_A) + B(p'_B)$, where the final momenta p'_A and p'_B are also on-shell and subject to the condition imposed by energy and momentum conservation: $p_A + p_B = p'_A + p'_B$. We know that in quantum physics elastic scattering is only part of the total process because, as we have seen already in Chapter 2, particles may be created or destroyed as a result of the interaction. Therefore the general process will be of the form $A(p_A) + B(p_B) \to C(q_1) + D(q_2) + \ldots$, where the particles C, D, etc. may or may not be one of the initial particles A or B. To give an example, at the LHC the average number of final particles in a proton–proton collision is on the order of 200. We often call the process, yielding a particular set of particles in the final state, *a channel* and we can define *a partial cross section* σ_i or *partial differential cross section* $\mathrm{d}\sigma_i$ corresponding to the channel i. The elastic cross section σ_{el} corresponds to the elastic scattering channel, which is one of the channels i. All the others correspond to *inelastic* processes. The sum over all channels gives the *total cross section* σ_{tot}.

Let us consider a particular channel consisting of r final particles with momenta $q_1, q_2, \ldots, q_r$. They all satisfy the mass shell conditions $q_a^2 = m_a^2$, $a = 1, 2, \ldots, r$, with m_a the mass of the a-th particle. The partial cross section σ_i will be given by an expression of the form:

$$\sigma_i = \int \mathrm{d}\Omega_{m_1} \cdots \int \mathrm{d}\Omega_{m_r} (2\pi)^4 \delta^4(p_A + p_B - q_1 - \cdots - q_r) \frac{\mathrm{d}\sigma_i}{\mathrm{d}\Omega_{m_1} \ldots} \tag{4.4}$$

where $\mathrm{d}\Omega_{m_a}$ is the integration measure on the mass hyperboloid of the a-th particle given by equation (2.19) and the four-dimensional delta function expresses the conservation of energy and momentum. The $\mathrm{d}\sigma_i$ is the differential cross section which is a positive definite probability density depending on the initial and final momenta. It is defined as the number of the corresponding events per unit time and per unit target volume divided by the incident flux and by the initial densities.[4] It is useful to extract all these kinematic factors: If the incoming beams contain one particle per unit volume, the division factor reduces to the incoming flux whose value is

$$v = |\boldsymbol{v}_A - \boldsymbol{v}_B| \tag{4.5}$$

where $\boldsymbol{v}_A$ and $\boldsymbol{v}_B$ are the velocities of the two incoming particles, assumed to be collinear, since this is the most interesting case in actual particle physics experiments. In general, the density of initial states depends on the way the beams are prepared, but here, although the discussion is classical, we shall adopt the normalisation given by $\varrho_{A,B} = 2\omega_{A,B} = 2\sqrt{\boldsymbol{p}_{A,B}^2 + m_{A,B}^2}$, as we found in equation (2.27). So, the overall normalisation factor is $(v 2\omega_A 2\omega_B)^{-1}$. Let us compute the expression $\varrho_A \varrho_B v$ for collinear initial momenta $\boldsymbol{p}_A$ and $\boldsymbol{p}_B$ in the centre-of-mass frame:

$$(\varrho_A \varrho_B v)^2 = 16\omega_A^2 \omega_B^2 \left| \frac{\boldsymbol{p}_A}{\omega_A} - \frac{\boldsymbol{p}_B}{\omega_B} \right|^2 = 16|\boldsymbol{p}_A|^2(\omega_A + \omega_B)^2 = 16\Big((p_A \cdot p_B)^2 - m_A^2 m_B^2\Big) \tag{4.6}$$

Therefore, the cross section can be written as:

$$\sigma_i = \frac{1}{4\sqrt{(p_A \cdot p_B)^2 - m_A^2 m_B^2}} \int \mathrm{d}\Omega_{m_1} \cdots \int \mathrm{d}\Omega_{m_r} (2\pi)^4 \delta^4(p_A + p_B - \Sigma q) |\mathcal{M}|^2 \tag{4.7}$$

where $|\mathcal{M}|^2$ is a positive definite quantity that encodes the dynamics of the particular process. Computing this function for any given scattering experiment is one of the main goals of theoretical particle physics and in this book we shall develop the necessary tools to do it. The $3r$-dimensional integral is called *the phase space* of the process and gives the cross section under the simplified assumption that $|\mathcal{M}|$ is a constant. We see that the cross section increases with the total volume of the phase space, which is

[4] The expression (4.4) is only formal. It assumes that we are able to identify all particles in the final state and define precisely the channel i. But in actual experiments there is no *universal detector* capable of identifying all particles in all kinematical regions. Especially at very high energies where the number of final particles is often very large, what we measure is a so-called *inclusive cross section* i.e. one summed over a large number of channels. We will ignore this problem for the moment and we shall give a more precise definition of inclusive processes in Chapter 21.