This book discusses the main concepts of the Standard Model of elementary particles in a compact and straightforward way. The work illustrates the unity of modern theoretical physics by combining approaches and concepts of the quantum field theory and modern condensed matter theory. The inductive approach allows a deep understanding of ideas and methods used for solving problems in this field.

THE SERIES: DE GRUYTER STUDIES IN MATHEMATICAL PHYSICS

The series is devoted to the publication of monographs and high-level texts in mathematical physics. They cover topics and methods in fields of current interest, with an emphasis on didactical presentation. The series will enable readers to understand, apply, and develop further, with sufficient rigor, mathematical methods to given problems in physics. For this reason, works with a few authors are preferred over edited volumes. The works in this series are aimed at advanced students and researchers in mathematical and theoretical physics. They can also serve as secondary reading for lectures and seminars at advanced levels.

www.degruyter.com
ISBN 978-3-11-027029-7

GSTMP
17

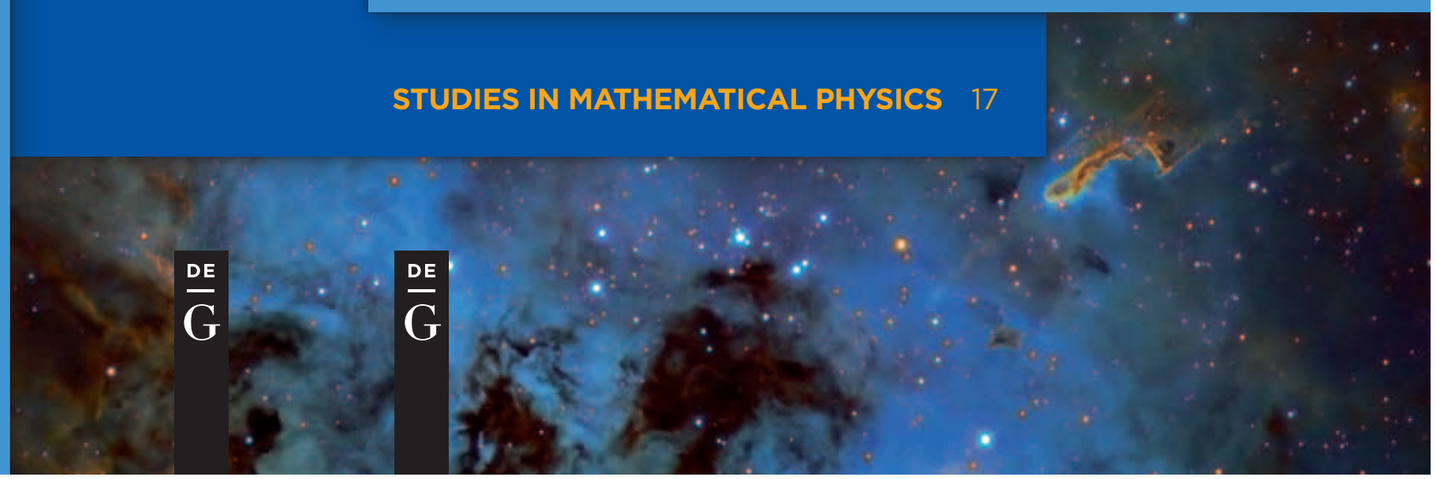
Michael V. Sadovskii
QUANTUM FIELD THEORY

DE GRUYTER

Michael V. Sadovskii

QUANTUM FIELD THEORY

STUDIES IN MATHEMATICAL PHYSICS 17



De Gruyter Studies in Mathematical Physics 17

Michael V. Sadovskii

Quantum Field Theory

De Gruyter

Physics and Astronomy Classification Scheme 2010: 03.70.+k, 03.65.Pm, 11.10.-z, 11.10.Gh, 11.10.Jj, 11.25.Db, 11.15.Bt, 11.15.Ha, 11.15.Ex, 11.30. -j, 12.20.-m, 12.38.Bx, 12.10.-g, 12.38.Cy.

ISBN 978-3-11-027029-7
e-ISBN 978-3-11-027035-8

Library of Congress Cataloging-in-Publication Data

A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.dnb.de.

Typesetting: PTP-Berlin Protago-TEX-Production GmbH, www.ptp-berlin.de
Printing and binding: Hubert & Co. GmbH & Co. KG, Göttingen
♾ Printed on acid-free paper

Printed in Germany

www.degruyter.com

Preface

This book is the revised English translation of the 2003 Russian edition of "Lectures on Quantum Field Theory", which was based on much extended lecture course taught by the author since 1991 at the Ural State University, Ekaterinburg. It is addressed mainly to graduate and PhD students, as well as to young researchers, who are working mainly in condensed matter physics and seeking a compact and relatively simple introduction to the major section of modern theoretical physics, devoted to particles and fields, which remains relatively unknown to the condensed matter community, largely unaware of the major progress related to the formulation the so-called "standard model" of elementary particles, which is at the moment the most fundamental theory of matter confirmed by experiments. In fact, this book discusses the main concepts of this fundamental theory which are basic and necessary (in the author's opinion) for everyone starting professional research work in other areas of theoretical physics, not related to high-energy physics and the theory of elementary particles, such as condensed matter theory. This is actually even more important, as many of the theoretical approaches developed in quantum field theory are now actively used in condensed matter theory, and many of the concepts of condensed matter theory are now widely used in the construction of the "standard model" of elementary particles. One of the main aims of the book is to illustrate this unity of modern theoretical physics, widely using the analogies between quantum field theory and modern condensed matter theory.

In contrast to many books on quantum field theory [2, 6, 8–10, 13, 25, 28, 53, 56, 59, 60], most of which usually follow rather deductive presentation of the material, here we use a kind of inductive approach (similar to that used in [59, 60]), when one and the same problem is discussed several times using different approaches. In the author's opinion such repetitions are useful for a more deep understanding of the various ideas and methods used for solving real problems. Of course, among the books mentioned above, the author was much influenced by [6, 56, 60], and this influence is obvious in many parts of the text. However, the choice of material and the form of presentation is essentially his own. For the present English edition some of the material was rewritten, bringing the content more up to date and adding more discussion on some of the more difficult cases.

The central idea of this book is the presentation of the basics of the gauge field theory of interacting elementary particles. As to the methods, we present a rather detailed derivation of the Feynman diagram technique, which long ago also became so important for condensed matter theory. We also discuss in detail the method of functional (path) integrals in quantum theory, which is now also widely used in many sections of theoretical physics.

We limit ourselves to this relatively traditional material, dropping some of the more modern (but more speculative) approaches, such as supersymmetry. Obviously, we also drop the discussion of some new ideas which are in fact outside the domain of the quantum field theory, such as strings and superstrings. Also we do not discuss in any detail the experimental aspects of modern high-energy physics (particle physics), using only a few illustrative examples.

Ekaterinburg, 2012 M.V. Sadovskii

Contents

We have no better way of describing elementary particles than quantum field theory. A quantum field in general is an assembly of an infinite number of interacting harmonic oscillators. Excitations of such oscillators are associated with particles … All this has the flavor of the 19th century, when people tried to construct mechanical models for all phenomena. I see nothing wrong with it, because any nontrivial idea is in a certain sense correct. The garbage of the past often becomes the treasure of the present (and vice versa*). For this reason we shall boldly investigate all possible analogies together with our main problem.*

A. M. Polyakov, "Gauge Fields and Strings", 1987 [51]

Chapter 1

Basics of elementary particles

1.1 Fundamental particles

Before we begin with the systematic presentation of the principles of quantum field theory, it is useful to give a short review of the modern knowledge of the world of elementary particles, as quantum field theory is the major instrument for describing the properties and interactions of these particles. In fact, historically, quantum field theory was developed as the principal theoretical approach in the physics of elementary particles. Below we shall introduce the basic terminology of particle physics, shortly describe the classification of elementary particles, and note some of the central ideas used to describe particle interactions. Also we shall briefly discuss some of the problems which will not be discussed at all in the rest of this book. All of these problems are discussed in more detail (on an elementary level) in a very well-written book [46] and a review [47]. It is quite useful to read these references *before* reading this book! Elementary presentation of the *theoretical principles* to be discussed below is given in [26]. A discussion of the world of elementary particles similar in spirit can be found in [23]. At the less elementary level, the basic results of the modern *experimental* physics of elementary particles, as well as basic theoretical ideas used to describe their classification and interactions, are presented in [24, 29, 50].

During many years (mainly in the 1950s and 1960s and much later in popular literature) it was a common theme to speak about a "crisis" in the physics of elementary particles which was related to an enormous number (hundreds!) of experimentally observed subnuclear ("elementary") particles, as well as to the difficulties of the theoretical description of their interactions. A great achievement of modern physics is the rather drastic simplification of this complicated picture, which is expressed by the so-called "standard model" of elementary particles. Now it is a well-established *experimental* fact, that the world of truly elementary particles[1] is rather simple and *theoretically* well described by the basic principles of modern quantum field theory.

According to most fundamental principles of relativistic quantum theory, all elementary particles are divided in two major classes, *fermions* and *bosons*. Experimentally, there are only 12 elementary fermions (with spin $s = 1/2$) and 4 bosons (with spin $s = 1$), plus corresponding antiparticles (for fermions). In this sense, our world is really rather simple!

[1] Naturally, we understand as "truly elementary" those particles which can not be shown to consist of some more elementary entities at the present level of experimental knowledge.

1.1.1 Fermions

All the known fundamental fermions ($s = 1/2$) are listed in Table 1.1. Of their properties in this table we show only the electric charge. These 12 fermions form three "generations"[2], with two *leptons* and two *quarks*[3]. To each charged fermion there is corresponding antiparticle, with an opposite value of electric charge. Whether or not there are corresponding antiparticles for neutrinos is at present undecided. It is possible that neutrinos are the so-called truly neutral particles.

Table 1.1. Fundamental fermions.

Generations	*1*	*2*	*3*	*Q*
Quarks	u	c	t	$+2/3$
("up" and "down")	d	s	b	$-1/3$
Leptons	ν_e	ν_μ	ν_τ	0
(neutrino and charged)	e	μ	τ	-1

All the remaining subnuclear particles are composite and are built of quarks. How this is done is described in detail, e. g., in [24,50][4], and we shall not deal with this problem in the following. We only remind the reader that *baryons*, i. e., fermions like protons, neutrons, and various hyperons, are built of three quarks each, while quark–antiquark pairs form *mesons*, i. e., Bosons like π-mesons, K-mesons, etc. Baryons and mesons form a large class of particles, known as *hadrons* – these particles take part in all types of interactions known in nature: strong, electromagnetic, and weak. Leptons participate only in electromagnetic and weak interactions. Similar particles originating from different generations differ only by their masses, all other quantum numbers are just the same. For example, the muon μ is in all respects equivalent to an electron, but its mass is approximately 200 times larger, and the nature of this difference is unknown. In Table 1.2 we show experimental values for masses of all fundamental fermions (in units of energy), as well as their lifetimes (or appropriate widths of resonances) for unstable particles. We also give the year of discovery of the appropriate particle[5]. The values of quark masses (as well as their lifetimes) are to be understood with some

[2] In particle theory there exists a rather well-established terminology; in the following, we use the standard terms without quotation marks. Here we wish to stress that almost all of these accepted terms have absolutely no relation to any common meaning of the words used.

[3] Leptons, such as electron and electron neutrino, have been well known for a long time. Until recently, in popular and general physics texts quarks were called "hypothetical" particles. This is wrong – quarks have been studied experimentally for a rather long time, while certain doubts have been expressed concerning their existence are related to their "theoretical" origin and impossibility of observing them in free states (confinement). It should be stressed that quarks are absolutely real particles which have been clearly observed inside hadrons in many experiments at high energies.

[4] Historical aspects of the origin of the quark model can be easily followed in older reviews [76, 77].

[5] The year of discovery is in some cases not very well defined, so that we give the year of theoretical prediction

Table 1.2. Masses and lifetimes of fundamental fermions.

$\nu_e < 10\,\text{eV}$ (1956)	$\nu_\mu < 170\,\text{KeV}$ (1962)	$\nu_\tau < 24\,\text{MeV}$ (1975)
$e = 0.5\,\text{MeV}$ (1897)	$\mu = 105.7\,\text{MeV}, 2 \cdot 10^{-6}\,\text{s}$ (1937)	$\tau = 1777\,\text{MeV}, 3 \cdot 10^{-13}\,\text{s}$ (1975)
$u = 2.5\,\text{MeV}$ (1964)	$c = 1266\,\text{MeV}, 10^{-12}\,\text{s}$ (1974)	$t = 173\,\text{GeV}, \Gamma = 2\,\text{GeV}$ (1994)
$d = 5\,\text{MeV}$ (1964)	$s = 105\,\text{MeV}$ (1964)	$b = 4.2\,\text{GeV}, 10^{-12}\,\text{s}$ (1977)

caution, as quarks are not observed as free particles, so that these values characterize quarks deep inside hadrons at some energy scale of the order of several Gev[6].

It is rather curious that in order to build the entire world around us, which consists of atoms, molecules, etc., i. e., nuclei (consisting of protons and neutrons) and electrons (with the addition of stable neutrinos), we need only fundamental fermions of the first generation! Who "ordered" two more generations, and for what purpose? At the same time, there are rather strong arguments supporting the claim, that there are only three (not more!) generations of fundamental fermions[7].

1.1.2 Vector bosons

Besides fundamental fermions, which are the basic building blocks of ordinary matter, experiments confirm the existence of four types of vector ($s = 1$) bosons, which are responsible for the transfer of basic interactions; these are the well-known *photon* γ, *gluons* g, neutral weak ("intermediate") boson Z^0, and charged weak bosons $W^\pm$ (which are antiparticles with respect to each other). The basic properties of these particles are given in Table 1.3.

Table 1.3. Fundamental bosons (masses and widths).

Boson	γ *(1900)*	g *(1973)*	Z *(1983)*	W *(1983)*
Mass	0	0	91.2 GeV	80.4 GeV
Width	0	0	2.5 GeV	2.1 GeV

The most studied of these bosons are obviously photons. These are represented by radio waves, light, X-rays, and γ-rays. The photon mass is zero, so that its energy

[6] Precise values of these and other parameters of the Standard Model, determined during the hard experimental work of recent decades, can be found in [67]

[7] In recent years it has become clear that the "ordinary" matter, consisting of atoms and molecules (built of hadrons (quarks) and leptons), corresponds to a rather small fraction of the whole universe we live in. Astrophysical and cosmological data convincingly show that most of the universe apparently consists of some unknown classes of matter, usually referred to as "dark" matter and "dark" energy, both having nothing to do with the "ordinary" particles discussed here [67]. In this book we shall discuss only "ordinary" matter.

spectrum (dispersion) is given by[8] $E = \hbar c|\mathbf{k}|$. Photons with $E \neq \hbar c|\mathbf{k}|$ are called virtual; for example the Coulomb field in the hydrogen atom creates virtual photons with $\hbar^2 c^2 \mathbf{k}^2 \gg E^2$. The source of photons is the electric charge. The corresponding dimensionless coupling constant is the well-known fine structure constant $\alpha = e^2/\hbar c \approx 1/137$. All electromagnetic interactions are transferred by the exchange of photons. The theory which describes electromagnetic interactions is called *quantum electrodynamics* (QED).

Massive vector bosons Z and $W^{\pm}$ transfer the short-range weak weak interactions. Together with photons they are responsible for the unified *electroweak* interaction. The corresponding dimensionless coupling constants are $\alpha_W = g_W^2/\hbar c \sim \alpha_Z = g_Z^2/\hbar c \sim \alpha$, of the order of the electromagnetic coupling constant.

Gluons transfer strong interactions. The sources of gluons are specific "color" charges. Each of the six types (or "flavors) of quarks u, d, c, s, t, b exists in three color states: red r, green g, blue b. Antiquarks are characterized by corresponding the anticolors: $\bar{r}, \bar{g}, \bar{b}$. The colors of quarks do not depend on their flavors. Hadrons are formed by symmetric or opposite color combinations of quarks – they are "white", and their color is zero. Taking into account antiparticles, there are 12 quarks, or 36 if we consider different colors. However, for each flavor, we are dealing simply with a different color state of each quark. Color symmetry is exact.

Color states of gluons are more complicated. Gluons are characterized not by one, but by two color indices. In total, there are eight colored gluons: $3 \times \bar{3} = 8 + 1$, one combination – $r\bar{r} + g\bar{g} + b\bar{b}$ – is white with no color charge (color neutral). Unlike in electrodynamics, where photons are electrically neutral, gluons possess color charges and interact both with quarks and among themselves, i. e., radiate and absorb other gluons ("luminous light"). This is one of the reasons for *confinement*: as we try to separate quarks, their interaction energy grows (in fact, linearly with interquark distance) to infinity, leading to nonexistence of free quarks. The theory of interacting quarks and gluons is called *quantum chromodynamics* (QCD).

1.2 Fundamental interactions

The physics of elementary particles deals with three types of interactions: strong, electromagnetic, and weak. The theory of strong interactions is based on quantum chromodynamics and describes the interactions of quarks inside hadrons. Electromagnetic and weak interactions are unified within the so-called electroweak theory. All these interactions are characterized by corresponding dimensionless coupling constants: $\alpha = e^2/\hbar c$, $\alpha_s = g^2/\hbar c$, $\alpha_W = g_W^2/\hbar c$, $\alpha_Z = g_Z^2/\hbar c$. Actually, it was

[8] Up to now we are writing $\hbar$ and c explicitly, but in the following we shall mainly use the natural system of units, extensively used in theoretical works of quantum field theory, where $\hbar = c = 1$. The main recipes to use such system of units are described in detail in Ref. [46]. In most cases $\hbar$ and c are easily restored in all expressions, when necessary.

already was recognized in the 1950s that $\alpha = e^2/\hbar c \approx 1/137$ is constant only at zero (or a very small) square of the momentum q^2, transferred during the interaction (scattering process). In fact, due to the effect of *vacuum polarization*, the value of α increases with the growth of q^2, and for large, though finite, values of q^2 can even become infinite (Landau–Pomeranchuk pole). At that time this result was considered to be a demonstration of the internal inconsistency of QED. Much later, after the creation of QCD, it was discovered that $\alpha_s(q^2)$, opposite to the case of $\alpha(q^2)$, tends to zero as $q^2 \to \infty$, which is the essence of the so-called *asymptotic freedom*. Asymptotic freedom leads to the possibility of describing gluon–quark interactions at small distances (large q^2!) by simple perturbation theory, similar to electromagnetic interactions. Asymptotic freedom is reversed at large interquark distances, where the quark–gluon interaction grows, so that perturbation theory cannot be applied: this is the essence of confinement. The difficulty in giving a theoretical description of the confinement of quarks is directly related to this inapplicability of perturbation theory at large distances (of the order of hadron size and larger). Coupling constants of weak interaction α_W, α_Z also change with transferred momentum – they grow approximately by 1% as q^2 increases from zero to $q^2 \sim 100\,\mathrm{GeV}^2$ (this is an experimental observation!). Thus, modern theory deals with the so-called "running" coupling constants. In this sense, the old problem of the size of an electric charge as a fundamental constant of nature, in fact, lost its meaning – the charge is not a constant, but a function of the characteristic distance at which particle interaction is analyzed. The theoretical extrapolation of all coupling constants to large q^2 demonstrates the tendency for them to become approximately equal for $q^2 \sim 10^{15} - 10^{16}\,\mathrm{GeV}^2$, where $\alpha \sim \alpha_s \sim \alpha_W \sim \frac{8}{3}\frac{1}{137} \approx \frac{1}{40}$. This leads to the hopes for a unified description of electroweak and strong interactions at large q^2, the so-called *grand unification theory* (GUT).

1.3 The Standard Model and perspectives

The *Standard Model* of elementary particles foundation is special relativity (equivalence of inertial frames of reference). All processes are taking place in four-dimensional Minkowski space-time $(x, y, z, t) = (\mathbf{r}, t)$. The distance between two points (events) A and B in this space is determined by a four-dimensional interval: $s^2_{AB} = c^2(t_A - t_B)^2 - (x_A - x_B)^2 - (y_A - y_B)^2 - (z_A - z_B)^2$. Interval $s^2_{AB} \geq 0$ for two events, which can be casually connected (time-like interval), while the space-like interval $s^2_{AB} < 0$ separates two events which cannot be casually related.

At the heart of the theory lies the concept of a *local* quantum field – field commutators in points separated by a space-like interval are always equal to zero: $[\psi(x_A), \psi(x_B)] = 0$ for $s^2_{AB} < 0$, which corresponds to the independence of the corresponding fields. Particles (antiparticles) are considered as quanta (excitations) of the corresponding fields. Most general principles of relativistic invariance and stability

of the ground state of the field system directly lead to the fundamental spin-statistics theorem: particles with halfinteger spins are fermions, while particles with integer spin are bosons. In principle, bosons can be assumed to be "built" of an even number of fermions; in this sense Fermions are "more fundamental".

Symmetries are of fundamental importance in quantum field theory. Besides the relativistic invariance mentioned above, modern theory considers a number of exact and approximate symmetries (symmetry groups) which are derived from the vast experimental material on the classification of particles and their interactions. Symmetries are directly related with the appropriate *conservation laws* (Noether theorem), such as energy-momentum conservation, angular momentum conservation, and conservation of different "charges". The principle of *local gauge invariance* is the key to the theory of particles interactions. Last but not least, the phenomenon of *spontaneous symmetry-breaking* (vacuum phase transitions) leads to the mechanism of mass generation for initially massless particles (Higgs mechanism)[9]. The rest of this book is essentially devoted to the explanation and deciphering of these and of some other statements to follow.

The Standard Model is based on *experimentally* established local gauge $SU(3)_c \otimes SU(2)_W \otimes U(1)_Y$ symmetry. Here $SU(3)_c$ is the symmetry of strong (color) interaction of quarks and gluons, while $SU(2)_W \otimes U(1)_Y$ describes electroweak interactions. If this last symmetry is not broken, all fermions and vector gauge bosons are massless. As a result of spontaneous $SU(2)_W \otimes U(1)_Y$ breaking, bosons responsible for weak interaction become massive, while the photon remains massless. Leptons also acquire mass (except for the neutrino?)[10]. The electrically neutral Higgs field acquires a nonzero vacuum value (Bose-condensate). The quanta of this field (the notorious Higgs bosons) are the scalar particles with spin $s = 0$, and up to now have not been discovered in experiments. The search for Higgs bosons is among the main tasks of the large hadron collider (LHC) at CERN. This task is complicated by rather indeterminate theoretical estimates [67] of Higgs boson mass, which reduce to some inequalities such as, e. g., $m_Z < m_h < 2m_Z$[11]. There is an interesting theoretical possibility that the Higgs boson could be a composite particle built of the fermions of the Standard Model (the so-called technicolor models). However, these ideas meet with serious difficulties of the selfconsistency of experimentally determined parameters of the Standard Model. In any case, the problem of experimental confirmation of the existence of the

[9] The Higgs mechanism in quantum field theory is the direct analogue of the Meissner effect in the Ginzburg–Landau theory of superconductivity.

[10] The problem of neutrino mass is somehow outside the Standard Model. There is direct evidence of finite, but very small masses of different neutrinos, following from the experiments on *neutrino oscillations* [67]. The absolute values of neutrino masses are unknown, are definitely very small (in comparison to electron mass): experiments on neutrino oscillations only measure differences of neutrino masses. The current (conservative) limitation is $m_{\nu_e} < 2\,\text{eV}$ [67]

[11] On July 4, 2012, the ATLAS and CMS collaborations at LHC announced the discovery of a new particle "consistent with the long-sought Higgs boson" with mass $m_h \sim 125.3 \pm 0.6\,\text{Gev}$. See details in *Physics Today*, September 2012, pp. 12–15. See also a brief review of experimental situation in [55].

Higgs boson remains the main problem of modern experimental particle physics. Its discovery will complete the experimental confirmation of the Standard Model. The nondiscovery of the Higgs boson within the known theoretical limits will necessarily lead to a serious revision of the Standard Model. The present-day situation of the experimental confirmation of the Standard Model is discussed in [67].

We already noted that the Standard Model (even taking into account only the first generation of fundamental fermions) is sufficient for complete understanding of the structure of matter in our world, consisting only of atoms and nuclei. All generalizations of the Standard Model up to now are rather speculative and are not supported by the experiments. There are a number of grand unification (GUT) models where multiplets of quarks and leptons are described within the single (gauge) symmetry group. This symmetry is assumed to be exact at very high transferred momenta (small distances) of the order of $q^2 \sim 10^{15} - 10^{16}\,\mathrm{GeV}^2$, where all coupling constants become (approximately) equal. Experimental confirmation of GUT is very difficult, as the energies needed to make scattering experiments with such momentum transfers are unlikely to be ever achievable by humans. The only verifiable, in principle, prediction of GUT models is the decay of the proton. However, the intensive search for proton instability during the last decades has produced no results, so that the simplest versions of GUT are definitely wrong. More elaborate GUT models predict proton lifetime one or two orders of magnitude larger, making this search much more problematic.

Another popular generalization is *supersymmetry* (SUSY), which unifies fermions and bosons into the same multiplets. There are several reasons for theorists to believe in SUSY:

- cancellation of certain divergences in the Standard Model;
- unification of all interactions, probably including gravitation (?);
- mathematical elegance.

In the simplest variant of SUSY, each known particle has the corresponding "superpartner", differing (in case of an exact SUSY) only by its spin: to a photon with $s = 1$ there corresponds a photino with $s = 1/2$, to an electron with $s = 1/2$ there corresponds an electrino with $s = 0$, to quarks with $s = 1/2$ there corresponds squarks with $s = 0$, etc. Supersymmetry is definitely strongly broken (by mass); the search for superpartners is also one of the major tasks for LHC. Preliminary results from LHC produced no evidence for SUSY, but the work continues. We shall not discuss sypersymmetry in this book.

Finally, beyond any doubt there should be one more fundamental particle – the *graviton*, i. e., the quantum of gravitational interactions with $s = 2$. However, gravitation is definitely outside the scope of experimental particle physics. Gravitation is too weak to be observed in particle interactions. It becomes important only for microprocesses at extremely high, the so-called *Planck energies* of the order of $E \sim m_P c^2 = (\hbar c/G)^{1/2} c^2 = 1.22 \cdot 10^{19}\,\mathrm{GeV}$. Here G is the Newtonian gravitational constant, and m_P is the so-called Planck mass ($\sim 10^{-5}$ Gramm!), which determines also the

characteristic Planck length: $\Lambda_P \sim \hbar/m_P c \sim \sqrt{\hbar G/c^3} \sim 10^{-33}$ cm. Experiments at such energies are simply unimaginable for humans. However, the effects of quantum gravitation were decisive during the Big Bang and determined the future evolution of the universe. Thus, quantum gravitation is of primary importance for relativistic cosmology. Unfortunately, quantum gravitation is still undeveloped, and for many serious reasons. Attempts to quantize Einstein's theory of gravitation (general relativity) meet with insurmountable difficulties, due to the strong nonlinearity of this theory. All variants of such quantization inevitably lead to a strongly *nonrenormalizable* theory, with no possibility of applying the standard methods of modern quantum field theory. These problems have been under active study for many years, with no significant progress. There are some elegant modifications of the standard theory of gravitation, such as e. g., supergravity. Especially beautiful is an idea of "induced" gravitation, suggested by Sakharov, when Einstein's theory is considered as the low-energy (phenomenological) limit of the usual quantum field theory in the curved space-time. However, up to now these ideas have not been developed enough to be of importance for experimental particle physics.

There are even more fantastic ideas which have been actively discussed during recent decades. Many people think that both quantum field theory and the Standard Model are just effective phenomenological theories, appearing in the low energy limit of the new microscopic *superstring* theory. This theory assumes that "real" microscopic theory should not deal with point-like particles, but with strings with characteristic sizes of the order of $\Lambda_P \sim 10^{-33}$ cm. These strings are moving (oscillating) in the spaces of many dimensions and possess fermion-boson symmetry (superstrings!). These ideas are now being developed for the "theory of everything".

Our aim in this book is a much more modest one. There is a funny terminology [47], according to which all theories devoted to particles which have been and will be discovered in the near future are called "phenomenological", while theories devoted to particles or any entities, which will *never* be discovered experimentally, are called "theoretical". In this sense, we are not dealing here with "fundamental" theory at all. However, we shall see that there are too many interesting problems even at this "low" level.

Chapter 2

Lagrange formalism. Symmetries and gauge fields

2.1 Lagrange mechanics of a particle

Let us recall first of all some basic principles of classical mechanics. Consider a particle (material point) with mass m, moving in some potential $V(x)$. For simplicity we consider one-dimensional motion. At the time moment t the particle is at point $x(t)$ of its trajectory, which connects the initial point $x(t_1)$ with the finite point $x(t_2)$, as shown in Figure 2.1(a). This trajectory is determined by the solution of Newton's equation of motion:

$$m\frac{d^2x}{dt^2} = F(x) = -\frac{dV(x)}{dx} \tag{2.1}$$

with appropriate initial conditions. This equation can be "derived" from the principle of least action. We introduce the Lagrange function as the difference between kinetic and potential energy:

$$L = T - V = \frac{m}{2}\left(\frac{dx}{dt}\right)^2 - V(x) \tag{2.2}$$

and the action integral

$$S = \int_{t_1}^{t_2} dt\, L(x, \dot{x})\,, \tag{2.3}$$

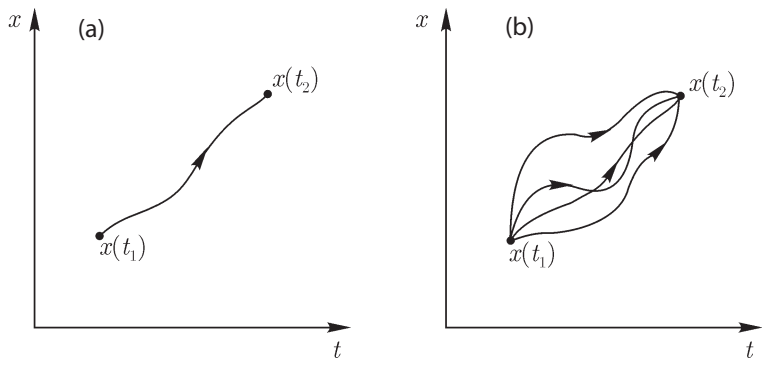


Figure 2.1. (a) Trajectory, corresponding to the least action. (b) The set of arbitrary trajectories of the particle.

where as usual $\dot{x}$ denotes velocity $\dot{x} = dx/dt$. The true trajectory of the particle corresponds to the minimum (in general extremum) of the action on the whole set of arbitrary trajectories, connecting points $x(t_1)$ and $x(t_2)$, as shown in Figure 2.1(b). From this principle we can immediately obtain the classical equations of motion. Consider the arbitrary small variation $a(t)$ of the true trajectory $x(t)$:

$$x(t) \to x'(t) = x(t) + a(t)\,. \tag{2.4}$$

At the initial and final points this variation is naturally assumed to be zero:

$$a(t_1) = a(t_2) = 0\,. \tag{2.5}$$

Substituting (2.4) into action (2.3) we obtain its variation as

$$\begin{aligned} S \to S' &= \int\limits_{t_1}^{t_2} dt \left[\frac{m}{2}(\dot{x}+\dot{a})^2 - V(x+a)\right] = \\ &= \int\limits_{t_1}^{t_2} dt \left[\frac{1}{2}m\dot{x}^2 + m\dot{x}\dot{a} - V(x) - aV'(x)\right] + O(a^2) = \\ &= S + \int\limits_{t_1}^{t_2} dt\,[m\dot{x}\dot{a} - aV'(x)] \equiv S + \delta S\,, \end{aligned} \tag{2.6}$$

where $V' = dV/dx$, so that

$$\delta S = \int\limits_{t_1}^{t_2} dt\,[m\dot{x}\dot{a} - aV'(x)]\,. \tag{2.7}$$

The action is extremal at $x(t)$ if $\delta S = 0$. Integrating the first term in (2.7) by parts, we get

$$\int\limits_{t_1}^{t_2} dt\,\dot{x}\dot{a} = \dot{x}a\Big|_{t_1}^{t_2} - \int\limits_{t_1}^{t_2} dt\,a\ddot{x} = -\int\limits_{t_1}^{t_2} dt\,a\ddot{x}\,, \tag{2.8}$$

as variations at the ends of trajectory are fixed by equation (2.5). Then

$$\delta S = -\int\limits_{t_1}^{t_2} dt\,a[m\ddot{x} + V'(x)] = 0\,, \tag{2.9}$$

Due to the arbitrariness of variation a we immediately obtain Newton's law (2.1):

$$m\ddot{x} = -V'(x)\,,$$

which determines the (single!) true trajectory of the classical particle.

2.2 Real scalar field. Lagrange equations

The transition from the classical mechanics of a particle to classical field theory reduces to the transition from particle trajectories to the space-time variations of field configurations, defined at each point in space-time. Analogue to the particle coordinate as a function of time $x(t)$ is the field function $\varphi(x^\mu) = \varphi(x, y, z, t)$.

Notes on relativistic notations

We use the following standard notations. Two space-time points (events) (x, y, z, t) and $x + dx, y + dy, z + dz, t + dt$ are separated by the interval

$$ds^2 = c^2dt^2 - (dx^2 + dy^2 + dz^2)\,.$$

The interval $ds^2 > 0$ is called *time-like* and the corresponding points (events) can be casually related. The interval $ds^2 < 0$ is called *space-like*; corresponding points (events) can not be casually related.

The set of coordinates

$$x^\mu = (x^0, x^1, x^2, x^3) \equiv (ct, x, y, z)$$

determines the *contravariant* components of 4-vector, while

$$x_\mu = (x_0, x_1, x_2, x_3) \equiv (ct, -x, -y, -z)$$

represents the corresponding *covariant* components. Then the interval can be written as

$$ds^2 = \sum_{\mu=0}^{3} dx^\mu dx_\mu \equiv dx^\mu dx_\mu = c^2dt^2 - dx^2 - dy^2 - dz^2\,.$$

There is an obvious relation:

$$x_\mu = g_{\mu\nu}x^\nu = g_{\mu0}x^0 + g_{\mu1}x^1 + g_{\mu2}x^2 + g_{\mu3}x^3\,,$$

where we have introduced the metric tensor in Minkowski space-time:

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}; \qquad g_{\mu\nu}g^{\nu\delta} = \delta^\delta_\mu\,.$$

For differential operators we shall use the following short notations:

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = (\partial_0, \partial_1, \partial_2, \partial_3) = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) = \left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right),$$

$$\partial^\mu = g^{\mu\nu}\partial_\nu = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right),$$

$$\Box \equiv \partial_\mu\partial^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \triangle\,.$$

For the energy-momentum vector of a particle with mass m we have

$$p^\mu = \left(\frac{E}{c}, \mathbf{p}\right), \qquad p_\mu = \left(\frac{E}{c}, -\mathbf{p}\right),$$

$$p^2 = p_\mu p^\mu = \frac{E^2}{c^2} - \mathbf{p}^2 = m^2 c^2 .$$

For typical combination, usually standing in Fourier integrals, we write

$$px = p_\mu x^\mu = Et - \mathbf{p}\cdot\mathbf{r} .$$

In the following almost everywhere we use the natural system of units with $\hbar = c = 1$. The advantages of this system, besides the obvious compactness of all expressions, and its connection with traditional systems of units, are well described in [46].

Consider the simplest example of a free *scalar* field $\varphi(x^\mu) = \varphi(x, y, z, t)$, which is attributed to particles with spin $s = 0$. This field satisfies the Klein–Gordon equation:

$$(\Box + m^2)\varphi = 0 . \tag{2.10}$$

Historically this equation was obtained as a direct relativistic generalization of the Schroedinger equation. If we consider $\varphi(x_\mu)$ as a wave function of a particle and take into account relativistic dispersion (spectrum)

$$E^2 = \mathbf{p}^2 + m^2 , \tag{2.11}$$

we can perform the standard Shroedinger replacement of dynamic variables by operators, acting on the wave function:

$$\mathbf{p} \to \frac{\hbar}{i}\frac{\partial}{\partial \mathbf{r}}, \qquad E \to i\hbar\frac{\partial}{\partial t}, \tag{2.12}$$

which immediately gives (2.10). Naturally, this procedure is not a *derivation*, and a more consistent procedure for obtaining relativistic field equations is based on the *principle of least action*.

Let us introduce the *action functional* as

$$S = \int d^4x\, \mathcal{L}(\varphi, \partial_\mu \varphi) , \tag{2.13}$$

where $\mathcal{L}$ is the *Lagrangian* (Lagrange function density) of the system of fields. The Lagrange function is $L = \int d^3\mathbf{r}\, \mathcal{L}$. It is usually assumed that $\mathcal{L}$ depends on the field φ and its first derivatives. The Klein–Gordon equation is easily derived from the following Lagrangian:

$$\mathcal{L} = \frac{1}{2}(\partial^\mu \varphi)(\partial_\mu \varphi) - \frac{m^2}{2}\varphi^2 = \frac{1}{2}\left[(\partial_0\varphi)^2 - (\nabla\varphi)^2 - m^2\varphi^2\right] . \tag{2.14}$$

This directly follows from the general Lagrange formalism in field theory. However, before discussing this formalism it is useful to read the following.

Notes on dimensionalities

In our system of units with $\hbar = c = 1$ dimensionalities of energy, mass, and inverse length are just the same: $[energy] = [mass] = l^{-1}$. To understand the last equality we remind that the Compton length for a particle with mass m is determined as $\hbar/mc$. The action $S = \int d^4x\, \mathcal{L}$ has the dimensionality of $\hbar$, so that in our system of units it is *dimensionless*! Then the dimensionality of Lagrangian is $[\mathcal{L}] = l^{-4}$. Accordingly, from equation (2.14) we obtain the dimensionality of the scalar field as $[\varphi] = l^{-1}$. This type of dimensionality analysis will be used many times in the following.

Now let us turn to the general Lagrange formalism of the field theory. Consider the field φ filling some space-time region (volume) $\mathcal{R}$ in Minkowski space. As initial and final hypersurfaces in this space we can take time slices at $t = t_1$ and $t = t_2$. Consider now arbitrary (small) variations of coordinates and fields:

$$x^\mu \to x'^\mu = x^\mu + \delta x^\mu\,, \tag{2.15a}$$

$$\varphi(x) \to \varphi'(x) = \varphi(x) + \delta\varphi(x)\,. \tag{2.15b}$$

Here we assume these variations δx^μ and $\delta\varphi(x)$ to be fixed at zero at the boundaries of our space-time region $\tilde{\mathcal{R}}$:

$$\delta\varphi(x) = 0\,, \quad \delta x^\mu = 0\,, \quad x \in \tilde{\mathcal{R}}\,. \tag{2.16}$$

Let us analyze the sufficiently general case, when the Lagrangian $\mathcal{L}$ is explicitly dependent of coordinates x^μ, which may correspond to the situation when our fields interact with external sources. Total variation of the field can be written as

$$\varphi'(x') = \varphi(x) + \Delta\varphi(x)\,, \tag{2.17}$$

where

$$\Delta\varphi = \varphi'(x') - \varphi(x') + \varphi(x') - \varphi(x) = \delta\varphi(x) + \delta x^\mu(\partial_\mu\varphi)\,. \tag{2.18}$$

Then action variation is given by

$$\delta S = \int_{\mathcal{R}} d^4x'\, \mathcal{L}(\varphi', \partial_\mu\varphi', x'_\mu) - \int_{\mathcal{R}} d^4x\, \mathcal{L}(\varphi, \partial_\mu\varphi, x_\mu)\,. \tag{2.19}$$

Here $d^4x' = J(x/x')d^4x$, where $J(x/x')$ is the Jacobian of transformation from x to x'. From equation (2.15a) we can see that

$$\frac{\partial x'^\mu}{\partial x^\lambda} = \delta^\mu_\lambda + \partial_\lambda \delta x^\mu \tag{2.20}$$

and for Jacobian we can write down the simple expression up to terms of the first order in δx^μ:

$$J(x/x') = \mathrm{Det}\left(\frac{\partial x'^\mu}{\partial x^\lambda}\right) = 1 + \partial_\mu(\delta x^\mu)\,. \tag{2.21}$$

Then

$$\delta S = \int_{\mathcal{R}} d^4x \left[\delta\mathcal{L} + \mathcal{L}\partial_\mu \delta x^\mu\right], \tag{2.22}$$

where

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta(\partial_\mu\varphi) + \frac{\partial\mathcal{L}}{\partial x^\mu}\delta x^\mu . \tag{2.23}$$

From equation (2.15a) it is clear that $\delta(\partial_\mu\varphi) = \partial_\mu\delta\varphi$, so that from equations (2.22) and (2.23) it immediately follows that

$$\delta S = \int_{\mathcal{R}} d^4x \left\{\frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\partial_\mu(\delta\varphi) + \partial_\mu(\mathcal{L}\delta x^\mu)\right\} . \tag{2.24}$$

The third term in figure brackets reduces to full divergence, so that this contribution is transformed (using the Gauss theorem) into the integral over the boundary surface $\mathcal{R}$. The second term in equation (2.24) can also be transformed to an expression containing full divergence:

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\partial_\mu(\delta\varphi) = \partial_\mu\left\{\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi\right\} - \partial_\mu\left\{\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right\}\delta\varphi . \tag{2.25}$$

As a result we rewrite the action variation (2.24) as

$$\delta S = \int_{\mathcal{R}} d^4x \left\{\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right]\right\}\delta\varphi + \int_{\tilde{\mathcal{R}}} d\sigma_\mu \left\{\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta\varphi + \mathcal{L}\delta x^\mu\right\} \tag{2.26}$$

Due limitations of equation (2.16), variations φ and x^μ on the boundary of integration region $\mathcal{R}$ are equal to zero, so that the surface integral in equation (2.26) reduces to zero. Then, demanding $\delta S = 0$ for arbitrary field and coordinate variations, we get

$$\frac{\partial\mathcal{L}}{\partial\varphi} - \frac{\partial}{\partial x^\mu}\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right] = 0 . \tag{2.27}$$

This is the general form of Lagrange equations (equations of motion) for the field φ[1].

Let us write down the Lagrangian of a scalar field (2.14) as a simplest quadratic form of the field and its first derivatives:

$$\mathcal{L} = \frac{1}{2}g^{\mu\nu}(\partial_\mu\varphi)(\partial_\nu\varphi) - \frac{1}{2}m^2\varphi^2 .$$

Then we have

$$\frac{\partial\mathcal{L}}{\partial\varphi} = -m^2\varphi , \qquad \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)} = g^{\mu\nu}(\partial_\nu\varphi) = \partial^\mu\varphi \tag{2.28}$$

and Lagrange equation reduces to the Klein–Gordon equation:

$$\partial_\mu\partial^\mu\varphi + m^2\varphi \equiv \Box\varphi + m^2\varphi = 0 . \tag{2.29}$$

[1] This derivation is actually valid for arbitrary fields, not necessarily scalar ones. In the case of vectors, tensors, or spinor fields, this equation is satisfied by all *components* of the field, which are numbered by the appropriate indices.

This is a linear differential equation, and it describes the free (noninteracting) field. If we add to the Lagrangian (2.28) higher order (higher power) invariants of field φ, we shall obtain a nonlinear equation for self-interacting scalar fields.

2.3 The Noether theorem

Let us return to equation (2.26) and rewrite the surface integral in a different form:

$$\delta S = \int_{\mathcal{R}} d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \right] \right\} \delta\varphi + \\ + \int_{\tilde{\mathcal{R}}} d\sigma_\mu \left\{ \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} [\delta\varphi + (\partial_\nu \varphi)\delta x^\nu] - \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} (\partial_\nu \varphi) - \delta^\mu_\nu \mathcal{L} \right] \delta x^\nu \right\} \quad (2.30)$$

where we just added and subtracted the same term. The expression in the first square brackets in the surface integral represents the full variation of the field, as defined in equation (2.18). The second square bracket, as we shall demonstrate below, defines the *energy-momentum tensor*:

$$\theta^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \partial_\nu \varphi - \delta^\mu_\nu \mathcal{L} \,. \quad (2.31)$$

Then δS is rewritten as

$$\delta S = \int_{\mathcal{R}} d^4x \left\{ \frac{\partial \mathcal{L}}{\partial \varphi} - \frac{\partial}{\partial x^\mu} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \right] \right\} \delta\varphi + \int_{\tilde{\mathcal{R}}} d\sigma_\mu \left\{ \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \Delta\varphi - \theta^\mu_\nu \delta x^\nu \right\} \quad (2.32)$$

Note that the first integral here is equal to zero (for arbitrary variations $\delta\varphi$) due to the validity of the equations of motion (2.27). Consider now the second term in equation (2.32). Assume that the action S is invariant with respect to some *continuous* group of transformations of x^μ and φ (Lie group). We can write the corresponding infinitesimal transformations as

$$\delta x^\mu = X^\mu_\nu \delta\omega^\nu \,. \quad \Delta\varphi = \Phi_\mu \delta\omega^\mu \,, \quad (2.33)$$

where $\delta\omega^\mu$ are infinitesimal *parameters* of group transformation ("rotation angles"), X^μ_ν is some matrix, and Φ_μ are some numbers. Note that in the general case indices here may be double, triple, etc. In particular we may consider some *multiplet* of fields φ_i, so that

$$\Delta\varphi_i = \Phi_{ij} \delta\omega_j \,, \quad (2.34)$$

where Φ is now also some matrix in some abstract ("isotopic") space.

Demanding the invariance of the action $\delta S = 0$ under transformations (2.33), from (2.32) (taking into account (2.27)) we obtain

$$\int_{\tilde{\mathcal{R}}} d\sigma_\mu \left\{ \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \Phi_\nu - \theta^\mu_\kappa X^\kappa_\nu \right\} \delta\omega^\nu = 0 \,, \quad (2.35)$$

which, due to the arbitrariness of $\delta\omega^\nu$, leads to

$$\int_{\tilde{\mathcal{R}}} d\sigma_\mu J^\mu_\nu = 0\,, \tag{2.36}$$

where

$$J^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)}\Phi_\nu - \theta^\mu_\kappa X^\kappa_\nu\,. \tag{2.37}$$

Using the Gauss theorem, from equation (2.36) we obtain the continuity equation

$$\partial_\mu J^\mu_\nu = 0\,, \tag{2.38}$$

so that J^μ_ν represents some *conserving current*. More precisely, conserving is the generalized *charge*:

$$Q_\nu = \int_\sigma d\sigma_\mu J^\mu_\nu\,, \tag{2.39}$$

where the integral is taken over the arbitrary space-like hypersurface σ. If we take σ as hyperplane $t = \text{const}$, we simply obtain the integral over the three-dimensional volume V:

$$Q_\nu = \int_V d^3\mathbf{r}\, J^0_\nu\,. \tag{2.40}$$

As usual [33], integrating (2.38) over the volume V, we have

$$\int_V d^3\mathbf{r}\, \partial_0 J^0_\nu + \int_V d^3\mathbf{r}\, \partial_i J^i_\nu = 0\,. \tag{2.41}$$

The second integral here is transformed, using the three-dimensional Gauss theorem, into the surface integral, which determines the flow of charge through this surface [33]. For the closed system (universe) this flow is zero and we obtain

$$\frac{d}{dt}\int_V d^3\mathbf{r}\, J^0_\nu = \frac{dQ_\nu}{dt} = 0\,. \tag{2.42}$$

This is the main statement of the Noether theorem: *invariance of the action with respect to some continuous symmetry group leads to the corresponding conservation law*.

Consider the simple example. Let symmetry transformations (2.33) be the simple *space-time translations*

$$\delta x^\mu = \varepsilon^\mu\,, \qquad \Delta\varphi = 0\,, \tag{2.43}$$

so that

$$X^\mu_\nu = \delta^\mu_\nu\,, \qquad \Phi_\mu = 0\,. \tag{2.44}$$

Then from equation (2.37) we immediately obtain

$$J^\mu_\nu = -\theta^\mu_\nu \tag{2.45}$$

and the corresponding conservation law is given by

$$\frac{d}{dt}\int_V d^3\mathbf{r}\,\theta_\nu^0 = 0\,, \tag{2.46}$$

which represents the conservation of energy and momentum, and confirms the definition of the energy-momentum tensor given above. Here

$$P_\nu = \int_V d^3\mathbf{r}\,\theta_\nu^0 \tag{2.47}$$

defines the 4-momentum of our field. This is also clear from the simple analogy with classical mechanics. In particular, from definition (2.31) it follows that

$$\int_V d^3\mathbf{r}\,\theta_0^0 = \int_V d^3\mathbf{r}\left\{\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\dot{\varphi} - \mathcal{L}\right\}, \tag{2.48}$$

which is similar to the well-known expression relating Lagrange function with the Hamiltonian of classical mechanics [34]:

$$H = \sum_i p_i\dot{q}_i - L\,, \quad p_i = \frac{\partial L}{\partial\dot{q}_i}\,, \tag{2.49}$$

so that equation (2.48) gives the energy of the field. Similarly, the value of $\int d^3\mathbf{r}\,\theta_i^0$ determines the momentum of the field.

Thus, energy-momentum conservation is valid for any system with the Lagrangian (action) independent of x^μ (explicitly).

For the Klein–Gordon Lagrangian (2.28) from (2.31) we immediately obtain the energy-momentum tensor as

$$\theta^{\mu\nu} = (\partial^\mu\varphi)(\partial^\nu\varphi) - g^{\mu\nu}\mathcal{L}\,. \tag{2.50}$$

This expression is explicitly symmetric over indices $\theta^{\mu\nu} = \theta^{\nu\mu}$. However, it is not always so if are using the definition of equation (2.31) for an arbitrary Lagrangian. At the same time, we can always add to (2.31) an additional term like $\partial_\lambda f^{\mu\lambda\nu}$, where $f^{\mu\lambda\nu} = -f^{\lambda\mu\nu}$, so that $\partial_\mu\partial_\lambda f^{\lambda\mu\nu} \equiv 0$ and conservation laws (2.38), (2.46) are not broken. We can use this indeterminacy and introduce

$$T^{\mu\nu} = \theta^{\mu\nu} + \partial_\lambda f^{\lambda\mu\nu}\,, \tag{2.51}$$

choosing some specific $f^{\lambda\mu\nu}$ to guarantee the symmetry condition $T^{\mu\nu} = T^{\nu\mu}$. In this case the energy-momentum tensor is called canonical. Naturally we have

$$\partial_\mu T^{\mu\nu} = \partial_\mu\theta^{\mu\nu} = 0\,. \tag{2.52}$$

The total 4-momentum in this case is also unchanged, as

$$\int_V d^3\mathbf{r}\,\partial_\lambda f^{\lambda 0\nu} = \int_V d^3\mathbf{r}\,\partial_i f^{i0\nu} = \int d\sigma_i\, f^{i0\nu} = 0\,. \tag{2.53}$$

The first equality in equation (2.53) follows from $f^{00\nu} = 0$, and the second one follows from the Gauss theorem. The zero in the right-hand side appears when the surface σ is moved to the infinity, where fields are assumed to be absent.

Thus, both the energy and momentum of the field are determined unambiguously, despite some indeterminacy of the energy-momentum tensor.

There are certain physical reasons to require the energy-momentum tensor to always be symmetric [33, 56]. An especially elegant argument follows from general relativity. Einstein's equations for gravitational field (space-time metric $g_{\mu\nu}$) has the form [33]

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -\frac{8\pi G}{c^2} T_{\mu\nu} \,, \tag{2.54}$$

where $R_{\mu\nu}$ is Riemann's curvature tensor, simplified by two indices (Ricci tensor), R is the scalar curvature of space, and G is the Newtonian gravitational constant. The left-hand side of equation (2.54) is built of the metric tensor $g_{\mu\nu}$ and its derivatives, and by definition it is a purely geometric object. It can be shown to be always symmetric over indices μ, ν [33]. Then, the energy-momentum tensor in the right-hand side, which is the source of the gravitational field, should also be symmetric.

2.4 Complex scalar and electromagnetic fields

Consider now the complex scalar field, which can be conveniently written as

$$\varphi = \frac{1}{\sqrt{2}} (\varphi_1 + i\varphi_2) \,, \tag{2.55a}$$

$$\varphi^* = \frac{1}{\sqrt{2}} (\varphi_1 - i\varphi_2) \,. \tag{2.55b}$$

In fact we are considering here *two* independent scalar fields φ_1, φ_2, which can be representing, e. g., two projections of some two-dimensional vector on axis 1 and 2 in some *isotopic*[2] space, associated with our field. Requiring the action to be real, the Lagrangian of our field, similar to (2.28), can be written as

$$\mathcal{L} = (\partial_\mu \varphi)(\partial^\mu \varphi^*) - m^2 \varphi^* \varphi \,. \tag{2.56}$$

Considering fields φ and φ^* to be independent variables, we obtain from the Lagrange equations (2.27) two Klein–Gordon equations:

$$(\Box + m^2)\varphi = 0 \,, \tag{2.57a}$$

$$(\Box + m^2)\varphi^* = 0 \,. \tag{2.57b}$$

[2] The term "isotopic" as used by us is in most cases not related to the isotopic symmetry of hadrons in nuclear and hadron physics [40]. In fact, we are speaking about some space of internal quantum numbers of fields (particles), conserving due to appropriate symmetry in this associated space (not related to space-time).

The Lagrangian (2.56) is obviously invariant with respect to the so-called global[3] gauge transformations:

$$\varphi \to e^{-i\Lambda}\varphi\,, \qquad \varphi^* \to e^{i\Lambda}\varphi^*\,, \tag{2.58}$$

where Λ is an arbitrary real constant. Equation (2.58) is the typical Lie group transformation (in this case it is the $U(1)$ group of two-dimensional rotations), accordingly; for small Λ we can always write

$$\delta\varphi = -i\Lambda\varphi\,, \qquad \delta\varphi^* = i\Lambda\varphi^* \tag{2.59}$$

i. e., as the infinitesimal gauge transformation. Due to the independence of Λ on space-time coordinates, the infinitesimal transformation of field derivatives has the same form:

$$\delta(\partial_\mu\varphi) = -i\Lambda\partial_\mu\varphi\,, \qquad \delta(\partial_\mu\varphi^*) = i\Lambda\partial_\mu\varphi^*\,. \tag{2.60}$$

In the notations of equation (2.33) we have

$$\Phi = -i\varphi\,, \qquad \Phi^* = i\varphi\,, \qquad X = 0\,, \tag{2.61}$$

so that conserving Noether current (2.37) in this case takes the following form:

$$J^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}(-i\varphi) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^*)}(i\varphi^*)\,. \tag{2.62}$$

With the account of (2.56) we get

$$J^\mu = i(\varphi^*\partial^\mu\varphi - \varphi\partial^\mu\varphi^*) \tag{2.63}$$

i. e., the explicit form of the current, satisfying the equation

$$\partial_\mu J^\mu = 0\,. \tag{2.64}$$

This may be checked also directly, using equations of motion (2.57). Accordingly, in this theory we get the conserving charge

$$Q = \int dV J^0 = i\int dV\left(\varphi^*\frac{\partial\varphi}{\partial t} - \varphi\frac{\partial\varphi^*}{\partial t}\right). \tag{2.65}$$

If the field is real, i. e., $\varphi = \varphi^*$, we obviously get $Q = 0$, so that the concept of conserving the charge with $dQ/dt = 0$ can be defined only for a complex field. This is the decisive role of $U(1)$ symmetry of Lagrangian (2.56), (2.58). Note that our entire discussion up to now is purely classical; accordingly Q may acquire arbitrary (noninteger) values.

[3] The term "global" means that the arbitrary phase Λ here is the same for fields, taken at different space-time points.

Let us rewrite (2.56), using (2.55), as the additive sum of Lagrangians for fields φ_1, φ_2:

$$\mathcal{L} = \frac{1}{2}\left[(\partial_\mu\varphi_1)(\partial^\mu\varphi_1) + (\partial_\mu\varphi_2)(\partial^\mu\varphi_2)\right] - \frac{1}{2}m^2(\varphi_1^2 + \varphi_2^2)\,. \tag{2.66}$$

Then, writing the field φ as a vector $\vec{\varphi}$ in two-dimensional isotopic space,

$$\vec{\varphi} = \varphi_1\vec{i} + \varphi_2\vec{j}\,, \tag{2.67}$$

where $\vec{i}$, $\vec{j}$ are unit vectors along axes in this space, we can write (2.66) as

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\vec{\varphi})(\partial^\mu\vec{\varphi}) - \frac{1}{2}m^2\vec{\varphi}\cdot\vec{\varphi}\,, \tag{2.68}$$

which clearly demonstrates the geometric meaning of this symmetry of the Lagrangian. The gauge transformations (2.58) can be written also as

$$\varphi_1' + i\varphi_2' = e^{-i\Lambda}(\varphi_1 + i\varphi_2)\,, \qquad \varphi_1' - i\varphi_2' = e^{i\Lambda}(\varphi_1 - i\varphi_2)\,,$$

or

$$\begin{aligned}\varphi_1' &= \quad \varphi_1\cos\Lambda + \varphi_2\sin\Lambda\,,\\ \varphi_2' &= -\varphi_1\sin\Lambda + \varphi_2\cos\Lambda\,,\end{aligned} \tag{2.69}$$

which describes the rotation of the vector $\vec{\varphi}$ by angle Λ in the $1,2$-plane. Our Lagrangian is obviously invariant with respect to these rotations, described by the two-dimensional rotation group $O(2)$, or the isomorphic $U(1)$ group. Transformation (2.58) is unitary: $e^{i\Lambda}(e^{i\Lambda})^* = 1$. Group space is defined as the set of all possible angles Λ, determined up to $2\pi n$ (where n is an integer and the rotation by angle Λ is equivalent to rotations by $\Lambda + 2\pi n$), which is topologically equivalent to a circle of unit radius.

Now we going to take a decisive step! We can ask rather the formal question of whether or not we can make our theory invariant with respect to *local* gauge transformations, similar to (2.58), but with a phase (angle) which is an arbitrary *function* of the space-time point, where our field is defined

$$\varphi(x) \to e^{-i\Lambda(x)}\varphi(x)\,, \qquad \varphi^*(x) \to e^{i\Lambda(x)}\varphi^*(x)\,. \tag{2.70}$$

There are no obvious reasons for such a wish. In principle, we can only say that the global transformation (2.58) does not look very beautiful from the point of view of relativistic "ideology", as we are "rotating" our field by the same angle (in isotopic space) in all space-time points, including those separated by space-like interval (which cannot be casually related to each other). At the same time, isotopic space is in no way related to Minkowski space-time. However, we shall see shortly that demanding the invariance of the theory with respect to (2.70) will immediately lead to rather remarkable results.

Naively, the invariance of the theory with respect to (2.70) is just impossible. Consider once again infinitesimal transformations with $\Lambda(x) \ll 1$. Then (2.70) reduces to

$$\varphi \to \varphi - i\Lambda\varphi\,, \quad \delta\varphi = -i\Lambda\varphi\,, \tag{2.71}$$

which is identical to (2.59). However, for field derivatives the situation is more complicated due to explicit dependence $\Lambda(x)$ on the coordinate:

$$\partial_\mu\varphi \to \partial_\mu\varphi - i(\partial_\mu\Lambda)\varphi - i\Lambda(\partial_\mu\varphi)\,, \quad \delta(\partial_\mu\varphi) = -i\Lambda(\partial_\mu\varphi) - i(\partial_\mu\Lambda)\varphi\,, \tag{2.72}$$

which, naturally, does not coincide with (2.60). For a complex conjugate field everything is similar:

$$\varphi^* \to \varphi^* + i\Lambda\varphi^*\,, \quad \delta\varphi^* = i\Lambda\varphi^*\,, \tag{2.73}$$

$$\partial_\mu\varphi^* \to \partial_\mu\varphi^* + i(\partial_\mu\Lambda)\varphi^* + i\Lambda(\partial_\mu\varphi^*)\,, \quad \delta(\partial_\mu\varphi^*) = i\Lambda(\partial_\mu\varphi^*) + i(\partial_\mu\Lambda)\varphi^*\,. \tag{2.74}$$

This means that field derivatives of φ are transformed (in contrast to the field itself) in a noncovariant way, i. e., not proportionally to itself. The problem is with the derivative of Λ! The Lagrangian (2.56) is obviously noninvariant to these transformations. Let us look, however, whether we can somehow guarantee it.

The change of the Lagrangian under arbitrary variations of fields and field derivatives is written as

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\delta(\partial_\mu\varphi) + (\varphi \to \varphi^*)\,. \tag{2.75}$$

Rewriting the first term using the Lagrange equations (2.27) and substituting (2.71) into (2.72), we obtain

$$\begin{aligned}\delta\mathcal{L} &= \partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right](-i\Lambda\varphi) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}(-i\Lambda\partial_\mu\varphi - i\varphi\partial_\mu\Lambda) - (\varphi \to \varphi^*)\\ &= -i\Lambda\partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\varphi\right] - i\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}(\partial_\mu\Lambda)\varphi - (\varphi \to \varphi^*)\,.\end{aligned} \tag{2.76}$$

The first term here is proportional to the divergence of the conserving current (2.62) and gives zero. The second term, using the explicit form of the Lagrangian, is rewritten as

$$\delta\mathcal{L} = i(\varphi^*\partial^\mu\varphi - \varphi\partial^\mu\varphi^*)\partial_\mu\Lambda = J^\mu\partial_\mu\Lambda\,, \tag{2.77}$$

where J^μ is again the same conserving current (2.63).

Thus, the action is noninvariant with respect to local gauge transformations. However, we *can* guarantee such invariance of the action by introducing the new *vector* field A_μ, directly *interacting* with current J^μ, adding to the Lagrangian the following interaction term:

$$\mathcal{L}_1 = -eJ^\mu A_\mu = -ie(\varphi^*\partial^\mu\varphi - \varphi\partial^\mu\varphi^*)A_\mu\,, \tag{2.78}$$

where e is a dimensionless coupling constant. Let us require that local gauge transformations of the field φ (2.70) are accompanied by the *gradient* transformations of A_μ:

$$A_\mu \to A_\mu + \frac{1}{e}\partial_\mu \Lambda\,. \tag{2.79}$$

Then we obtain

$$\delta\mathcal{L}_1 = -e(\delta J^\mu)A_\mu - eJ^\mu(\delta A_\mu) = -e(\delta J^\mu)A_\mu - J^\mu\partial_\mu\Lambda\,. \tag{2.80}$$

Now we see that the second term in (2.80) precisely cancels (2.77). But we also need to eliminate the first term in (2.80). With the help of (2.71) and (2.73) we can get

$$\delta J^\mu = i\delta(\varphi^*\partial^\mu\varphi - \varphi\partial^\mu\varphi^*) = 2\varphi^*\varphi\partial^\mu\Lambda\,, \tag{2.81}$$

so that

$$\delta\mathcal{L} + \delta\mathcal{L}_1 = -2eA_\mu(\partial^\mu\Lambda)\varphi^*\varphi\,. \tag{2.82}$$

But let us add to $\mathcal{L}$ one more term:

$$\mathcal{L}_2 = e^2A_\mu A^\mu\varphi^*\varphi\,. \tag{2.83}$$

Then, under the influence of (2.79) we have

$$\delta\mathcal{L}_2 = 2e^2A_\mu\delta A^\mu\varphi^*\varphi = 2eA_\mu(\partial^\mu\Lambda)\varphi^*\varphi\,. \tag{2.84}$$

Then, it is easily seen that

$$\delta\mathcal{L} + \delta\mathcal{L}_1 + \delta\mathcal{L}_2 = 0\,, \tag{2.85}$$

so that the invariance of the action with respect to local gauge transformations is guaranteed!

Let us now take into account that the new vector field A_μ should also produce the appropriate "free" contribution to the Lagrangian. This term should be invariant to gradient transformations (2.79). It is quite clear how we now proceed. Let us introduce the 4-vector of the curl of the field A_μ:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,, \tag{2.86}$$

which is obviously invariant with respect to (2.79). Then we can introduce

$$\mathcal{L}_3 = -\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}\,. \tag{2.87}$$

Collecting all terms of the new Lagrangian, we get

$$\begin{aligned}\mathcal{L}_{tot} = \mathcal{L} + \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 &= (\partial_\mu\varphi)(\partial^\mu\varphi^*) - m^2\varphi^*\varphi \\ &- ie(\varphi^*\partial^\mu\varphi - \varphi\partial^\mu\varphi^*)A_\mu + e^2A_\mu A^\mu\varphi^*\varphi - \frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu}\,,\end{aligned} \tag{2.88}$$

which is rewritten as

$$\mathcal{L}_{tot} = -\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu} + (\partial_\mu + ieA_\mu)\varphi(\partial^\mu - ieA^\mu)\varphi^* - m^2\varphi^*\varphi\,. \tag{2.89}$$

Thus, we obtained the Lagrangian of *electrodynamics* of the complex scalar field φ! It is easily obtained from the initial Klein–Gordon Lagrangian (2.56) by the standard replacement [33] of the usual derivative $\partial_\mu\varphi$ by the *covariant* derivative[4]:

$$D_\mu\varphi = (\partial_\mu + ieA_\mu)\varphi \tag{2.90}$$

and the addition of the term, corresponding to the free electromagnetic field (2.87).

The Lagrangian of an electromagnetic field (2.87) can be written as $\mathcal{L} = aF_{\mu\nu}F^{\mu\nu}$ [33], where the constant a can be chosen to be different, depending on the choice of the system of units. In the Gaussian system of units, used e. g., by Landau and Lifshitz, it is taken as $a = -1/16\pi$. In the Heaviside system of units (see e. g., [56]) $a = -1/4$, In this system there is no factor of 4π in field equations, but instead it appears in Coulomb's law. In a Gaussian system, on the opposite, 4π enters Maxwell equations, but is absent in Coulomb's law. In the literature on quantum electrodynamics, in most cases the Heaviside system is used. However, below we shall mainly use the Gaussian system, with special remarks, when using Heaviside system.

In contrast to $\partial_\mu\varphi$ the value of (2.90) is transformed under gauge transformation covariantly, i. e., as the field φ itself:

$$\delta(D_\mu\varphi) = \delta(\partial_\mu\varphi) + ie(\delta A_\mu)\varphi + ieA_\mu\delta\varphi = -i\Lambda(\partial_\mu\varphi + ieA_\mu\varphi) = -i\Lambda(D_\mu\varphi)\,. \tag{2.91}$$

The field φ is now associated with an electric charge e, the conjugate field φ^* corresponds to the charge $(-e)$:

$$\big(D_\mu\varphi\big)^* = (\partial_\mu - ieA_\mu)\varphi^*\,. \tag{2.92}$$

It is clear that $F_{\mu\nu}$, introduced above, represents the usual tensor of electromagnetic fields [33].

Maxwell equations follow from (2.89) as Lagrange equations for the A_μ field:

$$\frac{\partial\mathcal{L}}{\partial A_\mu} - \partial_\nu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\nu A_\mu)}\right] = 0\,, \tag{2.93}$$

which reduces to

$$\begin{aligned}\frac{1}{4\pi}\partial_\nu F^{\mu\nu} &= -ie(\varphi^*\partial^\mu\varphi - \varphi\partial^\mu\varphi^*) + 2e^2A^\mu|\varphi|^2 = \\ &= -ie\big[\varphi^*D^\mu\varphi - \varphi(D^\mu\varphi^*)\big] \equiv -e\mathcal{J}^\mu\,,\end{aligned} \tag{2.94}$$

[4] The constant e means the electric charge.

where

$$\mathcal{J}^{\mu} = i\left[\varphi^{*} D^{\mu}\varphi - \varphi(D^{\mu}\varphi^{*})\right] \tag{2.95}$$

is the covariant form of the current. From the antisymmetry of $F^{\mu\nu}$ it immediately follows that

$$\partial_{\mu}\mathcal{J}^{\mu} = 0\,, \tag{2.96}$$

so that in the presence of electromagnetic field the conserved current is $\mathcal{J}^{\mu}$, not J^{μ}.

Note that electromagnetic field is massless and that this is absolutely necessary – if we attribute to an electromagnetic field a finite mass M, we have to add to the Lagrangian (2.87) an additional term such as

$$\mathcal{L}_{M} = \frac{1}{8\pi} M^{2} A_{\mu} A^{\mu}\,. \tag{2.97}$$

It is obvious that such a contribution is noninvariant with respect to local gauge transformations (2.70), (2.79).

This way of introducing an electromagnetic field was used apparently for the first time by Weyl during his attempts to formulate the unified field theory in the 1920s. Electrodynamics corresponds to the Abelian gauge group $U(1)$, and the electromagnetic field is the simplest example of a *gauge* field.

2.5 Yang–Mills fields

Introducing the invariance to local gauge transformations of the $U(1)$ group, we obtain from the Lagrangian of a free Klein–Gordon field the Lagrangian of scalar electrodynamics, i. e., the field theory with quite nontrivial interaction. We can say that the symmetry "dictated" to us the form of interaction and leads to the necessity of introducing the gauge field A_{μ}, which is responsible for this interaction. Gauge group $U(1)$ is Abelian. The generalization of gauge field theory to non-Abelian gauge groups was proposed at the beginning of the 1950s by Yang and Mills. This opened the way for construction of the wide class of nontrivial theories of interacting quantum fields, which were quite successfully applied to the foundations of the modern theory of dynamics of elementary particles.

The simplest version of a non-Abelian gauge group, analyzed in the first paper by Yang and Mills, is the group of isotopic spin, $SU(2)$, which is isomorphic to the three-dimensional rotation group $O(3)$. Previously we considered the complex scalar field which is represented by the two-dimensional vector $\vec{\varphi} = (\varphi_1, \varphi_2)$ in "isotopic" space. Consider instead the scalar field, which is simultaneously a three-dimensional vector in some "isotopic" space: $\vec{\varphi} = (\varphi_1, \varphi_2, \varphi_3)$. The Lagrangian of this Klein–Gordon field, which is invariant to three-dimensional rotations in this "associated" space, can be written as

$$\mathcal{L} = \frac{1}{2}(\partial_{\mu}\vec{\varphi})(\partial^{\mu}\vec{\varphi}) - \frac{1}{2}m^{2}\vec{\varphi}\cdot\vec{\varphi}\,, \tag{2.98}$$

where the field $\vec{\varphi}$ enters only via its scalar products. Invariance with respect to rotations here is global – the field $\vec{\varphi}$ is rotated by an arbitrary angle in isotopic space, which is the same for fields in all space-time points. For example, we can consider rotation in the$1-2$-plane by angle Λ_3 around the axis 3:

$$\begin{aligned} \varphi_1' &= \quad \varphi_1 \cos\Lambda_3 + \varphi_2 \sin\Lambda_3 \,, \\ \varphi_2' &= -\varphi_1 \sin\Lambda_3 + \varphi_2 \cos\Lambda_3 \,, \\ \varphi_3' &= \quad \varphi_3 \,. \end{aligned} \tag{2.99a}$$

For infinitesimal rotation $\Lambda_3 \ll 1$ and we can write

$$\varphi_1' = \varphi_1 + \Lambda_3 \varphi_2 \,,$$

$$\varphi_2' = \varphi_2 - \Lambda_3 \varphi_1 \,, \tag{2.99b}$$

$$\varphi_3' = \varphi_3 \,. \tag{2.99c}$$

For infinitesimal rotation around an arbitrarily oriented axis we write

$$\vec{\varphi} \to \vec{\varphi}' = \vec{\varphi} - \vec{\Lambda} \times \vec{\varphi} \,, \quad \delta\vec{\varphi} = -\vec{\Lambda} \times \vec{\varphi} \,, \tag{2.99d}$$

where vector $\vec{\Lambda}$ is directed along the rotation axis and its value is equal to the rotation angle.

Consider now the local transformation, assuming $\vec{\Lambda} = \vec{\Lambda}(x_\mu)$. Then the field derivative $\vec{\varphi}$ is transformed in a noncovariant way:

$$\begin{aligned} \partial_\mu \vec{\varphi} \to \partial_\mu \vec{\varphi}' &= \partial_\mu \vec{\varphi} - \partial_\mu \vec{\Lambda} \times \vec{\varphi} - \vec{\Lambda} \times \partial_\mu \vec{\varphi} \,, \\ \delta(\partial_\mu \vec{\varphi}) &= -\vec{\Lambda} \times \partial_\mu \vec{\varphi} - \partial_\mu \vec{\Lambda} \times \vec{\varphi} \,. \end{aligned} \tag{2.100}$$

Let us again try to construct the covariant derivative, writing it as

$$D_\mu \vec{\varphi} = \partial_\mu \vec{\varphi} + g \vec{W}_\mu \times \vec{\varphi} \,. \tag{2.101}$$

where we have introduced the gauge field (Yang–Mills field) $\vec{W}_\mu$, which is the vector not only in Minkowski space, but also in an associated isotopic space, and g is the coupling constant.

Covariance means that

$$\delta(D_\mu \vec{\varphi}) = -\vec{\Lambda} \times (D_\mu \vec{\varphi}) \,. \tag{2.102}$$

What transformation rules for field $\vec{W}_\mu$ are necessary to guarantee covariance? The answer is

$$\begin{aligned} \vec{W}_\mu \to \vec{W}_\mu' &= \vec{W}_\mu - \vec{\Lambda} \times \vec{W}_\mu + \frac{1}{g} \partial_\mu \vec{\Lambda} \,, \\ \delta \vec{W}_\mu &= -\vec{\Lambda} \times \vec{W}_\mu + \frac{1}{g} \partial_\mu \vec{\Lambda} \,. \end{aligned} \tag{2.103}$$

To check this, use (2.99d), (2.100), and (2.101) to obtain

$$\begin{aligned}\delta(D_\mu\vec{\varphi}) &= \delta(\partial_\mu\vec{\varphi}) + g(\delta\vec{W}_\mu)\times\vec{\varphi} + g\vec{W}_\mu\times(\delta\vec{\varphi})\\ &= -\vec{\Lambda}\times\partial_\mu\vec{\varphi} - \partial_\mu\vec{\Lambda}\times\vec{\varphi} - g(\vec{\Lambda}\times\vec{W}_\mu)\times\vec{\varphi} + \partial_\mu\vec{\Lambda}\times\vec{\varphi} - g\vec{W}_\mu\times(\vec{\Lambda}\times\vec{\varphi})\\ &= -\vec{\Lambda}\times\partial_\mu\vec{\varphi} - g[(\vec{\Lambda}\times\vec{W}_\mu)\times\vec{\varphi} + \vec{W}_\mu\times(\vec{\Lambda}\times\vec{\varphi})]\,. \end{aligned} \tag{2.104}$$

Then use the Jacobi identity[5]:

$$(\vec{A}\times\vec{B})\times\vec{C} + (\vec{B}\times\vec{C})\times\vec{A} + (\vec{C}\times\vec{A})\times\vec{B} = 0\,, \tag{2.105}$$

Making here cyclic permutations we can obtain

$$(\vec{A}\times\vec{B})\times\vec{C} + \vec{B}\times(\vec{A}\times\vec{C}) = \vec{A}\times(\vec{B}\times\vec{C})\,. \tag{2.106}$$

Applying this identity to the expression in square brackets in (2.104), we get

$$\delta(D_\mu\vec{\varphi}) = -\vec{\Lambda}\times(\partial_\mu\vec{\varphi} + g\vec{W}_\mu\times\vec{\varphi}) = -\vec{\Lambda}\times D_\mu\vec{\varphi}\,, \tag{2.107}$$

Q.E.D.

Let us now discuss how we should write the analogue of the $F_{\mu\nu}$ tensor of electrodynamics. We shall denote it as $\vec{W}_{\mu\nu}$. In contrast to $F_{\mu\nu}$, which is a scalar with respect to $O(2)$ $(U(1))$ gauge group transformations, $\vec{W}_{\mu\nu}$ is the vector with respect to $O(3)$ $(SU(2))$. Accordingly, transformation rules should be the same, as for the field $\vec{\varphi}$:

$$\delta\vec{W}_{\mu\nu} = -\vec{\Lambda}\times\vec{W}_{\mu\nu}\,. \tag{2.108}$$

In fact, $\partial_\mu\vec{W}_\nu - \partial_\nu\vec{W}_\mu$ is not transformed in this way:

$$\begin{aligned}\delta(\partial_\mu\vec{W}_\nu - \partial_\nu\vec{W}_\mu) &= \partial_\mu\left(-\vec{\Lambda}\times\vec{W}_\nu + \frac{1}{g}\partial_\nu\vec{\Lambda}\right) - \partial_\nu\left(-\vec{\Lambda}\times\vec{W}_\mu + \frac{1}{g}\partial_\mu\vec{\Lambda}\right)\\ &= -\vec{\Lambda}\times(\partial_\mu\vec{W}_\nu - \partial_\nu\vec{W}_\mu) - (\partial_\mu\vec{\Lambda}\times\vec{W}_\nu - \partial_\nu\vec{\Lambda}\times\vec{W}_\mu)\,. \end{aligned} \tag{2.109}$$

We have here an "extra" second term. Note now that

$$\delta(g\vec{W}_\mu\times\vec{W}_\nu) = g\left(-\vec{\Lambda}\times\vec{W}_\mu + \frac{1}{g}\partial_\mu\vec{\Lambda}\right)\times\vec{W}_\nu + g\vec{W}_\mu\times\left(-\vec{\Lambda}\times\vec{W}_\nu + \frac{1}{g}\partial_\nu\vec{\Lambda}\right), \tag{2.110}$$

The first and third terms here can be united with the use of (2.106), which gives

$$\delta(g\vec{W}_\mu\times\vec{W}_\nu) = -g\vec{\Lambda}\times(\vec{W}_\mu\times\vec{W}_\nu) + (\partial_\mu\vec{\Lambda}\times\vec{W}_\nu - \partial_\nu\vec{\Lambda}\times\vec{W}_\mu)\,. \tag{2.111}$$

---

[5] This identity is easily proven using the well-known rule $(\vec{A}\times\vec{B})\times\vec{C} = \vec{B}(\vec{A}\cdot\vec{C}) - \vec{A}(\vec{B}\cdot\vec{C})$.

We see that the second term here coincides with the "extra" term in (2.109). Thus, we have to define the tensor of Yang–Mills fields as

$$\vec{W}_{\mu\nu} = \partial_\mu \vec{W}_\nu - \partial_\nu \vec{W}_\mu + g\vec{W}_\mu \times \vec{W}_\nu \,, \tag{2.112}$$

which is transformed in a correct way, i. e., according to (2.108).

Now we can write the Lagrangian of Yang–Mills theory:

$$\mathcal{L} = \frac{1}{2}(D_\mu \vec{\varphi})(D^\mu \vec{\varphi}) - \frac{1}{2}m^2 \vec{\varphi}\cdot\vec{\varphi} - \frac{1}{16\pi}\vec{W}_{\mu\nu}\cdot\vec{W}^{\mu\nu} \,. \tag{2.113}$$

Equations of motion are derived in the usual way from Lagrange equations:

$$\frac{\partial \mathcal{L}}{\partial (W^i_\mu)} = \partial_\nu \left\{ \frac{\partial \mathcal{L}}{\partial(\partial_\nu W^i_\mu)} \right\} , \tag{2.114}$$

where i is the vector index in isotopic space. Then we have

$$\partial^\nu \vec{W}_{\mu\nu} + g\vec{W}^\nu \times \vec{W}_{\mu\nu} = 4\pi g\big[(\partial_\mu \vec{\varphi}) \times \vec{\varphi} + g(\vec{W}_\mu \times \vec{\varphi}) \times \vec{\varphi}\big] \tag{2.115}$$

or, taking into account (2.101),

$$D^\nu \vec{W}_{\mu\nu} = 4\pi g (D_\mu \vec{\varphi}) \times \vec{\varphi} \equiv 4\pi g \vec{J}_\mu \,. \tag{2.116}$$

These equations are similar to Maxwell equations (2.94), but are *nonlinear* in the field $\vec{W}_\mu$. The second equality in (2.116) in fact determines the current of the field $\vec{\varphi}$, which plays the role of the "source" of the gauge (Yang–Mills) field $\vec{W}_\mu$. In the absence of "matter", i. e., for $\vec{\varphi} = 0$, from (2.115), (2.116) we have

$$D^\nu \vec{W}_{\mu\nu} = 0 \qquad \text{or} \qquad \partial^\nu \vec{W}_{\mu\nu} = -g\vec{W}^\nu \times \vec{W}_{\mu\nu} \,, \tag{2.117}$$

so that the Yang–Mills field (non-Abelian gauge field) is the source of itself[6] ("luminous light")! This is radically different from the case of the Abelian gauge field (electromagnetic field), where (Maxwell) field equations are linear [33]:

$$\partial^\nu F_{\mu\nu} = 0 \quad \text{or} \quad \operatorname{div} \mathbf{E} = 0 \,, \quad \frac{\partial \mathbf{E}}{\partial t} - \operatorname{rot} \mathbf{H} = 0 \,. \tag{2.118}$$

In standard electrodynamics we also have an additional homogeneous Maxwell equation [33]:

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0 \,, \tag{2.119}$$

from which, in three-dimensional notations, we get the second pair of electromagnetic field equations:

$$\operatorname{div} \mathbf{H} = 0 \,, \quad \frac{\partial \mathbf{H}}{\partial t} + \operatorname{rot} \mathbf{E} = 0 \,. \tag{2.120}$$

[6] The situation here is similar to general relativity, where the gravitational field is also the source of itself due to the nonlinearity of Einstein's equations [33].

The first of these equations, in particular, signifies the absence of magnetic charges (monopoles). Similar equations also exist in Yang–Mills theory (its derivation will be presented a little bit later):

$$D_\lambda \vec{W}_{\mu\nu} + D_\mu \vec{W}_{\nu\lambda} + D_\nu \vec{W}_{\lambda\mu} = 0\,. \tag{2.121}$$

The tensor of Yang–Mills fields $\vec{W}_{\mu\nu}$ can be written via corresponding non-Abelian "electric" and "magnetic" fields, in a similar way to electrodynamics [33]:

$$\vec{W}_{\mu\nu} = \begin{pmatrix} 0 & \vec{E}_x & \vec{E}_y & \vec{E}_z \\ -\vec{E}_x & 0 & -\vec{H}_z & \vec{H}_y \\ -\vec{E}_y & \vec{H}_z & 0 & -\vec{H}_x \\ -\vec{E}_z & -\vec{H}_y & \vec{H}_x & 0 \end{pmatrix}. \tag{2.122}$$

Then, it follows from (2.121) that

$$\operatorname{div} \vec{\mathbf{H}} \neq 0\,, \tag{2.123}$$

which directly leads to the existence of the so-called t'Hooft–Polyakov monopoles in Yang–Mills theory [56]. Due to the lack of space, we shall not further analyze these interesting solutions of field equations here.

The Yang–Mills field, similar to the electromagnetic field, should be massless. For the massive case we have to add to the Lagrangian (2.113) an additional term such as

$$\mathcal{L}_M = \frac{1}{8\pi} M^2 \vec{W}_\mu \cdot \vec{W}^\mu\,, \tag{2.124}$$

which will lead to the replacement of equation (2.116) by

$$D^\nu \vec{W}_{\mu\nu} = 4\pi g \vec{J}_\mu + M^2 \vec{W}_\mu\,, \tag{2.125}$$

which is explicitly noninvariant with respect to local gauge transformations.

For a rather long time, the zero mass of Yang–Mills fields under conditions of strict gauge invariance was considered to be a primary obstacle for physical applications of gauge field theories. The initial idea of these theories was [75] that, using one or another (exact or approximate and experimentally confirmed) internal symmetry of elementary particles (e. g., conservation of baryon number or isotopic spin), one can introduce *local* invariance with respect to appropriate group transformations and obtain quite nontrivial interaction Lagrangians with corresponding (Abelian or non-Abelian) gauge fields.

The gauge principle was proposed as a foundation for the theory of interacting fields. But it seems that difficulties appeared from the very beginning. The appearance of a massless gauge field immediately leads to the existence of *long range* forces, associated with this field. A typical case is electrodynamics and its long range Coulomb interaction. However, it is rather easily demonstrated that an electromagnetic field is probably the only long range force in nature (except, obviously, for gravitation!). We can see this using very simple estimates, due to Lee and Yang [39].

Consider the simplest case of an Abelian gauge field, which may be related to conservation of a baryon charge. It will lead to an additional long-range B-force, acting upon baryons. Let us compare the usual Newtonian gravitation potential with the potential energy of this hypothetical field, due to its interaction with nucleons of the earth. Consider a test-particle p with mass m_p, which is placed above the earth's surface at distance r from the earth's center. Then

$$V_{gr} = -\frac{G m_p M_E}{r}\,, \tag{2.126}$$

where G is the Newtonian gravitational constant, and M_E is the earth mass. Let the baryon charge of our test particle be N_p and the nucleon mass be m_N. Assume the density of nucleons on the earth to be constant (and there are no antinucleons at all) and equal to

$$\rho = \frac{M_E}{m_N \frac{4}{3}\pi R_E^3}\,, \tag{2.127}$$

where R_E is the earth's radius. Then the potential V_B, due to B-forces of nucleons, forming the earth, can be calculated as

$$V_B = \frac{g_B^2 M_E N_p}{\frac{4}{3}\pi R_E^3 m_N} \int \frac{d^3\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|} = \frac{g_B^2 M_E N_p}{m_N r}\,, \tag{2.128}$$

where the integration is made over the earth's volume; g_B is the coupling constant of B-forces. It is seen that equation (2.128) is similar to gravitation potential. Thus, the total potential, acting upon our test particle, is equal to

$$V = -G\frac{m_p M_E}{r} + g_B^2 \frac{M_E N_p}{m_N r} = -G\frac{m_p M_E}{r}\left[1 - \frac{g_B^2}{G}\frac{N_p}{m_N m_p}\right]. \tag{2.129}$$

Thus, the presence of B-forces leads to $V \neq \bar{V}$, where $\bar{V}$ is the potential acting upon antiparticle $\bar{p}$ with the opposite baryon charge: $N_{\bar{p}} = -N_p$. In principle, this effect can be observable in the case of

$$\frac{g_B^2}{m_N^2} \sim G\,. \tag{2.130}$$

However, it is unobservable experimentally: particles and antiparticles fall in the gravitational field of the earth with the same acceleration (with rather high accuracy). This fact leads to an estimate of $g_B^2 < 10^{-38}$ as $G m_N^2 \sim 10^{-38}$. Even such a small value of g_B can be excluded! The equation of motion of a test particle in the gravitational field can be written as

$$m_p g = -G\frac{m_p M_E}{r^2} \tag{2.131}$$

and mass m_p is canceled here, so that free-fall acceleration g does not depend on it (the equivalence of inertional and gravitational masses!). If we neglect the mass of the electrons (compared to nucleons), we have

$$m_p = m_N N_p - \epsilon\,, \tag{2.132}$$

where ϵ is the coupling energy in the nuclei of the substance of our test particle. Then

$$N_p = \frac{m_p}{m_N} + \frac{\epsilon}{m_N}\,. \tag{2.133}$$

In the presence of B-forces, the Newtonian equation of motion takes the form

$$m_p g = -\frac{m_p M_E}{r^2} C + \frac{g_B^2}{r^2} \frac{M_E \epsilon}{m_N^2}, \tag{2.134}$$

where $C = G - g_B^2/m_N^2$ can be identified with the measured gravitational constant G_{exp}. In other words, equation (2.134) can be rewritten as

$$m_p g = -\frac{m_p M_E}{r^2} G_{exp} \left[1 - \frac{g_B^2}{G_{exp} m_N^2} \frac{\epsilon}{m_p} \right]. \tag{2.135}$$

The second term here breaks the equivalence of inertional and gravitational masses, which is experimentally established with an accuracy of the order of 10^{-8} in classical Eotvos experiments for different substances. A typical modern estimate from similar experiments gives

$$\frac{g_B^2}{G m_N^2} \frac{\epsilon}{m_p} \sim 10^{-3} \frac{g_B^2}{G m_N^2} < 10^{-12}, \tag{2.136}$$

where we have taken into account that the difference of ϵ/m_p for different substances is of the order of 10^{-3}. Accordingly,

$$\frac{g_B^2}{G m_N^2} < 10^{-9}. \tag{2.137}$$

Thus, the experimentally established equivalence of inertional and gravitational masses leads to the following upper bound of B-forces coupling constant: $g_B^2 < 10^{-47}$! Accordingly, B-forces (if they exist at all) are much weaker even than gravitation! Thus, in every practical sense, we can exclude the existence of any massless gauge fields except the electromagnetic field. Experimentally observed, vector mesons are massive and break the local gauge invariance. Thus, it seems that the beautiful idea of the introduction of new gauge fields becomes rather doubtful. Later we shall see how this problem is solved in modern particle theory.

2.6 The geometry of gauge fields

Let us make some generalizations. We have seen above that the rotation of the *vector* in isotopic space on some small angle $\vec{\Lambda}$ ($|\vec{\Lambda}| \ll 1$) can be written as (cf. (2.99d))

$$\vec{\varphi} \to \vec{\varphi}' = \vec{\varphi} - \vec{\Lambda} \times \vec{\varphi}, \tag{2.138}$$

which is an infinitesimal version of the general transformation

$$\vec{\varphi} \to \vec{\varphi}' = \exp(i \vec{I} \cdot \vec{\Lambda}) \vec{\varphi}, \tag{2.139}$$

where $\vec{I}$ are matrix *generators*:

$$I_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad I_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.140}$$

Here the matrix elements can be written as

$$(I_i)_{mn} = -i \varepsilon_{imn}, \tag{2.141}$$

where ε_{imn} is the antisymmetric Levi–Civita symbol. Accordingly, by components, equation (2.138) can be written as

$$\varphi'_m = (1 + i I_i \Lambda_i)_{mn} \varphi_n = (\delta_{mn} + \varepsilon_{imn}\Lambda_i)\varphi_n = \varphi_m - \varepsilon_{min}\Lambda_i\varphi_n = (\vec{\varphi} - \vec{\Lambda} \times \vec{\varphi})_m \,. \tag{2.142}$$

Local transformations have the form

$$\vec{\varphi} \to \vec{\varphi}' = \exp\left(i\vec{I} \cdot \vec{\Lambda}(x)\right)\vec{\varphi} = S(x)\vec{\varphi} \,, \tag{2.143}$$

where $S(x)$ denotes the operator of local rotation. Matrices I are generators of the vector representation of rotation group $O(3)$ (or $SU(2)$) and satisfy the well-known (angular momentum) commutation relations:

$$[I_i, I_j] = i\varepsilon_{ijk} I_k = C_{ijk} I_k \,. \tag{2.144}$$

Here C_{ijk} denote *structural constants* of the $SU(2)$ group, in this case $C_{ijk} = i\varepsilon_{ijk}$. Naturally, structural constants for other Lie groups are different, but commutation relations for generators are always written as in equation (2.144).

For an arbitrary Lie group, generators satisfy the Jacobi identity

$$[[I_i, I_j], I_k] + [[I_j, I_k], I_i] + [[I_k, I_i], I_j] = 0 \,, \tag{2.145}$$

which reduces (for structural constants) to

$$C_{ijl}C_{lkm} + C_{jkl}C_{lim} + C_{kil}C_{ljm} = 0 \,. \tag{2.146}$$

So far we have analyzed the *isovector* field. A more fundamental approach requires introduction of *isospinors* for the same $SU(2)$ group[7]. Rotation of fundamental two-dimensional spinor $\psi = \binom{\psi_1}{\psi_2}$ can be written as

$$\psi' = \exp\left[\frac{i}{2}\vec{\tau} \cdot \vec{\Lambda}(x)\right]\psi(x) = S(x)\psi(x) \,, \tag{2.147}$$

where $S(x)$ is 2×2 matrix and $\vec{\tau}$ are Pauli matrices in isotopic space, $\tau_i/2$ satisfy commutation relations (2.144), and from the beginning we are writing the local transformation. For the general n-dimensional case we have

$$\psi(x) \to \psi'(x) = \exp[i M^a \Lambda^a(x)]\psi(x) = S(x)\psi(x) \,, \tag{2.148}$$

where a takes the values $1, 2, 3$ ($SU(2)$ group!); here ψ is the n-component spinor, and M^a are $n \times n$ matrices, satisfying commutation relations like (2.144).

[7] Below we shall return to the detailed analysis of spinors; here it is sufficient to remember some elementary information from the standard course on quantum mechanics.

If we consider local transformations of fields, the field derivative $\partial_\mu \psi$, as we have seen above, is transformed in a noncovariant way:

$$\partial_\mu \psi' = S(\partial_\mu \psi) + (\partial_\mu S)\psi \,. \tag{2.149}$$

The reason for this is purely of a "geometrical" nature. The fields $\psi(x)$ and $\psi(x + dx) = \psi(x) + d\psi$ in nearby (infinitesimally close) points of space-time are measured relative to different (rotated by local gauge transformations) axes in isotopic space, shown in Figure 2.2(a). Thus, the value of $d\psi$ contains information not only on field change with distance, as we move from point x to $x + dx$, but also on the appropriate change due to the rotation of the axes in isotopic fields. To construct covariant derivative, we have to compare $\psi(x + dx)$ not with $\psi(x)$, but with the value which the field $\psi(x)$ acquires due to translation from x to $x + dx$ with fixed directions of axes in isotopic space, which is denoted below as $\psi + \delta\psi$ and called the field obtained as a result of "parallel" transfer, as shown in Figure 2.2(b). Let us assume that $\delta\psi$ is proportional to field ψ itself and also to translation dx_μ, so that it can be written as

$$\delta\psi = igM^a A^a_\mu dx^\mu \psi \,, \tag{2.150}$$

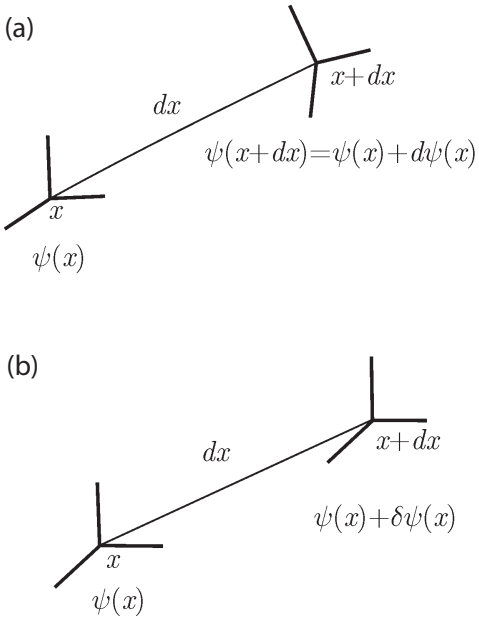


Figure 2.2. (a) The value of $d\psi$ contains information both on the change of ψ and on the transformation of coordinate axes in isotopic space during the transfer from point x to $x + dx$. (b) The value of $\delta\psi$, determined by "parallel" transfer.

where g is some constant and A^a_μ is the gauge field, which in some sense determines how axes in isotopic space change during the transfer from one point to the other. The "true" or covariant derivative of ψ is now determined by the difference

$$D\psi = (\psi + d\psi) - (\psi + \delta\psi) = d\psi - \delta\psi = d\psi - igM^a A^a_\mu dx^\mu \psi \qquad (2.151)$$

and equal to

$$\frac{D\psi}{dx^\mu} = D_\mu \psi = \partial_\mu \psi - igM^a A^a_\mu \psi \,. \qquad (2.152)$$

The situation here is similar to that in the theory of gravitation [33], where the covariant derivative of some vector V^μ is defined as

$$D_\nu V^\mu = \partial_\nu V^\mu + \Gamma^\mu_{\lambda\nu} V^\lambda \,, \qquad (2.153)$$

where Christoffel coefficients $\Gamma^\mu_{\lambda\nu}$ connect the components of the vector in a given point with its components in a nearby point, from which this vector is transferred by parallel translation in Riemann space.

Equation (2.152) gives the general definition of the covariant derivative in the Yang–Mills theory for an arbitrary field ψ, which is transformed under some irreducible representation of some gauge group with generators M^a [56]. Consider the following simple examples:

- $U(1)$ *group*:

$$\varphi \to e^{-i\Lambda}\varphi \,, \qquad \varphi^* \to e^{i\Lambda}\varphi^* \,, \qquad M = -1 \,, \qquad D_\mu = \partial_\mu + igA_\mu \qquad g = e \qquad (2.154)$$

  – electrodynamics.

- $SU(2)$ *group*:

  vector representation:

$$\begin{aligned} (M^a)_{mn} &= -i\varepsilon_{amn} \,, \quad (a,m,n = 1,2,3) \,, \\ D_\mu \varphi_m &= \partial_\mu \varphi_m - ig(M^a)_{mn} A^a_\mu \varphi_n = \partial_\mu \varphi_m - g\varepsilon_{amn} A^a_\mu \varphi_n \\ &= (\partial_\mu \vec{\varphi} + g\vec{A}_\mu \times \vec{\varphi})_m \,, \end{aligned} \qquad (2.155)$$

  where $\vec{A}$ is the same gauge field, which was denoted as $\vec{W}$ above.

  spinor representation:

$$\begin{aligned} M^a &= \frac{1}{2}\tau^a \,, \quad (a = 1,2,3) \,, \\ D_\mu \psi &= \partial_\mu \psi - i\frac{g}{2}\vec{\tau}\cdot\vec{A}_\mu \psi \end{aligned} \qquad (2.156)$$

  – Yang–Mills theory.

Thus, under an arbitrary rotation in isotopic space the field is transformed as

$$\psi \to S(x_\mu)\psi\,, \tag{2.157}$$

and the covariant derivative is transformed as the field itself:

$$D_\mu\psi \to D'_\mu\psi' = S(x_\mu)D_\mu\psi\,. \tag{2.158}$$

It is convenient to introduce matrix notations

$$\hat{A}_\mu = M^a A^a_\mu\,, \tag{2.159}$$

so that equation (2.152) takes the form

$$D_\mu\psi = (\partial_\mu - ig\hat{A}_\mu)\psi\,. \tag{2.160}$$

Transformation to a new coordinate system in isotopic space, with the account of (2.158), gives

$$(\partial_\mu - ig\hat{A}'_\mu)\psi' = S(\partial_\mu - ig\hat{A}_\mu)\psi\,. \tag{2.161}$$

Assuming here $\psi' = S\psi$, we obtain

$$\hat{A}'_\mu = S\hat{A}_\mu S^{-1} - \frac{i}{g}(\partial_\mu S)S^{-1}\,, \tag{2.162}$$

which gives the general form of the gauge transformations of Yang–Mills fields (generalized gradient transformation). Consider again the same examples:

- $U(1)$ *group*:

$$S = e^{-i\Lambda}\,,\quad \partial_\mu S = -i(\partial_\mu\Lambda)e^{-i\Lambda}\,,$$
$$A'_\mu = A_\mu + \frac{1}{e}\partial_\mu\Lambda \quad (g = e,\ M = -1)\,,\quad D_\mu = \partial_\mu + igA_\mu\,. \tag{2.163}$$

- $SU(2)$ *group*:

  spinor representation:

$$S = \exp\left(\frac{i}{2}\vec{\tau}\cdot\vec{\Lambda}\right)\,,\qquad \partial_\mu S = \frac{i}{2}\vec{\tau}\cdot\partial_\mu\vec{\Lambda}S\,, \tag{2.164}$$

$$\vec{A}'_\mu = \vec{A}_\mu - \vec{\Lambda}\times\vec{A}_\mu + \frac{1}{g}\partial_\mu\vec{\Lambda}\,, \tag{2.165}$$

  which follows from (2.162) for $|\vec{\Lambda}| \ll 1$, with the account of commutation relations $[\tau_a, \tau_b] = i2\varepsilon_{abc}\tau_c$, and coincides with (2.103).

Consider now the succession of "parallel transfers" of our field around the closed path $ABCD$, shown in Figure 2.3. Let us start from point A, where the field is assumed to

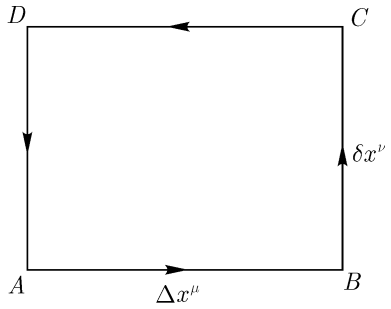


Figure 2.3

be equal to $\psi_{A,0}$. Then its change due to the transfer to point B is determined by the covariant derivative (cf. (2.158), (2.152)), which gives

$$\psi_B = \psi_{A,0} + D_\mu \psi_{A,0} \Delta x^\mu + \frac{1}{2} D_\mu D_\nu \psi_{A,0} \Delta x^\mu \Delta x^\nu + \cdots = (1 + \Delta x^\mu D_\mu + \cdots)\psi_{A,0} .$$

Next, performing transfer to point C, up to the terms of first order, we get

$$\psi_C = \psi_B + \delta x^\nu D_\nu \psi_B = (1 + \delta x^\nu D_\nu)\psi_B = (1 + \delta x^\nu D_\nu)(1 + \Delta x^\mu D_\mu)\psi_{A,0} .$$

Next, the transfer to point D and the return to initial point A give

$$\psi_D = (1 - \Delta x^\rho D_\rho)\psi_C , \tag{2.166}$$

$$\begin{aligned} \psi_{A,1} &= (1 - \delta x^\sigma D_\sigma)\psi_D \\ &= (1 - \delta x^\sigma D_\sigma)(1 - \Delta x^\rho D_\rho)(1 + \delta x^\nu D_\nu)(1 + \Delta x^\mu D_\mu)\psi_{A,0} \qquad (2.167) \\ &= \left\{1 + \delta x^\mu \Delta x^\nu [D_\mu, D_\nu]\right\} \psi_{A,0} , \qquad (2.168) \end{aligned}$$

where the commutator of operators of covariant differentiation appeared:

$$[D_\mu, D_\nu] = [\partial_\mu - ig\hat{A}_\mu, \partial_\nu - ig\hat{A}_\nu] = -ig\left\{\partial_\mu \hat{A}_\nu - \partial_\nu \hat{A}_\mu - ig[\hat{A}_\mu, \hat{A}_\nu]\right\} . \tag{2.169}$$

Let us introduce the field tensor

$$G_{\mu\nu} = \partial_\mu \hat{A}_\nu - \partial_\nu \hat{A}_\mu - ig[\hat{A}_\mu, \hat{A}_\nu] , \tag{2.170}$$

so that

$$[D_\mu, D_\nu] = -igG_{\mu,\nu} . \tag{2.171}$$

Equation (2.170), in fact, gives the general definition of the tensor of Yang–Mills fields for an arbitrary gauge group. Accordingly, equation (2.168) can be written as

$$\psi_{A,1} = (1 - ig\Delta S^{\mu\nu} G_{\mu\nu})\psi_{A,0} , \quad \Delta S^{\mu\nu} = \delta x^\mu \Delta x^\nu \tag{2.172}$$

and we obtain

$$\psi_{A,1} - \psi_{A,0} = -i\,g\Delta S^{\mu\nu} G_{\mu\nu}\psi_{A,0}\,. \tag{2.173}$$

Thus, *a nonzero gauge field tensor leads to a finite physical result as we go around the closed path, which is proportional to the flux of the gauge field $G_{\mu\nu}$ through the path (contour) area $\Delta S^{\mu\nu}$*: the field ψ is rotated in isotopic space. It is easy to see that the field tensor $G_{\mu\nu}$ is invariant relative to gauge transformations:

$$G_{\mu\nu} = S G_{\mu\nu} S^{-1} \tag{2.174}$$

so that it cannot be reduced to zero, using only such transformations. At the same time, if $G_{\mu\nu}$ is zero for some gauge, it remains zero for all other gauges.

Consider again our examples:

- $U(1)$ *group*:

$$G_{\mu\nu} \equiv F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{2.175}$$

  – the usual field tensor of electrodynamics.

- $SU(2)$ *group*:

$$[M^a, M^b] = i\,\varepsilon_{abc} M^c\,, \qquad G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g\varepsilon_{abc} A^b_\mu A^c_\nu\,, \tag{2.176}$$

  which in vector notations in isotopic space

$$\vec{G}_{\mu\nu} = \partial_\mu \vec{A}_\nu - \partial_\nu \vec{A}_\mu + g\vec{A}_\mu \times \vec{A}_\nu \tag{2.177}$$

  coincides with definition (2.112) given above.

Here we again can note an analogy to the theory of gravitation. The tensor of Yang–Mills fields is, in some sense, analogous to the Riemann–Christoffel curvature tensor [33]:

$$R^\kappa_{\lambda\mu\nu} = \partial_\nu \Gamma^\kappa_{\lambda\mu} - \partial_\mu \Gamma^\kappa_{\lambda\nu} + \Gamma^\rho_{\lambda\mu}\Gamma^\kappa_{\rho\nu} - \Gamma^\rho_{\lambda\nu}\Gamma^\kappa_{\rho\mu}\,. \tag{2.178}$$

The parallel transfer of an arbitrary vector V^μ around the closed contour in Riemann space leads to the following difference between the initial and final components of the vector:

$$\Delta V^\mu = \frac{1}{2} R^\mu_{\rho\sigma\lambda} V^\rho \Delta S^{\sigma\lambda}\,, \tag{2.179}$$

where $\Delta S^{\sigma\lambda}$ again denotes the area of the contour. The value of ΔV^μ is different from zero only in the space of with finite curvature. In general relativity, this corresponds to the presence of a nontrivial gravitational field.

Analyzing the transfer around the path forming the parallelepiped shown in Figure 2.4, Feynman has given a simple derivation of the following identity for the field $G_{\mu\nu}$:

$$D_\rho G_{\mu\nu} + D_\mu G_{\nu\rho} + D_\nu G_{\rho\mu} = 0\,, \tag{2.180}$$

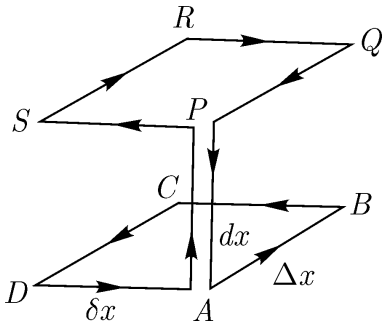


Figure 2.4

which in fact determines the second pair of "Maxwell equations" for the Yang–Mills field (2.121). In the case of $U(1)$ gauge symmetry, this reduces simply to (2.119):

$$\partial_\rho F_{\mu\nu} + \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} = 0\,. \tag{2.181}$$

Briefly, the derivation goes as follows. In Figure 2.4 we show the path (contour) $ABCDAPSRQPA$. There are another two paths of the same type along the borders of two pairs of opposite facets of the parallelepiped, so that along the borders of all six facets we can draw the path $(ABCDAPSRQPA) + (ADSPABQRCBA) + (APQBADCRSDA)$. All parts of this contour are now passed twice in two opposite directions. Accordingly, the field ψ is *not changed* as we go around our closed path, which immediately gives the identity (2.180).

In the theory of gravitation there exists the similar Bianchi identity for the Riemann–Christoffel tensor:

$$D_\rho R^\kappa_{\lambda\mu\nu} + D_\mu R^\kappa_{\lambda\nu\rho} + D_\nu R^\kappa_{\lambda\rho\mu} = 0\,. \tag{2.182}$$

The analogy of gauge field theories and the theory of gravitation can be expressed as in Table 2.1.

Table 2.1. Analogies between gauge field theories and gravitation.

Gauge theories	*General relativity*
Gauge transformations	Coordinate transformations
Gauge group	Group of all coordinate transformations
Potential of gauge field A_μ	Christoffel coefficients $\Gamma^\kappa_{\mu\nu}$
Field tensor $G_{\mu\nu}$	Tensor of curvature $R^\kappa_{\lambda\mu\nu}$

All these analogies actually exist on a deeper level. Even during early stages of the development of gauge field theories, it was shown by Utiyama [68] that the equations of Einstein's general relativity theory can be derived using the idea and general scheme of gauge field theory, if we take the Lorentz group (coordinate transformations of the special theory of relativity)

and demand the invariance of the theory with respect to corresponding local transformations (when parameters of the Lorentz group are considered as arbitrary functions of coordinate in Minkowski space).

2.7 A realistic example – chromodynamics

Let us briefly consider the structure of quantum chromodynamics (QCD) as an example of realistic non-Abelian gauge theory. Quantum chromodynamics is based on the fundamental experimental discovery: each quark of the given "flavor" u, d, s, c, t, b possesses an additional quantum number, which is called "color", and which can take three possible values ($1, 2, 3$ or R, G, B)[8]. Then, each quark field is represented by the fundamental spinor of the $SU(3)$ group[9]:

$$q = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} . \tag{2.183}$$

The color symmetry is *exact*, and QCD Lagrangian should be invariant to $SU(3)$ group transformations:

$$q \to Uq\,, \tag{2.184}$$

where the 3×3-matrices U are unitary and unimodular:

$$\begin{gathered} U^+U = 1\,, \quad \operatorname{Det} U = 1\,, \\ U = e^{iT}\,, \quad T = T^+\,, \quad \operatorname{Sp} T = 0\,. \end{gathered} \tag{2.185}$$

These matrices (transformations) depend on eight parameters ("rotation angles") ε_a, and accordingly there are eight generators $\lambda_i/2 \quad (i = 1, \ldots, 8)$:

$$\begin{gathered} \lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\ \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \end{gathered} \tag{2.186}$$

[8] The necessity of this quantum number was clear from the very beginning of the quark model, as it allowed to lift certain contradictions with the Pauli principle.

[9] A rather clear and compact presentation of irreducible representations of this group, though in relation to the other problem of particle physics (approximate symmetry of hadrons and their quark structure), can be found in [40].

which are a kind of "generalization" of Pauli matrices to three dimensions. These generators satisfy the following commutation relations:

$$\left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2}\right] = i f_{abc} \frac{\lambda_c}{2}, \tag{2.187}$$

where the nonzero structural constants f_{abc} are given by

$$f_{123} = 1, \quad f_{147} = -f_{156} = f_{246} = f_{257} = f_{345} = -f_{367} = \frac{1}{2},$$
$$f_{458} = f_{678} = \frac{\sqrt{3}}{2}. \tag{2.188}$$

The basic approach of QCD is to make color symmetry the local gauge symmetry!

As a result, using the recipes of gauge field theory, we introduce eight gauge fields (gluons), which transfer interactions between quarks. These can be conveniently written in the following matrix form (as in (2.159)):

$$\hat{A}_\mu = A^a_\mu \frac{\lambda^a}{2} = \frac{1}{2} \begin{pmatrix} A^3_\mu + \frac{1}{\sqrt{3}} A^8_\mu & A^1_\mu - i A^2_\mu & A^4_\mu - i A^5_\mu \\ A^1_\mu + i A^2_\mu & -A^3_\mu + \frac{1}{\sqrt{3}} A^8_\mu & A^6_\mu - i A^7_\mu \\ A^4_\mu + i A^5_\mu & A^6_\mu + i A^7_\mu & -\frac{2}{\sqrt{3}} A^8_\mu \end{pmatrix}. \tag{2.189}$$

The explicit form of the gluon field tensor can be obtained from (2.170) or from (2.176), substituting into the last expression instead of ε_{abc} the structural constants f_{abc} of the $SU(3)$ group. In accordance with the general ideology of gauge theories, gluons are massless. The absence of long-range forces due to gluons is explained by the phenomenon of confinement, which will be discussed in the final part of this book.

Chapter 3

Canonical quantization, symmetries in quantum field theory

3.1 Photons

3.1.1 Quantization of the electromagnetic field

Now we have to move from classical to quantum field theory. The procedure of canonical field quantization is done in complete analogy to similar procedures for mechanical systems. First of all, we shall consider the quantum field theory of free (noninteracting) fields, and we shall start with the case of the electromagnetic field – not the simplest case, but nonetheless physically quite important. We have already seen above that the electromagnetic field is an example of an (Abelian) gauge field. This leads to some additional complications related to the correct account of gauge invariance. For the electromagnetic field these problems are solved in a relatively simple way, within the canonical quantization procedure, while for non-Abelian Yang–Mills fields we need a much a more complicated scheme of quantization, based on functional integration, which will be discussed much later. The presentation in this chapter is essentially based on [6].

From a mechanical point of view, the field is represented as the system with an infinite (continuous) number of degrees of freedom. However, it is convenient to start from the classical description of the field, which deals with an infinite, but discrete, set of variables. We shall consider the electromagnetic field in the so-called Coulomb gauge, when its vector potential $\mathbf{A}(\mathbf{r}, t)$ satisfies the condition of transversality:

$$\operatorname{div} \mathbf{A} = 0 \,. \tag{3.1}$$

The scalar potential is taken as $\varphi = 0$, while the electric $\mathbf{E}$ and magnetic $\mathbf{H}$ field are defined as[1]

$$\mathbf{E} = -\dot{\mathbf{A}} \,, \qquad \mathbf{H} = \operatorname{rot} \mathbf{A} \,. \tag{3.2}$$

Maxwell equations reduce, in this case, to the wave equation for the vector potential $\mathbf{A}$

$$\nabla^2 \mathbf{A} - \frac{\partial^2 \mathbf{A}}{\partial t^2} = 0 \,. \tag{3.3}$$

It is well known that the six components of the electromagnetic field are written in the form of an antisymmetric tensor:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \,, \tag{3.4}$$

[1] Let us recall that we are using the system of units with the speed of light $c = 1$.

which immediately leads to homogeneous Maxwell equations:

$$\partial^\lambda F^{\mu\nu} + \partial^\mu F^{\nu\lambda} + \partial^\nu F^{\lambda\mu} = 0\,. \tag{3.5}$$

In a vacuum (in the absence of sources) inhomogeneous Maxwell equations are written as

$$\partial_\mu F^{\mu\nu} = 0 \tag{3.6}$$

or

$$\Box A^\nu - \partial^\nu(\partial_\mu A^\mu) = 0\,. \tag{3.7}$$

We know that these equations follow from the variational principle with a Lagrangian

$$\mathcal{L} = -\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu}\,, \tag{3.8}$$

where A^μ is considered as a dynamic variable. However, for the given values of field strength $F_{\mu\nu}$, the 4-vector potential A^μ is not single valued, but is determined only up to the gradient transformation

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \Lambda(x)\,. \tag{3.9}$$

If we require for $\Lambda(x)$ the validity of an additional condition $\Box\Lambda = -\partial_\mu A^\mu$, we easily obtain for the field, transformed by (3.9), $\partial_\mu A^{\mu\prime} = 0$. Now we can just drop the prime over A^μ and write the so-called *Lorentz condition*:

$$\partial_\mu A^\mu = 0\,. \tag{3.10}$$

Then (3.7) is transformed to

$$\Box A^\nu = 0 \tag{3.11}$$

i. e., the wave function for the 4-vector potential. The Lorentz gauge (3.10) gives one equation for four components of the potential, reducing the number of independent components of the field to three. However, this condition still does not make A_μ singly defined. For A_μ satisfying Lorentz condition, we may introduce $A'_\mu = A_\mu + \partial_\mu\Lambda$, which also satisfies it due to $\Box\Lambda(x) = 0$. Let us now choose $\Lambda(x)$ to satisfy $\frac{\partial\Lambda}{\partial t} = -\varphi$, then obviously $\varphi' = 0$, so that equation (3.10) gives $\nabla\cdot\mathbf{A} = \operatorname{div}\mathbf{A} = 0$. Thus, we come to the Coulomb gauge, with only two independent components of the electromagnetic field (transversality condition), in agreement with reality.

Transformation to a discrete set of field variables is achieved by considering the field system in a finite spatial volume V (below, for shortness of notation, we just put $V = 1$) [33]. The vector potential is represented by a Fourier expansion over plane waves:

$$\mathbf{A} = \sum_{\mathbf{k}} (\mathbf{a}_\mathbf{k} e^{i\mathbf{kr}} + \mathbf{a}^*_\mathbf{k} e^{-i\mathbf{kr}})\,, \tag{3.12}$$

where expansion coefficients $\mathbf{a}_\mathbf{k}$ depend on time according to

$$\mathbf{a}_\mathbf{k} \sim \mathrm{e}^{-i\omega_\mathbf{k} t}\,, \qquad \omega_\mathbf{k} = |\mathbf{k}|\,. \tag{3.13}$$

Due to transversality condition (3.1), we have

$$\mathbf{a_k} \cdot \mathbf{k} = 0\,. \tag{3.14}$$

In (3.12) summation is done over the infinite discrete set of k_x, k_y, k_z. As usual, we can make transformation from summation to integration over k_x, k_y, k_z, introducing $\frac{d^3\mathbf{k}}{(2\pi)^3}$, as the number of possible values of $\mathbf{k}$, in the infinitesimal volume of $\mathbf{k}$-space $d^3\mathbf{k} = dk_x dk_y dk_z$. Finally, the state of the field is completely determined by amplitudes $\mathbf{a_k}$, which are considered as the set of classical field variables.

Let us introduce canonical field variables as

$$\mathbf{Q_k} = \frac{1}{\sqrt{4\pi}}(\mathbf{a_k} + \mathbf{a_k^*})\,, \tag{3.15}$$

$$\mathbf{P_k} = -\frac{i\omega_\mathbf{k}}{\sqrt{4\pi}}(\mathbf{a_k} - \mathbf{a_k^*}) = \dot{\mathbf{Q}}_\mathbf{k}\,. \tag{3.16}$$

Obviously, these variables are real. Then the series (3.12) can be rewritten as

$$\mathbf{A} = \sqrt{4\pi}\sum_\mathbf{k}\left[\mathbf{Q_k}\cos\mathbf{kr} - \frac{1}{\omega_\mathbf{k}}\mathbf{P_k}\sin\mathbf{kr}\right]. \tag{3.17}$$

To determine the field Hamiltonian H, we calculate the total energy:

$$E = \frac{1}{8\pi}\int d^3\mathbf{r}(\mathbf{E}^2 + \mathbf{H}^2) \tag{3.18}$$

and express it via the variables $\mathbf{Q_k}$ and $\mathbf{P_k}$. To do this, we find $\mathbf{E}$ and $\mathbf{H}$ from (3.2) and (3.17), substitute the appropriate expressions into (3.18) and obtain, after the integration over coordinates,

$$H = \frac{1}{2}\sum_\mathbf{k}\left(\mathbf{P_k^2} + \omega_\mathbf{k}^2\mathbf{Q_k^2}\right). \tag{3.19}$$

From the condition of transversality, both $\mathbf{P_k}$ and $\mathbf{Q_k}$ are orthogonal to vector $\mathbf{k}$, so that they, in fact, possess only two independent components. The directions of these vectors are defined by the polarization directions of the appropriate wave. Let us denote two components of $\mathbf{P_k}$ and $\mathbf{Q_k}$ in the plane orthogonal to $\mathbf{k}$ as $P_{\mathbf{k}\alpha}$ and $Q_{\mathbf{k}\alpha}$, with $\alpha = 1, 2$. Then (3.19) can be rewritten as

$$H = \frac{1}{2}\sum_{\mathbf{k}\alpha}\left(P_{\mathbf{k}\alpha}^2 + \omega_\mathbf{k}^2 Q_{\mathbf{k}\alpha}^2\right). \tag{3.20}$$

Thus, the Hamiltonian H is represented by the sum of independent terms, each having the form of the Hamiltonian of the harmonic oscillator.

Now we can perform *quantization*. The way to quantize an oscillator is well known from quantum mechanics [35]. Quantization reduces to the change of the generalized coordinates $Q_{\mathbf{k}\alpha}$ and generalized momenta $P_{\mathbf{k}\alpha}$ by corresponding *operators*, satisfy-

ing the standard commutation relations[2]:

$$Q_{\mathbf{k}\alpha} P_{\mathbf{k}\alpha} - P_{\mathbf{k}\alpha} Q_{\mathbf{k}\alpha} \equiv [Q_{\mathbf{k}\alpha}, P_{\mathbf{k}\alpha}] = i \,. \tag{3.21}$$

For different values of $\mathbf{k}\alpha$, the corresponding operators just commute. Accordingly, the fields $\mathbf{A}, \mathbf{E}, \mathbf{H}$ also become operators.

The Eigenvalues of the Hamiltonian (3.20) obviously are

$$E = \sum_{\mathbf{k}\alpha} \left(N_{\mathbf{k}\alpha} + \frac{1}{2} \right) \omega_{\mathbf{k}} \,, \tag{3.22}$$

where $N_{\mathbf{k}\alpha}$ are integer numbers, representing the number of *photons* in quantum states, characterized by $\mathbf{k}\alpha$. The matrix elements of operator $Q_{\mathbf{k}\alpha}$ are also well known from quantum mechanics [35]:

$$\langle N_{\mathbf{k}\alpha} | Q_{\mathbf{k}\alpha} | N_{\mathbf{k}\alpha} - 1 \rangle = \langle N_{\mathbf{k}\alpha} - 1 | Q_{\mathbf{k}\alpha} | N_{\mathbf{k}\alpha} \rangle = \sqrt{\frac{N_{\mathbf{k}\alpha}}{2\omega_{\mathbf{k}}}} \,. \tag{3.23}$$

The matrix elements $P_{\mathbf{k}\alpha} = \dot{Q}_{\mathbf{k}\alpha}$ differ from (3.23) by a factor $\pm i\,\omega_{\mathbf{k}}$.

Let us introduce new operators:

$$c_{\mathbf{k}\alpha} = \frac{1}{\sqrt{2\omega_{\mathbf{k}}}} (\omega_{\mathbf{k}} Q_{\mathbf{k}\alpha} + i P_{\mathbf{k}\alpha}) \,, \qquad c^{+}_{\mathbf{k}\alpha} = \frac{1}{\sqrt{2\omega_{\mathbf{k}}}} (\omega_{\mathbf{k}} Q_{\mathbf{k}\alpha} - i P_{\mathbf{k}\alpha}) \,. \tag{3.24}$$

Then, from (3.23) and (3.24) we obtain

$$\langle N_{\mathbf{k}\alpha} - 1 | c_{\mathbf{k}\alpha} | N_{\mathbf{k}\alpha} \rangle = \langle N_{\mathbf{k}\alpha} | c^{+}_{\mathbf{k}\alpha} | N_{\mathbf{k}\alpha} - 1 \rangle = \sqrt{N_{\mathbf{k}\alpha}} \,. \tag{3.25}$$

From (3.24) and (3.21) we immediately get commutation relations for operators $c_{\mathbf{k}\alpha}$ and $c^{+}_{\mathbf{k}\alpha}$:

$$c_{\mathbf{k}\alpha} c^{+}_{\mathbf{k}\alpha} - c^{+}_{\mathbf{k}\alpha} c_{\mathbf{k}\alpha} \equiv [c_{\mathbf{k}\alpha}, c^{+}_{\mathbf{k}\alpha}] = 1 \,. \tag{3.26}$$

For different $\mathbf{k}$ and α these operators simply commute. The operators $c_{\mathbf{k}\alpha}$ and $c^{+}_{\mathbf{k}\alpha}$ are called operators of annihilation and creation of photons in the state with wave vector (momentum) $\mathbf{k}$ and polarization α. The origin of these terms is obvious from (3.25). For historical reasons, the formalism, based on the use of such operators is called *second quantization*.

The *operator* of the vector potential (with the use of (3.12), (3.15), (3.16), and (3.24)) can now be written as

$$\mathbf{A} = \sum_{\mathbf{k}\alpha} \left(c_{\mathbf{k}\alpha} \mathbf{A}_{\mathbf{k}\alpha} + c^{+}_{\mathbf{k}\alpha} \mathbf{A}^{*}_{\mathbf{k}\alpha} \right) , \tag{3.27}$$

where

$$\mathbf{A}_{\mathbf{k}\alpha} = \sqrt{4\pi} \frac{\mathbf{e}^{(\alpha)}}{\sqrt{2\omega_{\mathbf{k}}}} e^{i\mathbf{k}\mathbf{r}} \,, \tag{3.28}$$

[2] Note that here we use $\hbar = 1$.

where $\mathbf{e}^{(\alpha)}$ is the unit vector of polarization for the given field oscillator. Obviously, we have $\mathbf{e}^{(\alpha)} \cdot \mathbf{k} = 0$, so that this vector is orthogonal to the photon momentum $\mathbf{k}$. To each value of $\mathbf{k}$ we have two independent directions of polarization $\alpha = 1, 2$.

Similarly we can write down expansions for electric $\mathbf{E}$ and magnetic $\mathbf{H}$ field operators:

$$\mathbf{E} = \sum_{\mathbf{k}\alpha} (c_{\mathbf{k}\alpha} \mathbf{E}_{\mathbf{k}\alpha} + c^+_{\mathbf{k}\alpha} \mathbf{E}^*_{\mathbf{k}\alpha}) , \tag{3.29}$$

$$\mathbf{H} = \sum_{\mathbf{k}\alpha} (c_{\mathbf{k}\alpha} \mathbf{H}_{\mathbf{k}\alpha} + c^+_{\mathbf{k}\alpha} \mathbf{H}^*_{\mathbf{k}\alpha}) , \tag{3.30}$$

where

$$\mathbf{E}_{\mathbf{k}\alpha} = i\omega_{\mathbf{k}} \mathbf{A}_{\mathbf{k}\alpha} , \quad \mathbf{H}_{\mathbf{k}\alpha} = [\mathbf{n} \times \mathbf{E}_{\mathbf{k}\alpha}] , \tag{3.31}$$

where $\mathbf{n} = \mathbf{k}/\omega_{\mathbf{k}}$ is the unit vector directed along photon propagation. The vectors $\mathbf{A}_{\mathbf{k}\alpha}$, introduced in (3.28), satisfy the following orthonormality condition:

$$\int d^3\mathbf{r} \mathbf{A}_{\mathbf{k}\alpha} \mathbf{A}^*_{\mathbf{k}'\alpha'} = \frac{2\pi}{\omega_{\mathbf{k}}} \delta_{\alpha\alpha'} \delta_{\mathbf{k}\mathbf{k}'} , \tag{3.32}$$

where we have taken into account that two independent polarization vectors are orthogonal: $\mathbf{e}^{(\alpha)} \cdot \mathbf{e}^{(\alpha')*} = 0$. In fact, the values of $\mathbf{A}_{\mathbf{k}\alpha}$ (plane waves) can be treated as wave functions of a photon with momentum $\mathbf{k}$ and polarization $\mathbf{e}^{(\alpha)}$[3].

From (3.32) and (3.31) it is easy to obtain

$$\frac{1}{4\pi} \int d^3\mathbf{r} (\mathbf{E}_{\mathbf{k}\alpha} \mathbf{E}^*_{\mathbf{k}'\alpha'} + \mathbf{H}_{\mathbf{k}\alpha} \mathbf{H}^*_{\mathbf{k}'\alpha'}) = \omega_{\mathbf{k}} \delta_{\mathbf{k}\mathbf{k}'} \delta_{\alpha\alpha'} . \tag{3.33}$$

Substituting (3.29) and (3.30) into (3.18) and using (3.33), we find

$$\begin{aligned} H &= \sum_{\mathbf{k}\alpha} \frac{1}{2} (c_{\mathbf{k}\alpha} c^+_{\mathbf{k}\alpha} + c^+_{\mathbf{k}\alpha} c_{\mathbf{k}\alpha}) \frac{1}{4\pi} \int d^3\mathbf{r} (\mathbf{E}_{\mathbf{k}\alpha} \mathbf{E}^*_{\mathbf{k}'\alpha'} + \mathbf{H}_{\mathbf{k}\alpha} \mathbf{H}^*_{\mathbf{k}'\alpha'}) \\ &= \sum_{\mathbf{k}\alpha} \frac{1}{2} (c_{\mathbf{k}\alpha} c^+_{\mathbf{k}\alpha} + c^+_{\mathbf{k}\alpha} c_{\mathbf{k}\alpha}) \omega_{\mathbf{k}} \end{aligned} \tag{3.34}$$

or, using commutation relations (3.26),

$$H = \sum_{\mathbf{k}\alpha} \left(c^+_{\mathbf{k}\alpha} c_{\mathbf{k}\alpha} + \frac{1}{2} \right) \omega_{\mathbf{k}} \tag{3.35}$$

which gives the secondary quantized Hamiltonian of the system of photons. After comparison with (3.22) it becomes clear that

$$\hat{N}_{\mathbf{k}\alpha} = c^+_{\mathbf{k}\alpha} c_{\mathbf{k}\alpha} \tag{3.36}$$

[3] Let us stress that these wave functions cannot be understood as probability amplitudes of spatial localization of photon, as there is no sense in defining the coordinate of a particle moving with the velocity of light.

represents the operator of the number of photons in $\mathbf{k}\alpha$ state, which is diagonal in occupation number representation, with integer eigenvalues. Note that (3.33) corresponds to the wave function normalization to a "single photon per volume $V = 1$".

In classical theory of electromagnetic field its momentum is defined as [33]

$$\mathbf{P} = \frac{1}{4\pi} \int d^3\mathbf{r}\, [\mathbf{E} \times \mathbf{H}]\,. \tag{3.37}$$

Replacing $\mathbf{E}$ and $\mathbf{B}$ by operators (3.29) and (3.30), we obtain

$$\mathbf{P} = \sum_{\mathbf{k}\alpha} \left(c^+_{\mathbf{k}\alpha} c_{\mathbf{k}\alpha} + \frac{1}{2} \right) \mathbf{k}\,, \tag{3.38}$$

which corresponds to each photon carrying the momentum $\mathbf{k}$.

The presence in (3.35) and (3.38) of the terms, independent of occupation numbers ($1/2$ contribution in parenthesis), corresponds to an *infinite* contribution of vacuum fluctuations ("zero-level" oscillations) of the electromagnetic field. This is the first example we meet of a typical "field theory divergence". In most cases, in this situation we can simply shift the origin of an energy scale (or the origin of momentum scale) and write

$$H = \sum_{\mathbf{k}\alpha} c^+_{\mathbf{k}\alpha} c_{\mathbf{k}\alpha} \omega_{\mathbf{k}}\,, \qquad \mathbf{P} = \sum_{\mathbf{k}\alpha} c^+_{\mathbf{k}\alpha} c_{\mathbf{k}\alpha} \mathbf{k}\,. \tag{3.39}$$

The origin of energy or momentum scale is "renormalized" here by infinite ("vacuum") constants, which are independent of excitations of the field system. However, we must stress that the presence of an infinite energy (momentum) of the vacuum (zero-level oscillations) is absolutely real physically and reflects the quantum nature of the field, leading to some *finite* experimental effects. One of the best examples is the so-called Casimir effect, which we shall discuss below.

3.1.2 Remarks on gauge invariance and Bose statistics

The choice of potentials in electrodynamics, as is well known, is not unique. Above we have used the Coulomb gauge (3.1). In the general case, components of vector potential A_μ can undergo the gradient transformation like

$$A_\mu \to A_\mu + \partial_\mu \Lambda\,. \tag{3.40}$$

For plane waves, limiting ourselves to transformations, which do not change this form of potential (i. e., its proportionality to $\exp(-ik^\mu x_\mu)$), this nonuniqueness reduces to the possibility of adding to the wave amplitude an arbitrary 4-vector proportional to k^μ.

In the case of an arbitrary gauge, the 4-potential of the field can be written in the form, generalizing (3.27),

$$A^\mu = \sum_{k\alpha} (c_{k\alpha} A^\mu_{k\alpha} + c^+_{k\alpha} A^{\mu *}_{k\alpha})\,, \tag{3.41}$$

where wave functions of photon are

$$A_k^\mu = \sqrt{4\pi}\frac{e^\mu}{\sqrt{2\omega}}e^{-ik_\nu x^\nu}\,, \tag{3.42}$$

where e^μ is a space-like 4-vector of polarization, which satisfies the condition $e_\mu e^{\mu *} = -1$. The space-like nature of the 4-vector of polarization is obvious from the condition of four-dimensional transversality, as the wave vector (momentum) of a real photon always belongs to the light cone. In these notations our gradient (gauge) transformation reduces to

$$e_\mu \to e_\mu + \Lambda k_\mu\,, \tag{3.43}$$

where $\Lambda = \Lambda(k^\mu)$ is an arbitrary scalar function of k^μ. Transversality of polarization means that we always can choose the gauge, guaranteeing three-dimensional transversality, when we choose

$$e^\mu = (0, \mathbf{e})\,, \qquad \mathbf{e}\cdot\mathbf{k} = 0\,. \tag{3.44}$$

Four-dimensional transversality, equivalent to the Lorentz condition (3.10), can be written in an invariant form as

$$e_\mu k^\mu = 0\,. \tag{3.45}$$

This condition, as well as $e_\mu e^{\mu *} = -1$, is not violated by transformation (3.43), as for a real photon we always have $k^2 = 0$ (massless photon on the light-cone!). The measurable physical characteristics should obviously be invariant to gauge transformations.

Photons are described by Bose statistics. This is obvious from the fact that the number of photons $N_{\mathbf{k}\alpha}$ in $\mathbf{k}\alpha$ state may be an arbitrary integer, as well as from the form of commutation relations (3.26). A Bose field can acquire the classical limit. It is well known that the properties of the quantum system approach that of the classical, when quantum numbers, determining the system state, become large. For an electromagnetic field this means that the number of photons $N_{\mathbf{k}\alpha}$ is to be large enough. In this case we can neglect unity in the right-hand side of commutation relations (3.26) (obviously, this corresponds to the limit of $\hbar \to 0$ for the usual system of units) and write

$$c^+_{\mathbf{k}\alpha}c_{\mathbf{k}\alpha} \approx c_{\mathbf{k}\alpha}c^+_{\mathbf{k}\alpha}\,, \tag{3.46}$$

so that operators $c_{\mathbf{k}\alpha}$, $c^+_{\mathbf{k}\alpha}$ can be considered as classical filed amplitudes. However, some care is needed, as in the case of all $N_{\mathbf{k}\alpha} \gg 1$ we shall get the infinity after the summation over $\mathbf{k}\alpha$ for the field energy (3.22).

In fact, from a physical point of view it is sensible to consider the values of the fields, averaged over some finite time intervals Δt. In Fourier expansion of a such averaged field $\mathbf{E}$ the main contribution comes from the frequency region $\omega\Delta t < 1$. Now, to derive the conditions of quasiclassicality we have to consider only field oscillators with $\omega < 1/\Delta t$. The number of oscillators with frequencies from zero to $\omega \sim 1/\Delta t$,

by the order of magnitude, is equal to ($V = 1$):

$$\left(\frac{\omega}{c}\right)^3 \sim \frac{1}{(c\Delta t)^3}\,. \tag{3.47}$$

The energy of the field in the unit volume is of the order of $\mathbf{E}^2$. Dividing this energy by the number of oscillators and by the average photon energy $\sim \hbar\omega$, we get the following estimate for the number of photons:

$$N \sim \frac{\mathbf{E}^2 c^3}{\hbar\omega^4}\,. \tag{3.48}$$

Then from the condition $N \gg 1$ and (3.47) we obtain

$$|\mathbf{E}| \gg \frac{\sqrt{\hbar c}}{(c\Delta t)^2}\,, \tag{3.49}$$

which determines the criterion of quasiclassicality[4]. We see that the field is be strong enough, and stronger for smaller time intervals Δt. For the time-dependent field $\Delta t \sim \omega^{-1}$, so that a sufficiently weak alternating field cannot be described quasiclassically. Only static fields, for which $\Delta t \to \infty$, can always be treated as classical.

On the measurability of fields in quantum electrodynamics

The existence of a finite limit for velocity of propagation of interactions (speed of light) in relativistic theory leads to a number of additional limitations for the measurability of physical characteristics (variables). At the early stages of the development of quantum field theory these limitations were discussed by Landau and Peierls. The qualitative discussion of these limitations can be found in the Introduction to [6]. During this analysis, Landau and Peierls formulated the fundamental question of the possibility of measuring an electromagnetic field itself. They claimed that the measurement of any component of (say) an electric field requires the determination of the momentum of a charged test particle, so that the imminent action of the field, radiated during this operation, will always lead to unavoidable limitations of field measurements. They concluded that the precise measurement of field strength becomes impossible, in contradiction with the basic points of quantum electrodynamics discussed above. This fact, as well as a number of similar difficulties to be discussed later, were the reason for a long period of Landau's rather skeptical opinion on quantum field theory in general.

The problem of fields measurability was analyzed in more detail by Bohr and Rosenfeld (cf. an interesting review of this problem by Rosenfeld in [49]). It was demonstrated, that all the difficulties are essentially solved (in the spirit of the Copenhagen interpretation of quantum mechanics) if we use the finite (not point-like) test particles. For example, consider the measurement of the E_x component of an electric field, averaged over some volume and time intervals. Let us use the test particle with volume V and homogeneous charge density ρ and measure its momenta p'_x and p''_x at the beginning and end of time interval T. Making this test particle heavy enough, we can achieve its arbitrarily small displacement during this interval, and obtain for the average value of the field $\bar{E}_x$

$$\bar{E}_x \rho V T = p''_x - p'_x\,. \tag{3.50}$$

[4] For better understanding here we explicitly write down both c and $\hbar$.

However, the measurement of the momentum of the test particle inevitably leads to some error Δx in the determination of its position, according to the usual indeterminacy relation: $\Delta p_x \sim \hbar/\Delta x$. This leads to indeterminacy $\Delta \bar{E}_x$ for the field value $\bar{E}_x$, which is of the order of

$$\Delta \bar{E}_x \sim \frac{\hbar}{\rho V T \Delta x} . \tag{3.51}$$

However, it is obvious that this error can be made arbitrarily small by just increasing the charge density of a test particle.

In a similar way, we can analyze the measurability of charges and currents [49]. In the opinion of Bohr and Rosenfeld, such arguments demonstrate the absence of any contradictions in the basic principle of quantum electrodynamics. However, we should note that the Copenhagen interpretation of quantum theory, using the classical concepts as its inevitable part, at present is not commonly accepted (nor is it considered to be absolutely satisfactory by many researchers). The modern situation with the quantum limitations of field measurements is discussed in [44].

3.1.3 Vacuum fluctuations and Casimir effect

The reality of vacuum ("zero-level") fluctuations of an electromagnetic field is beautifully illustrated by the so-called Casimir effect [28]. Consider two big ideally conducting metallic planes, placed in a *vacuum*, at the distance a from each other, as it is shown in Figure 3.1. Let these metallic plates be just squares with sides L and $L \gg a$. Consider the modes of an oscillating electromagnetic field in the volume $L^2 a$. Boundary conditions require that the

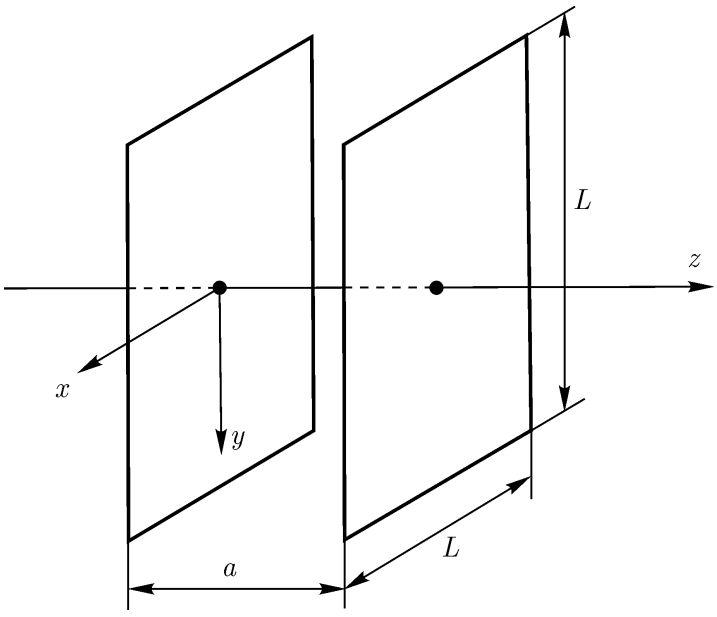


Figure 3.1

vector of electric field **E** be perpendicular, while the vector of magnetic field **B** is parallel to the internal surface of the plate. Only transversal modes contribute to energy. If the wave vector component k_z, orthogonal to the surface of plates, is nonzero, it can acquire only discrete values $k_z = n\pi/a \quad (n = 1, 2, \ldots)$, so that the nodes of the field are at the plates. We also have to take into account two polarization states. If $k_z = 0$, we remain with only one mode (the electric field component of this mode is just zero, as a tangential electric field is absent on the surface of an ideal conductor). Then, the energy of zero-level oscillations of electromagnetic field in the volume between plates is given by

$$E = \sum_{\mathbf{k}\alpha} \frac{1}{2}\hbar\omega_{\mathbf{k}\alpha} = \sum_{\mathbf{k}\alpha} \frac{1}{2}\hbar c|\mathbf{k}_\alpha| = \frac{\hbar c}{2} L^2 \int \frac{d^2\mathbf{k}_\parallel}{(2\pi)^2} \left[|\mathbf{k}_\parallel| + 2\sum_{n=1}^{\infty} \sqrt{\mathbf{k}_\parallel^2 + \frac{n^2\pi^2}{a^2}} \right]. \quad (3.52)$$

This expression is obviously infinite. However, let us subtract from (3.52) the similar expression for the energy of vacuum fluctuations in the same volume, but in the absence of metallic plates:

$$E_0 = \frac{\hbar c}{2} L^2 \int \frac{d^2\mathbf{k}_\parallel}{(2\pi)^2} a \int_{-\infty}^{\infty} \frac{dk_z}{2\pi} 2\sqrt{\mathbf{k}_\parallel^2 + k_z^2} = \frac{\hbar c}{2} L^2 \int \frac{d^2\mathbf{k}_\parallel}{(2\pi)^2} 2 \int_0^{\infty} dn \sqrt{\mathbf{k}_\parallel^2 + n^2\pi^2/a^2}. \quad (3.53)$$

Then, the change of the vacuum energy due to introduction of metallic plates (per unit surface of the plates) is given by

$$\mathcal{E} = \frac{E - E_0}{L^2} = \frac{\hbar c}{2\pi} \int_0^{\infty} dk\, k \left(\frac{k}{2} + \sum_{n=1}^{\infty} \sqrt{k^2 + n^2\pi^2/a^2} - \int_0^{\infty} dn \sqrt{k^2 + n^2\pi^2/a^2} \right). \quad (3.54)$$

This expression is still infinite, due to ultraviolet (large k!) divergences. However, we can take into account that for wavelengths smaller than atomic size, the approximation of an ideal conductors (considered as continuous medium) becomes inapplicable, Thus we have to introduce in the integrand of (3.54) some smooth cutoff function $f(k)$, which is equal to unity for $k < k_m$ and tends to zero for $k \gg k_m$, where k_m is of the order of the inverse atomic size. Then we can write

$$\mathcal{E} = \hbar c \frac{\pi^2}{4a^3} \int_0^{\infty} du \left[\frac{\sqrt{u}}{2} f\left(\frac{\pi}{a}\sqrt{u}\right) + \sum_{n=1}^{\infty} \sqrt{u + n^2} f\left(\frac{\pi}{a}\sqrt{u + n^2}\right) - \int_0^{\infty} dn \sqrt{u + n^2} f\left(\frac{\pi}{a}\sqrt{u + n^2}\right) \right], \quad (3.55)$$

where we have introduced the dimensionless integration variable $u = a^2k^2/\pi^2$. The last expression can be rewritten as

$$\mathcal{E} = \hbar c \frac{\pi^2}{4a^3} \left[\frac{1}{2}F(0) + F(1) + F(2) + \cdots - \int_0^{\infty} dn\, F(n) \right], \quad (3.56)$$

where we have defined the function

$$F(n) = \int_0^{\infty} du \sqrt{u + n^2} f\left(\frac{\pi}{a}\sqrt{u + n^2}\right). \quad (3.57)$$

For $n \to \infty$ we have $F(n) \to 0$ due to the properties of the cutoff function. To calculate the difference between the sum and the integral in square brackets in equation (3.56), we may use

the Eiler–Maclaurin summation formula, writing it as

$$\frac{1}{2}F(0)+F(1)+F(2)+\ldots-\int_0^\infty dn\,F(n)=-\frac{1}{2!}B_2F'(0)-\frac{1}{4!}B_4F'''(0)+\cdots\,, \tag{3.58}$$

where B_ν are Bernoulli numbers, defined by the series

$$\frac{y}{e^y-1}=\sum_{\nu=0}^{\infty}B_\nu\frac{y^\nu}{\nu!}\,. \tag{3.59}$$

In particular, $B_2=1/6$, $B_4=-1/30,\ldots$. We have

$$F(n)=\int_{n^2}^\infty du\sqrt{u}\,f\left(\frac{\pi\sqrt{u}}{a}\right),\quad F'(n)=-2n^2f\left(\frac{\pi n}{a}\right). \tag{3.60}$$

Assuming that $f(0)=1$ and all its derivatives are zero at the same value of its argument, we have $F'(0)=0$, $F'''(0)=-4$, while all higher order derivatives of F are zero. Thus, the value of the cutoff does not enter into the final results and we get

$$\mathcal{E}=\frac{\hbar c\pi^2}{a^3}\frac{B_4}{4!}=-\frac{\pi^2}{720}\frac{\hbar c}{a^3}\,. \tag{3.61}$$

Then, the force (per unit square) acting upon the plates are

$$\mathcal{F}=-\frac{\pi^2}{240}\frac{\hbar c}{a^4}\,. \tag{3.62}$$

The negative sign here corresponds to attraction. It is remarkable that the existence of this (quite weak!) attractive force, due to vacuum fluctuations of the electromagnetic field, was experimentally confirmed, and the theoretical expression (3.62) was directly checked. It is even more surprising that the existence of the Casimir force has to be taken into account [32] during construction and work of modern micromachines! This proves, beyond any doubt, that "zero-level" oscillations of an electromagnetic field are quite real.

3.2 Bosons

3.2.1 Scalar particles

Consider particles with spin 0. The state of a free spinless particle is completely determined by its momentum $\mathbf{p}$. Its energy $\varepsilon_{\mathbf{p}}$ is defined by

$$\varepsilon_{\mathbf{p}}^2=\mathbf{p}^2+m^2\quad\text{or}\quad p^2=m^2 \tag{3.63}$$

or, as usually expressed, the particle is on its "mass surface". Energy-momentum conservation follows from the homogeneity of space-time. In quantum mechanics, the requirement of the symmetry towards an arbitrary translation of a coordinate system means that the wave function of a particle with 4-momentum p is multiplied (as a result of translation) by a phase factor (with modulus 1). This requirement is obeyed only by the plane wave:

$$\text{const}\cdot e^{-ipx}\,,\quad px=\varepsilon_{\mathbf{p}}t-\mathbf{pr}\,. \tag{3.64}$$

The wave equation for our particles should have (3.64) as a partial solution for any p, satisfying equation (3.63). This equation should also be linear, expressing the superposition principle: any linear combination of solutions also describes the possible state of a free particle. And finally, this equation should be of a sufficiently low order in derivatives.

The spin of a particle is its angular momentum in a coordinate system at rest, and the state of a particle in a system at rest is described by nonrelativistic quantum mechanics. Then, if a particle spin in a resting system is equal to s, its wave function in this coordinate system should have $2s + 1$ components (i. e., be represented by a three-dimensional spinor of rank $2s$) [35]. The particle with spin $s = 0$ in a resting system is described by a three-dimensional scalar. However, this three-dimensional scalar can have a double four-dimensional "origin" [6]: it can be a four-dimensional scalar φ, but it also can be a time component ψ_0 of some time-like 4-vector ψ_μ, such that in a system at rest only a ψ_0 component is different from zero. Tensors of higher ranks are not to be taken into account, as they will lead to differential equations of higher orders.

For a free particle, the only differential operator which can enter the wave equation is the operator of 4-momentum p:

$$p^\mu = i\,\partial^\mu = \left(i\frac{\partial}{\partial t}, -i\nabla\right). \tag{3.65}$$

A wave equation can be written as a differential relation between φ and ψ_μ, constructed with the help of operator p^μ, and satisfying the condition of relativistic invariance. Obviously, the simplest variant of such relation has the following form:

$$p_\mu\varphi = m\psi_\mu\,, \quad p^\mu\psi_\mu = m\varphi\,, \tag{3.66}$$

where m is a scalar, characterizing the particle[5]. Substituting ψ_μ from the first equation in (3.66) to the second one, we get

$$(p^2 - m^2)\varphi = 0\,, \tag{3.67}$$

which coincides with the Klein–Gordon equation (2.10), (2.29) for the *scalar* field φ. Substituting $\varphi \sim e^{-ipx}$ into (3.67), we obtain $p^2 = m^2$, so that (3.63) is satisfied, and the scalar m is just the rest mass of our particle. As (3.63) is valid for a relativistic particle with arbitrary spin, the Klein–Gordon equation is obeyed, in fact, by wave function components of particles with any spin.

The properties of a scalar field, satisfying the Klein–Gordon equation, were already discussed in detail above. For generality we shall consider here, from the very beginning, the case of a complex field. Its energy-momentum tensor, similar to (2.50), is given by

$$T^{\mu\nu} = (\partial^\mu\varphi^*)(\partial^\nu\varphi) + (\partial^\mu\varphi)(\partial^\nu\varphi^*) - g^{\mu\nu}\mathcal{L}\,, \tag{3.68}$$

[5] There is no sense in introducing two scalars m_1, m_2, as they can always be made equal by the appropriate redefinition of φ, ψ_μ

where the Lagrangian $\mathcal{L}$ is defined in (2.56). In particular,

$$T_{00} = \frac{\partial\varphi^*}{\partial t}\frac{\partial\varphi}{\partial t} + \nabla\varphi^* \cdot \nabla\varphi + m^2\varphi^*\varphi\,, \tag{3.69}$$

$$T_{i0} = \frac{\partial\varphi^*}{\partial t}\frac{\partial\varphi}{\partial x^i} + \frac{\partial\varphi^*}{\partial x^i}\frac{\partial\varphi}{\partial t}\,. \tag{3.70}$$

Then the 4-momentum of the field is determined by the integral

$$P_\mu = \int d^3\mathbf{r}T_{\mu 0}\,. \tag{3.71}$$

From (3.69) we can see that $T_{00} > 0$, so that energy is positively defined, which, in fact, determines the choice of signs in the Lagrangian.

Equation (3.69) can be used to normalize the field. The plane wave, normalized to "single particle in volume $V = 1$" can be written as

$$\psi_{\mathbf{p}} = \frac{1}{\sqrt{2\varepsilon_{\mathbf{p}}}}e^{-ipx}\,. \tag{3.72}$$

Calculating (3.69) with (3.72), we obtain $T_{00} = \varepsilon_{\mathbf{p}}$, so that the total energy in the volume $V = 1$ is equal to the energy of a single particle.

Let us now proceed to quantization. Let us consider an expansion of an arbitrary wave function (field) over the complete set of eigenfunctions of a free particle, e. g., the plane waves ψ_p from (3.72):

$$\varphi = \sum_{\mathbf{p}} a_{\mathbf{p}}\psi_{\mathbf{p}}\,, \quad \varphi^* = \sum_{\mathbf{p}} a^*_{\mathbf{p}}\psi^*_{\mathbf{p}}\,. \tag{3.73}$$

Quantization reduces to the replacement of the coefficients $a_{\mathbf{p}}, a^*_{\mathbf{p}}$ by the corresponding *operators* of annihilation and creation of particles $\hat{a}_{\mathbf{p}}\,, \hat{a}^+_{\mathbf{p}}$.

The principal aspect of relativistic theory is the existence of *two* solutions for equation (3.63), which gives for the energy of a particle

$$\varepsilon_{\mathbf{p}} = \pm\sqrt{\mathbf{p}^2 + m^2}\,. \tag{3.74}$$

Physically sensible are only $\varepsilon_{\mathbf{p}} > 0$, as the negative particle energies correspond to the instability of the system (absence of the ground state). We cannot just drop the solutions with $\varepsilon_{\mathbf{p}} < 0$, as the general solution of the wave equation is given by the superposition of all independent partial solutions, and the expansion of the field should be performed over the complete set of eigenfunctions. Let us write

$$\varphi = \sum_{\mathbf{p}} \frac{1}{\sqrt{2\varepsilon_{\mathbf{p}}}}a^{(+)}_{\mathbf{p}}e^{i(\mathbf{pr}-\varepsilon_{\mathbf{p}}t)} + \sum_{\mathbf{p}} \frac{1}{\sqrt{2\varepsilon_{\mathbf{p}}}}a^{(-)}_{\mathbf{p}}e^{i(\mathbf{pr}+\varepsilon_{\mathbf{p}}t)}\,, \tag{3.75}$$

where in the first sum the plane waves correspond to positive, while in the second they correspond to negative frequencies. Here and below we take $\varepsilon_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}$, i. e., the positive definite energy of the physical particle.

The recipe for the correct transition to the second quantization can now be formulated in the following way:

- $a_{\mathbf{p}}^{(+)} \to \hat{a}_{\mathbf{p}}$ is the annihilation operator of a particle with momentum $\mathbf{p}$;
- $a_{\mathbf{p}}^{(-)} \to \hat{b}_{-\mathbf{p}}^{+}$ is the creation operator of an *antiparticle* with momentum $-\mathbf{p}$.

The last change to be made is due to the time dependence in the second sum in (3.75) being $e^{i\varepsilon_{\mathbf{p}}t} = (e^{-i\varepsilon_{\mathbf{p}}t})^*$, which corresponds to the appearance of one "extra" particle with energy $\varepsilon_{\mathbf{p}}$ in the final state (during the calculation of any matrix element, which includes φ). Now, replacing in the second sum $\mathbf{p} \to -\mathbf{p}$, we write

$$\hat{\varphi} = \sum_{\mathbf{p}} \frac{1}{\sqrt{2\varepsilon_{\mathbf{p}}}} \left(\hat{a}_{\mathbf{p}} e^{-ipx} + \hat{b}_{\mathbf{p}}^{+} e^{ipx} \right),$$
$$\hat{\varphi}^{+} = \sum_{\mathbf{p}} \frac{1}{\sqrt{2\varepsilon_{\mathbf{p}}}} \left(\hat{a}_{\mathbf{p}}^{+} e^{ipx} + \hat{b}_{\mathbf{p}} e^{-ipx} \right). \tag{3.76}$$

Now operators $\hat{a}_{\mathbf{p}}$ and $\hat{b}_{\mathbf{p}}$ in expansion (3.76) are multiplied by "correct" factors like $e^{-i\varepsilon_{\mathbf{p}}t}$, while operators $\hat{a}_{\mathbf{p}}^{+}$ and $\hat{b}_{\mathbf{p}}^{+}$ are multiplied by complex conjugate factors like $e^{i\varepsilon_{\mathbf{p}}t}$. Both types of particles (particles and antiparticles), represented by the creation and annihilation operators entering the field operator $\widehat{\varphi}$ have the same masses.

Substituting operator expansion (3.76) into (3.69) and integral $\int d^3\mathbf{r} T_{00}$, determining the energy of the field, we obtain the Hamiltonian of the field as

$$H = \sum_{\mathbf{p}} \varepsilon_{\mathbf{p}} \left(\hat{a}_{\mathbf{p}}^{+} \hat{a}_{\mathbf{p}} + \hat{b}_{\mathbf{p}} \hat{b}_{\mathbf{p}}^{+} \right). \tag{3.77}$$

A physically reasonable result for the eigenvalues of this operator (positive definite energy) is obtained only if the creation and annihilation operators satisfy *Bose* commutation relations:

$$[\hat{a}_{\mathbf{p}}, \hat{a}_{\mathbf{p}}^{+}] = [\hat{b}_{\mathbf{p}}, \hat{b}_{\mathbf{p}}^{+}] = 1\,, \quad [\hat{a}_{\mathbf{p}}, \hat{b}_{\mathbf{p}}] = [\hat{a}_{\mathbf{p}}, \hat{b}_{\mathbf{p}}^{+}] = \cdots = 0\,. \tag{3.78}$$

In fact, using these commutation relations, we can write the Hamiltonian (3.77) in the following form:

$$H = \sum_{\mathbf{p}} \varepsilon_{\mathbf{p}} (a_{\mathbf{p}}^{+} a_{\mathbf{p}} + b_{\mathbf{p}}^{+} b_{\mathbf{p}} + 1)\,. \tag{3.79}$$

We have already seen above that in occupation number representation the eigenvalues of Bose operators $a_{\mathbf{p}}^{+} a_{\mathbf{p}}$ and $b_{\mathbf{p}}^{+} b_{\mathbf{p}}$ are given by nonnegative integers, which we shall denote as $N_{\mathbf{p}}$ and $\bar{N}_{\mathbf{p}}$, respectively (the numbers of particles and antiparticles in the state with a given momentum). Then both energy and momentum of the field can be

written (dropping the infinite energy of the vacuum) as

$$E = \sum_{\mathbf{p}} \varepsilon_{\mathbf{p}}(N_{\mathbf{p}} + \bar{N}_{\mathbf{p}}) , \tag{3.80}$$

$$\mathbf{P} = \sum_{\mathbf{p}} \mathbf{p}(N_{\mathbf{p}} + \bar{N}_{\mathbf{p}}) . \tag{3.81}$$

Formal derivation of the last expression can be performed with the help of (3.70) and (3.71). Assuming anticommutation (Fermi-like) relations for creation and annihilation operators, we obtain, instead of (3.79), an expression like $H = \sum_{\mathbf{p}} \varepsilon_{\mathbf{p}}(a^+_{\mathbf{p}} a_{\mathbf{p}} - b^+_{\mathbf{p}} b_{\mathbf{p}} + 1)$, which is not positively defined (leading to the absence of the ground state of the system). Thus, the particles with spin 0 (scalar particles) are Bosons. This is actually a proof of the spin-statistics theorem for this simplest case of the scalar field.

We have seen above, that for complex scalar field we have charge conservation (2.65). Replacing in the expression (2.63) for current density classical fields φ, φ^* by operators $\hat{\varphi}$, $\hat{\varphi}^+$ from (3.76), and making elementary calculations, we obtain from (2.65)

$$Q = \sum_{\mathbf{p}} (a^+_{\mathbf{p}} a_{\mathbf{p}} - b_{\mathbf{p}} b^+_{\mathbf{p}}) = \sum_{\mathbf{p}} (a^+_{\mathbf{p}} a_{\mathbf{p}} - b^+_{\mathbf{p}} b_{\mathbf{p}} - 1) , \tag{3.82}$$

where, while transforming to the last equality, we have again used the commutation relations (3.78). The eigenvalues of this operator, without the vacuum contribution, are written as

$$Q = \sum_{\mathbf{p}} (N_{\mathbf{p}} - \bar{N}_{\mathbf{p}}) , \tag{3.83}$$

so that the charges of particles and antiparticles are opposite in sign. Note that now (after quantization!) the charge can change only in a discrete way.

3.2.2 Truly neutral particles

Above we have considered operators $\hat{a}_{\mathbf{p}}$ and $\hat{b}_{\mathbf{p}}$ as referring to different particles. This is not always so – we may consider a specific case, when operators entering the expansion of $\hat{\varphi}$ refer to the same particles (we have already met this situation in case of photons). Then

$$\hat{\varphi} = \sum_{\mathbf{p}} \frac{1}{\sqrt{2\varepsilon_{\mathbf{p}}}} \left(\hat{c}_{\mathbf{p}} e^{-ipx} + \hat{c}^+_{\mathbf{p}} e^{ipx} \right) , \tag{3.84}$$

so that the particle just coincides with its antiparticle, and we are dealing with the so-called truly neutral particles. Now the field operator is Hermitian: $\hat{\varphi} = \hat{\varphi}^+$, which is an analogue of the real field in classical field theory. Naturally, this field possesses twice less degrees of freedom in comparison with complex field, and its Lagrangian takes the form as in (2.28). Accordingly, we can calculate the energy-momentum tensor and

obtain for the energy density the following expression:

$$T_{00} = \frac{1}{2}\left\{\left(\frac{\partial\varphi}{\partial t}\right)^2 + (\nabla\varphi)^2 + m^2\varphi^2\right\}. \tag{3.85}$$

Then, substituting expansion (3.84) into $\int d^3\mathbf{r}T_{00}$ we get the Hamiltonian as

$$H = \frac{1}{2}\sum_{\mathbf{p}} \varepsilon_{\mathbf{p}}\left(\hat{c}_{\mathbf{p}}^{+}\hat{c}_{\mathbf{p}} + \hat{c}_{\mathbf{p}}\hat{c}_{\mathbf{p}}^{+}\right). \tag{3.86}$$

Again we see the necessity to quantize using Bose rules, so that commutation relations for creation and annihilation operators are written as

$$[\hat{c}_{\mathbf{p}}, \hat{c}_{\mathbf{p}}^{+}] = 1\,, \quad [\hat{c}_{\mathbf{p}}, \hat{c}_{\mathbf{p}}] = [\hat{c}_{\mathbf{p}}^{+}, \hat{c}_{\mathbf{p}}^{+}] = 0\,. \tag{3.87}$$

The Hamiltonian is

$$H = \sum_{\mathbf{p}} \varepsilon_{\mathbf{p}}\left(c_{\mathbf{p}}^{+}c_{\mathbf{p}} + \frac{1}{2}\right), \tag{3.88}$$

so that, after dropping the vacuum contribution, its eigenvalues are given by

$$E = \sum_{\mathbf{p}} \varepsilon_{\mathbf{p}} N_{\mathbf{p}}\,. \tag{3.89}$$

It is obvious that for a Hermitian (real in classical limit) field both current density and charge are zero.

Note that from previously discussed physical particles an example of a truly neutral particle was the photon, and the Hermiticity of corresponding quantum field was relevant to the measurability of quantum electric and magnetic fields.

Remarks on the Lorentz group

According to the special theory of relativity, all inertial reference systems are equivalent. If two coordinate systems move relative to each other along direction x_1 with velocity v, the connection between corresponding coordinates is expressed by Lorentz transformation [33]:

$$\begin{aligned} x'^0 &= \gamma(x^0 - \beta x^1) = x^0 chu - x^1 shu\,, \\ x'^1 &= \gamma(x^1 - \beta x^0) = x^1 chu - x^0 shu\,, \\ x'^2 &= x^2\,, \qquad x'^3 = x^3\,, \end{aligned} \tag{3.90}$$

where

$$\gamma = \frac{1}{\sqrt{1-\beta^2}}\,, \quad \beta = \frac{v}{c}\,, \quad thu = \beta\,. \tag{3.91}$$

In the general case, we postulate the invariance of physical laws with respect to linear coordinate transformations (inhomogeneous Lorentz transformations):

$$x^\mu \to x'^\mu = \Lambda^\mu_\nu x^\nu + a^\mu\,, \tag{3.92}$$

which conserve the square of the interval

$$(x-y)^2 = (x_\mu - y_\mu)(x^\mu - y^\mu) = g_{\mu\nu}(x^\mu - y^\mu)(x^\nu - y^\nu)\,. \tag{3.93}$$

In (3.92) translation is performed after the homogeneous transformation. Inhomogeneous Lorentz transformations are also called Poincaré transformations.

Among possible coordinate transformations, we may consider not only translations and rotations in pseudo-Euclidean space-time, but also space and time inversions, which we shall denote by P, T, and PT:

$$\begin{aligned} Px^k &= -x^k\,, \quad Px^0 = x^0\,,\\ Tx^k &= x^k\,, \quad Tx^0 = -x^0\,,\\ PTx^\mu &= -x^\mu\,. \end{aligned} \tag{3.94}$$

The interval (3.93) does not change under transformations (3.92) if

$$\Lambda^\nu_\mu \Lambda^\mu_\sigma = \delta^\nu_\sigma\,, \qquad \Lambda^\nu_\mu = g_{\mu\rho}\Lambda^\rho_\beta g^{\beta\nu}\,. \tag{3.95}$$

In matrix form the last relation is written as

$$\tilde{\Lambda} g \Lambda = g\,, \tag{3.96}$$

where the tilde denotes matrix transposition. Then, it is clear that

$$\mathrm{Det}\,\Lambda = \pm 1\,. \tag{3.97}$$

From (3.95) it also follows that

$$(\Lambda^{00})^2 - \sum_k (\Lambda^{0k})^2 = 1\,, \tag{3.98}$$

so that $(\Lambda^{00})^2 \geq 1$. Correspondingly, there are two possibilities:

$$\Lambda^{00} \geq 1\,, \qquad \Lambda^{00} \leq -1\,. \tag{3.99}$$

Thus, the general transformations (3.92) can be divided into four classes:

1. $\mathcal{P}^\uparrow_+$: $\mathrm{Det}\,\Lambda = 1$, $\Lambda^{00} \geq 1$
   No time and space inversions. Only rotations and translations in pseudo-Euclidean space, which form a proper orthochronous Poincare group.
2. $\mathcal{P}^\downarrow_+$: $\mathrm{Det}\,\Lambda = 1$, $\Lambda^{00} \leq -1$
   Here the T-operation is included. Due to the unimodular nature of transformations, the P-operation is also included. Any transformation from $\mathcal{P}^\downarrow_+$ can be represented by the product $\mathcal{P}^\uparrow_+$ è PT. In particular, 4-inversion $PT \in \mathcal{P}^\downarrow_+$, while P and T do not enter $\mathcal{P}^\downarrow_+$ separately, due to $\mathrm{Det}\,\Lambda = 1$. $\mathcal{P}^\uparrow_+$ and $\mathcal{P}^\downarrow_+$ transformations together form the proper Poincare group $\mathcal{P}_+$.
3. $\mathcal{P}^\uparrow_-$: $\mathrm{Det}\,\Lambda = -1$, $\Lambda^{00} \geq 1$
   Corresponding transformations have the form $P\mathcal{P}^\uparrow_+$. Together with $\mathcal{P}^\uparrow_+$ they form an orthochronous Poincaré group.

4. $\mathcal{P}_{-}^{\downarrow}$: $\mathrm{Det}\,\Lambda = -1$, $\Lambda^{00} \leq -1$

Time direction changes. Any transformation from this class can be written as $T\mathcal{P}_{+}^{\uparrow}$.

The general Poincaré group can now be represented by the sum

$$\mathcal{P} = \mathcal{P}_{+}^{\uparrow} + PT\mathcal{P}_{+}^{\uparrow} + P\mathcal{P}_{+}^{\uparrow} + T\mathcal{P}_{+}^{\uparrow}. \tag{3.100}$$

Of all these components of the Poincaré group, only $\mathcal{P}_{+}^{\uparrow}$ contains the unit transformation. Thus, transformations from different classes cannot be connected by the continuous transformation belonging to $\mathcal{P}_{+}^{\uparrow}$. Transformations from the same class can be obtained from each other by transformations from $\mathcal{P}_{+}^{\uparrow}$.

3.2.3 *CPT*-transformations

Space inversion

Discrete symmetries, like space or time inversions and charge conjugation (replacement particles by antiparticles), are of major importance in quantum field theory. For example, space inversion is defined as

$$P\mathbf{r} = -\mathbf{r}. \tag{3.101}$$

Under this transformation the scalar field can be transformed as

$$P\varphi(t,\mathbf{r}) = \pm\varphi(t,-\mathbf{r}), \tag{3.102}$$

where signs correspond to the usual scalar or pseudoscalar. In nonrelativistic quantum mechanics, the behavior of the wave function of the system under space inversion is related simply to its coordinate dependence, which leads to the concept of orbital parity [35]:

$$\psi(t,-\mathbf{r}) = \pm\psi(t,\mathbf{r}). \tag{3.103}$$

In quantum field theory we are speaking about behavior of the field at a given point in space, and equation (3.102) defines the *internal* parity of corresponding particles. Total parity of the particle system is equal to the product of their internal parities and orbital parity of their relative motion. "Internal" symmetry properties of different particles become manifest only in the processes of particle transmutations during reaction between particles.

For the second quantized fields, internal parity is expressed via appropriate behavior of $\widehat{\varphi}$-operators. For scalar or pseudoscalar fields we have

$$P\widehat{\varphi}(t,\mathbf{r}) = \pm\widehat{\varphi}(t,-\mathbf{r}). \tag{3.104}$$

The action of the P-operation on the $\widehat{\varphi}$-operator can be formulated as transformation rules for creation and annihilation operators of particles, which correspond to (3.104).

Using (3.76) it is easy to find that these rules take the form

$$P: \quad \begin{aligned} &a_{\mathbf{p}} \to \pm a_{-\mathbf{p}}\,, \quad b_{\mathbf{p}} \to \pm b_{-\mathbf{p}}\,, \\ &a^{+}_{\mathbf{p}} \to \pm a^{+}_{-\mathbf{p}}\,, \quad b^{+}_{\mathbf{p}} \to \pm b^{+}_{-\mathbf{p}}\,. \end{aligned} \tag{3.105}$$

In fact, we can write

$$\varphi(t,\mathbf{r}) = \sum_{\mathbf{p}} \frac{1}{\sqrt{2\varepsilon_{\mathbf{p}}}}\left(a_{\mathbf{p}} e^{-i\varepsilon_{\mathbf{p}} t + i\mathbf{pr}} + b^{+}_{\mathbf{p}} e^{i\varepsilon_{\mathbf{p}} t - i\mathbf{pr}}\right), \tag{3.106}$$

and after the operation (3.105) and the change of the summation variable $\mathbf{p} \to -\mathbf{p}$ we immediately obtain $\pm\varphi(t,-\mathbf{r})$[6]. Note that transformation (3.105) is pretty obvious – inversion simply changes the sign of polar vector $\mathbf{p}$.

Charge conjugation

Replacement of particles by antiparticles can be made in the field operator (3.76) by an obvious operation:

$$C: \quad a_{\mathbf{p}} \to b_{\mathbf{p}}\,, \quad b_{\mathbf{p}} \to a_{\mathbf{p}}\,. \tag{3.107}$$

Then $\varphi \to \varphi^{C}$, where

$$\varphi^{C}(t,\mathbf{r}) = \varphi^{+}(t,\mathbf{r})\,. \tag{3.108}$$

The meaning of this transformation does not change if we introduce an arbitrary phase factor:

$$a_{\mathbf{p}} \to e^{i\alpha} b_{\mathbf{p}}\,, \quad b_{\mathbf{p}} \to e^{-i\alpha} a_{\mathbf{p}}\,, \tag{3.109}$$

so that

$$\varphi \to e^{i\alpha}\varphi^{+}\,, \quad \varphi^{+} \to e^{-i\alpha}\varphi\,. \tag{3.110}$$

If we perform the charge conjugation twice, we obtain the identical transformation $\varphi \to \varphi$. Symmetry towards the replacement of particles by antiparticles, in the general case, does not lead to any new particle characteristics, and operator C does not have eigenstates and eigenvalues. The only exception is the system containing the equal number of particles and antiparticles. Operator C transforms such a system into itself, and in this case it has eigenvalues $C = \pm 1$ (as $C^2 = 1$, which is obvious). The same is valid for truly neutral particles, when $\varphi^{C} = \pm\varphi$, and we can speak of charge parity.

Four-dimensional inversion and the inversion of time

Four-dimensional inversion is defined as

$$x \to -x\,, \quad \text{where} \quad x = (\mathbf{r}, t)\,. \tag{3.111}$$

[6] Note that below, in most cases, we shall not use the cap-sign of operator for creation and annihilation operators, as well as for other filed operators, hoping that this will not lead to any misunderstandings.

This operation can be considered as some four-dimensional rotation or, in other words, as some Lorentz transformation, because the determinant of the transformation matrix in both cases is equal to unity. The situation here is different from the case of three-dimensional (spatial) inversion, where the determinant is equal to -1. Thus, any expression invariant with respect to Lorentz transformations is also invariant to four-dimensional inversion. With respect to the operator of the scalar field (scalar with respect to four-dimensional rotations) this means that

$$\varphi(t,\mathbf{r}) \to \varphi(-t,-\mathbf{r})\,. \tag{3.112}$$

In terms of the creation and annihilation operators, transformation (3.112) is achieved by interchanging the coefficients before e^{-ipx} and e^{ipx} in equation (3.76), which gives

$$CPT: \quad a_{\mathbf{p}} \to b_{\mathbf{p}}^{+}\,, \quad b_{\mathbf{p}} \to a_{\mathbf{p}}^{+}\,. \tag{3.113}$$

Thus, this transformation includes the replacement of particles by antiparticles, so that in relativistic field theory we *automatically* obtain the invariance with respect to transformation when we simultaneously perform P and T, and also the charge conjugation C. This is the content of the so-called CPT-theorem, which is one of the most general statements of quantum field theory: nothing in nature will change if we simultaneously with 4-inversion (inversion of both space coordinates and time) replace all particles by antiparticles! Transformation (3.113) can be written also in the form

$$\varphi^{CPT}(t,\mathbf{r}) = \varphi(-t,-\mathbf{r})\,. \tag{3.114}$$

Then it is easy to formulate the recipe for T-inversion (inversion of time). This operation should be defined so that, together with CP, it reduces to CPT-transformation (3.113). Taking into account (3.105) and (3.107), we find

$$T: \quad a_{\mathbf{p}} \to \pm a_{-\mathbf{p}}^{+}\,, \quad b_{\mathbf{p}} \to \pm b_{-\mathbf{p}}^{+}\,, \tag{3.115}$$

where the signs correspond to the signs in equation (3.105). Thus, time inversion not only transforms the motion with momentum $\mathbf{p}$ to the motion with momentum $-\mathbf{p}$, but also interchanges the initial and final states in all matrix elements, which leads to the replacement of the annihilation operators of particles with momentum $\mathbf{p}$ by the creation operators of particles with momentum $-\mathbf{p}$ (and vice versa). From (3.115) and (3.106), with replacement $\mathbf{p} \to -\mathbf{p}$, we get

$$\varphi^{T}(t,\mathbf{r}) = \pm\varphi^{+}(-t,\mathbf{r})\,. \tag{3.116}$$

In fact, here we have the full correspondence with time inversion in quantum mechanics [35]: if some state is described by the wave function $\psi(t,\mathbf{r})$, the time-inverted state is described by $\psi^{*}(-t,\mathbf{r})$.

Transformations T and CPT-interchange initial and final states, and for these transformations there are no notions like eigenstates and eigenvalues. They do not lead

to any new characteristics of particles. Due to relativistic invariance, the operator of CPT-transformation should commute with the arbitrary Hamiltonian (Lagrangian) of relativistic field theory. As to C and P (i. e., also T) separately, this is not so in general. In particular, weak interactions of elementary particles are not invariant with respect to spatial inversion P, and even with respect to combined CP transformation. This last (very small!) breaking of symmetry, according to the CPT-theorem, leads to a weak nonequivalence of time directions in nature, which leads to some significant consequences for cosmology. For example, Sakharov proposed an idea that this symmetry-breaking can explain the overwhelming domination of matter over antimatter in the modern state of the universe.

Discrete transformations of current

Consider the operator of conserving current of the scalar field, which, with the help of (2.63), can be written as

$$j^\mu = i(\varphi^+ \partial^\mu \varphi - \varphi \partial^\mu \varphi^+) \,. \tag{3.117}$$

Transformation (3.104), with the obvious replacement $(\partial_0, \boldsymbol{\partial}) \to (\partial_0, -\boldsymbol{\partial})$, gives

$$P: \quad (j^0, \mathbf{j})_{t,\mathbf{r}} \to (j^0, -\mathbf{j})_{t,-\mathbf{r}} \,, \tag{3.118}$$

as it should be for a true 4-vector.

Similarly, charge conjugation (3.108) gives

$$C: \quad (j^0, \mathbf{j})_{t,\mathbf{r}} \to (-j^0, -\mathbf{j})_{t,\mathbf{r}} \,, \tag{3.119}$$

if operators φ and φ^+ commute. Strictly speaking, they do not commute, but this is irrelevant – this noncommutativity appears only due to the noncommutativity of the creation and annihilation operators with the same $\mathbf{p}$, which leads to the appearance of terms, independent of occupation numbers, i. e., independent of the state of the field. Dropping these terms we still obtain (3.119). From (3.119) it is seen that the change of particles by antiparticles leads to the change of the sign of all the components of the current.

The operation of time inversion is accompanied by the interchange of the initial and final states, so that being applied to the product of operators it changes the order of the operators in this product, for example

$$(\varphi^+ \partial_\mu \varphi)^T = (\partial_\mu \varphi)^T (\varphi^+)^T \,. \tag{3.120}$$

According to the remark after equation (3.119), this is irrelevant, and the return to the initial order does not change the results. Taking into account that under T-inversion $(\partial_0, \boldsymbol{\partial}) \to (-\partial_0, \boldsymbol{\partial})$, with the help of (3.116), we obtain

$$T: \quad (j^0, \mathbf{j})_{t,\mathbf{r}} \to (j^0, -\mathbf{j})_{-t,\mathbf{r}} \,, \tag{3.121}$$

so that the three-dimensional current $\mathbf{j}$ changes its sign, in accordance with the classical meaning of the current, while the charge density j^0 does not change.

Finally, under 4-inversion (3.112) we easily obtain

$$CPT: \quad (j^0, \mathbf{j})_{t,\mathbf{r}} \to (-j^0, -\mathbf{j})_{-t,-\mathbf{r}} \tag{3.122}$$

in accordance with the CPT-nature of this transformation.

The operator of electromagnetic interaction is proportional to $j_\mu A^\mu$ and is invariant to CPT, as any other relativistic interaction. Accordingly, using (3.118), (3.119), and (3.121) it is not difficult to obtain transformation rules for electromagnetic potential $A^\mu = (A_0, \mathbf{A})$:

$$\begin{aligned} C: &\quad (A_0, \mathbf{A})_{t,\mathbf{r}} \to (-A_0, -\mathbf{A})_{t,\mathbf{r}} \\ P: &\quad (A_0, \mathbf{A})_{t,\mathbf{r}} \to (A_0, -\mathbf{A})_{t,-\mathbf{r}} \\ T: &\quad (A_0, \mathbf{A})_{t,\mathbf{r}} \to (A_0, -\mathbf{A})_{-t,\mathbf{r}} \\ CPT: &\quad (A_0, \mathbf{A})_{t,\mathbf{r}} \to (-A_0, -\mathbf{A})_{-t,-\mathbf{r}}\,. \end{aligned} \tag{3.123}$$

Similar transformation rules are also valid for Yang–Mills gauge fields.

3.2.4 Vector bosons

The particle with spin 1 in its rest system is described by the three-component wave function – a three-dimensional vector (vector Boson). By its four-dimensional "origin" these may be three spatial components of 4-vector ψ^μ (space-like!) or three components of the antisymmetric second rank tensor $\psi^{\mu\nu}$, for which in a rest system the corresponding time-components ψ^0, ψ^{00} and spatial ψ^{ik} components become zero.

The wave equation can again be written as a differential relation between ψ^μ and $\psi^{\mu\nu}$ in the form

$$i\psi_{\mu\nu} = p_\mu \psi_\nu - p_\nu \psi_\mu\,, \tag{3.124}$$

$$p^\nu \psi_{\mu\nu} = i m^2 \psi_\mu\,, \tag{3.125}$$

where $p_\mu = i\partial_\mu$ is the momentum operator. These are Proca equations for a vector field. Applying p^μ to both sides of (3.125), we obtain (due to the antisymmetry of $\psi_{\mu\nu}$)

$$p^\mu \psi_\mu = 0\,. \tag{3.126}$$

Then, excluding $\psi_{\mu\nu}$ from (3.124), (3.125) (substituting the first equation into the second) and taking into account (3.126), we obtain

$$(p^2 - m^2)\psi_\mu = 0\,, \tag{3.127}$$

so that m, as usual, represents the particle mass. Thus, the free particle with spin 1 is described by a single 4-vector ψ_μ, the components of which satisfy the "Klein–Gordon"-like equation (3.127) and an *additional condition*, similar to the Lorentz condition (of four-dimensional transversality) (3.126), which excludes from ψ^μ a "part, belonging to spin 0".

In the rest system, ψ_μ does not depend on spatial coordinates ($\mathbf{p} = 0$), and we simply have $p^0\psi_0 = 0$, at the same time taking into account that in the rest system $p^0 = m$ we have $p^0\psi_0 = m\psi_0$. Then it is clear that in the rest system $\psi_0 = 0$, as it should be for a particle with spin 1. Together with ψ_0 in the rest system both ψ_{ik} and ψ_{00} also become zero.

The particle with spin 1 may have a different internal parity, depending on ψ^μ being a true vector or pseudovector:

$$P\psi^\mu = (\psi^0, -\boldsymbol{\psi}) \quad \text{or} \quad P\psi^\mu = (-\psi^0, \boldsymbol{\psi})\,. \tag{3.128}$$

The plane wave, normalized to a single particle in volume $V = 1$, is written as

$$\psi_\mu = \frac{1}{\sqrt{2\varepsilon_{\mathbf{p}}}} u_\mu e^{-ipx}\,, \qquad u_\mu u^{\mu *} = -1\,, \tag{3.129}$$

where u_μ is a unit 4-vector of polarization, normalized by the requirement of the space-like nature of ψ_μ, also satisfying the condition of four-dimensional transversality:

$$u_\mu p^\mu = 0\,. \tag{3.130}$$

Note that, in contrast to the case of photons, vector Bosons with spin 1 have three independent polarizations.

The Lagrangian of vector field can be written as

$$\mathcal{L} = -(\partial_\mu \psi^*_\nu)(\partial^\mu \psi^\nu) + m^2 \psi^*_\mu \psi^\mu\,. \tag{3.131}$$

The structure of this Lagrangian is similar to the case of a scalar field, but note the different overall sign! The thing is that ψ_μ is the space-like vector, so that $\psi^*_\mu \psi^\mu < 0$, while for scalar field $\varphi^* \varphi > 0$, so that the sign is chosen to guarantee the positive definiteness of energy in the classical limit. In fact, the practical use of the Lagrangian (3.131) reduces not only to the derivation of equations of motion, but also to the introduction of the energy-momentum tensor and current. It is easy to find that

$$T_{\mu\nu} = -\partial_\mu \psi^{\lambda *} \partial_\nu \psi_\lambda - \partial_\nu \psi^{\lambda *} \partial_\mu \psi_\lambda - \mathcal{L} g_{\mu\nu}\,, \tag{3.132}$$

$$j_\mu = -i[\psi^*_\lambda \partial_\mu \psi^\lambda - (\partial_\mu \psi^*_\lambda)\psi^\lambda]\,. \tag{3.133}$$

These expressions are similar to those obtained for scalar field and do not require further commenting.

Quantization can be performed similarly to the case of the scalar field. Again, to guarantee the physically obvious requirement of $T_{00} > 0$ and the arbitrariness of the sign of charge density j^0, we have to use Bose-like rules of quantization (commutation relations)!

Let us stress that, due to $m \neq 0$, gradient (gauge) invariance is absent. Because of this the massive vector field possesses three independent components. The absence of gauge invariance of this theory is most clearly seen from the second of the Proca equations (3.125): the value of $\psi_{\mu\nu}$ is invariant with respect to gradient transformations, so that the left-hand side of this equation is invariant, while the right-hand side is obviously noninvariant and changes under these transformations.

Particles with arbitrary integer spin

the wave function of a particle with integer spin s is represented by an irreducible 4-tensor of the rank s, i. e., tensor symmetric over all of its indices which becomes zero under contraction over any pair of its indices:

$$\psi_{\dots\mu\dots\nu\dots} = \psi_{\dots\nu\dots\mu\dots}\,, \quad \psi^{\mu\dots}_{\dots\mu\dots} = 0\,. \tag{3.134}$$

This tensor should satisfy any extra condition of four-dimensional transversality:

$$p^\mu \psi_{\dots\mu\dots} = 0\,, \tag{3.135}$$

and any of its components should satisfy the equation

$$(p^2 - m^2)\psi_{\dots\mu\dots} = 0\,. \tag{3.136}$$

In the rest system, equation (3.135) leads to the zeroes of all the components of the 4-tensor, with any of the indices equal to 0. Thus, in the rest system our field is reduced to irreducible three-dimensional tensor of rank s, with the number of independent components equal to $2s+1$.

The Lagrangian, energy-momentum tensor, and current density for the field with integer spin s differ from those just written above for the case of $s = 1$ only by the replacement ψ_μ with $\psi_{\dots\mu\dots\nu\dots}$. The normalized plane wave is written as

$$\psi^{\mu\nu\dots} = \frac{1}{\sqrt{2\varepsilon_{\mathbf{p}}}} u^{\mu\nu\dots} e^{-ipx}\,, \quad u^*_{\mu\nu\dots} u^{\mu\nu} = -1\,, \tag{3.137}$$

with

$$u^{\dots\mu\dots} p_\mu = 0\,. \tag{3.138}$$

There are in total $2s + 1$ independent polarizations.

Quantization is performed as an obvious generalization of the cases of $s = 0$ and $s = 1$.

The scheme presented above is sufficient for the description of *free* particles with integer spins. For the interacting case the situation becomes more complicated. For all integer spins with $s > 1$ it is actually impossible to formulate a variational principle, using only one (tensor) field function, with the rank, corresponding to this spin. It becomes necessary to introduce additional tensor (or spinor) entities of lower rank. Then the Lagrangian is chosen in such a way that these additional fields reduce to zero due to the equations of motion (following from variational principle) for free particles.

Note that the problem of particles with spin $s > 1$ is of rather "academic" interest, as there are no such elementary particles within the Standard Model (and forgetting about gravitons!).

3.3 Fermions

3.3.1 Three-dimensional spinors

Let us recall the description of particles with half-integer spin (Fermions) in nonrelativistic quantum mechanics [35]. A particle with spin $s = 1/2$ is described by a two-component wave function – the *spinor*, which is conveniently written as the following column:

$$\psi = \begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix} = \begin{pmatrix} \psi(1/2) \\ \psi(-1/2) \end{pmatrix}, \tag{3.139}$$

where components ψ^1 and ψ^2 correspond to spin projections $s^z = \pm 1/2$. Under an arbitrary rotation of the coordinate system, spinor components are transformed by linear transformation:

$$\psi'^1 = a\psi^1 + b\psi^2\,, \quad \psi'^2 = c\psi^1 + d\psi^2\,. \tag{3.140}$$

In other words,

$$\psi' = U\psi\,, \quad U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{3.141}$$

Transformation coefficients (matrix elements of U) are, in general, complex and are functions of the angles of rotation.

Consider a bilinear form

$$\psi^1\varphi^2 - \psi^2\varphi^1\,, \tag{3.142}$$

where ψ and φ are two spinors. Simple calculation gives

$$\psi'^1\varphi'^2 - \psi'^2\varphi'^1 = (ad - bc)(\psi^1\varphi^2 - \psi^2\varphi^1)\,, \tag{3.143}$$

so that (3.142) under coordinate system rotations (3.140) is transformed into itself. Consider now the bilinear form (3.142) as some wave function of the composite system. However, if we have a single component wave function, which is transformed under rotations into itself, it obviously corresponds to spin zero, i. e., it is a scalar and can not change under rotation at all. Thus, the coefficients of our transformation should satisfy the condition

$$ad - bc = 1\,, \quad \mathrm{Det}\, U = 1\,. \tag{3.144}$$

Then (3.142) is simply a wave function of a particle with spin $s = 0$, composed of two particles with spin $s = 1/2$. At the same time, we can introduce one more scalar, composed of spinor components (3.139):

$$\psi^1\psi^{1*} + \psi^2\psi^{2*} \tag{3.145}$$

which is just the probability density needed to find a particle in a given point of space. Transformation, which conserves the sum of squares of modules of transformed variables, is unitary, so that

$$U^+ = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix} = U^{-1}\,. \tag{3.146}$$

Taking (3.144) into account, the inverse transformation matrix takes the form

$$U^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \tag{3.147}$$

so that from unitarity we obtain

$$a = d^*\,, \quad b = -c^*\,. \tag{3.148}$$

Due to conditions (3.144) and (3.148), of the four complex coefficients a, b, c, d (i. e., of eight real numbers), in fact, only three (real) are independent, which corresponds to the three rotation angles of three-dimensional coordinate system.

Comparing scalars (3.142) and (3.145) we see that ψ^{1*} and ψ^{2*} should transform correspondingly as ψ^2 and $-\psi^1$.

Besides the *contravariant* spinor components ψ^1, ψ^2, introduced above, we may define the *covariant* components:

$$\psi_1 = \psi^2\,, \qquad \psi_2 = -\psi^1\,. \tag{3.149}$$

The invariant (3.142) can be written now as the scalar product:

$$\psi^\lambda \varphi_\lambda = \psi^1\varphi_1 + \psi^2\varphi_2 = \psi^1\varphi^2 - \psi^2\varphi^1\,. \tag{3.150}$$

Now, take into account that

$$\psi^\lambda \varphi_\lambda = \psi^1\varphi_1 + \psi^2\varphi_2 = -\psi_2\varphi^2 - \psi_1\varphi^1\,, \tag{3.151}$$

so that the following antisymmetry condition is always valid:

$$\psi^\lambda \varphi_\lambda = -\psi_\lambda \varphi^\lambda\,. \tag{3.152}$$

Then it is obvious that

$$\psi^\lambda \psi_\lambda = 0\,. \tag{3.153}$$

We can also define spinors of a higher rank. For example, we can introduce spinors of the second rank as

$$\psi^{\lambda\mu} \sim \psi^\lambda \varphi^\mu\,, \quad \psi_{\lambda\mu} \sim \psi_\lambda \varphi_\mu\,, \quad \psi^\mu_\lambda \sim \psi_\lambda \varphi^\mu\,. \tag{3.154}$$

Higher rank spinors are defined in a similar way.

Transformation from contravariant to covariant spinors can be made with the help of a "metric tensor":

$$g_{\lambda\mu} = g^{\lambda\mu} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{3.155}$$

as it is easily seen that we can write

$$\psi_\lambda = g_{\lambda\mu}\psi^\mu\,, \qquad \psi^\lambda = g^{\mu\lambda}\psi_\mu\,. \tag{3.156}$$

Consider now the multiplication and contraction of spinors. Multiplication of two spinors of the second and third ranks $\psi_{\lambda\mu}\psi^{\nu\rho\sigma}$ produces a spinor of the fifth rank. Construction of $\psi^{\nu\rho\sigma}_{\lambda\mu}$ over the pair of indices μ and ν gives the spinor of third rank $\psi^{\mu\rho\sigma}_{\lambda\mu}$. In particular, contraction of ψ^μ_λ produces the scalar ψ^λ_λ. Here we have to take into account (3.152) and (3.153), so that $\psi^\lambda_\lambda = -\psi^\lambda_\lambda$. Then it follows that the contraction over two indices of any *symmetric* (to permutation of indices) spinor produces zero! In particular, for the symmetric spinor of the second rank $\psi_{\lambda\mu} = \psi_{\mu\lambda}$ we have

$\psi^{\lambda}_{\lambda} = 0$. A spinor symmetric over all indices of any rank can always be constructed by the appropriate symmetrization (i. e., by taking the sum of the spinors with all the permutations of the indices). We have shown that the contraction over a pair of indices of a symmetric spinor can not produce spinors of a lower rank. From a mathematical point of view these spinors realize irreducible representations of the three-dimensional rotation group $SU(2)$.

By definition of the angular momentum (spin) s operator $1 + i\,\delta\varphi(\mathbf{n}\cdot\mathbf{s})$ describes the rotation by infinitesimal angle $\delta\varphi$ around an axis, oriented along the unit vector $\mathbf{n}$ [35]. For spin $s = 1/2$ we have $\mathbf{s} = \frac{1}{2}\sigma$, where σ is the set of three Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.157}$$

The corresponding operator for a finite angle rotation is given by

$$U_{\mathbf{n}} = \exp\left(\frac{i}{2}(\mathbf{n}\cdot\sigma)\varphi\right), \tag{3.158}$$

or, in another form:

$$U_{\mathbf{n}} = \cos\frac{\varphi}{2} + i\,(\mathbf{n}\cdot\boldsymbol{\sigma})\sin\frac{\varphi}{2}. \tag{3.159}$$

Then, for rotation around z-axis we have

$$U_z(\varphi) = \cos\frac{\varphi}{2} + i\,\sigma_z \sin\frac{\varphi}{2} = \begin{pmatrix} e^{i\varphi/2} & 0 \\ 0 & e^{-i\varphi/2} \end{pmatrix}, \tag{3.160}$$

so that

$$\psi'^{1} = \psi^{1} e^{i\varphi/2}, \qquad \psi'^{2} = \psi^{2} e^{-i\varphi/2}. \tag{3.161}$$

Now we can observe an unusual property of a spinor of the first rank – *under the rotation by angle* 2π *its components change sign* (nonclassical behavior). A similar property is characteristic for all spinors of odd rank.

For rotations around the x-axis and y-axis in a similar way we obtain

$$U_x(\varphi) = \cos\frac{\varphi}{2} + i\,\sigma_x \sin\frac{\varphi}{2} = \begin{pmatrix} \cos\frac{\varphi}{2} & i\sin\frac{\varphi}{2} \\ i\sin\frac{\varphi}{2} & \cos\frac{\varphi}{2} \end{pmatrix}, \tag{3.162}$$

$$U_y(\varphi) = \cos\frac{\varphi}{2} + i\,\sigma_y \sin\frac{\varphi}{2} = \begin{pmatrix} \cos\frac{\varphi}{2} & \sin\frac{\varphi}{2} \\ -\sin\frac{\varphi}{2} & \cos\frac{\varphi}{2} \end{pmatrix}. \tag{3.163}$$

Spin properties of wave functions for a particle with spin s and the system of $n = 2s$ particles with spin $s = 1/2$, oriented to obtain the total spin $2s$, are identical. The number of independent components of symmetric spinor of rank $2s$ is equal to $2s + 1$, as only those of its components are different which contain $2s$ indices equal to 1 and 0 indices 2, $2s - 1$ indices equal to 1 and one index equal to 2, . . . , 0 indices equal to 1 and $2s$ indices equal to 2. As we noted above, symmetric spinors are transformed via irreducible representations of the rotation group.

In particular, spinors of *even* rank are transformed as *tensors* of the rank which is half of that of spinors. Components of these tensors can be explicitly expressed via corresponding components of these spinors. As an important example we present in explicit form the relation between the components of a second rank spinor and corresponding vector [35]:

$$\psi_{12} = \frac{i}{\sqrt{2}} a_z\,, \quad \psi_{11} = -\frac{i}{\sqrt{2}}(a_x + i a_y)\,, \quad \psi_{22} = \frac{i}{\sqrt{2}}(a_x - i a_y)\,, \tag{3.164}$$

$$\psi^{12} = -\frac{i}{\sqrt{2}} a_z\,, \quad \psi^{11} = \frac{i}{\sqrt{2}}(a_x - i a_y)\,, \quad \psi^{22} = -\frac{i}{\sqrt{2}}(a_x + i a_y) \tag{3.165}$$

and

$$a_z = i\sqrt{2}\psi^{12} = \frac{i}{\sqrt{2}}(\psi^{12} + \psi^{21})\,, \quad a_x = \frac{i}{\sqrt{2}}(\psi^{22} - \psi^{11})\,,$$
$$a_y = -\frac{1}{\sqrt{2}}(\psi^{11} + \psi^{22})\,. \tag{3.166}$$

Using Pauli matrices, these relations can be rewritten in a more transparent and compact form:

$$\psi^{\mu}_{\lambda} = -\frac{i}{\sqrt{2}} \mathbf{a} \cdot \sigma^{\mu}_{\lambda}\,, \tag{3.167}$$

$$\mathbf{a} = \frac{i}{\sqrt{2}} \sigma^{\lambda}_{\mu} \psi^{\mu}_{\lambda}\,. \tag{3.168}$$

The scalar product of two vectors can be directly expressed via the scalar product of corresponding spinors as

$$\mathbf{a} \cdot \mathbf{b} = \psi_{\lambda\mu} \varphi^{\lambda\mu}\,. \tag{3.169}$$

3.3.2 Spinors of the Lorentz group

Thus, in nonrelativistic theory, a particle with spin s is described by a $(2s + 1)$-component symmetric spinor of rank $2s$, i. e., by a mathematical object which is transformed according to the corresponding irreducible representation of rotation group $SU(2)$. The rotation group is a subgroup of the Lorentz group (rotation group in four-dimensional space-time). Let us limit ourselves to a proper Lorentz group (without spatial inversions). The theory of four-dimensional spinors is constructed similarly to the theory of spinors in three dimensions.

Spinor ξ^{α} is a two-component object, and $\alpha = 1, 2$, in correspondence with two spin projections $s = \pm 1/2$. Under the action of an arbitrary Lorentz transformation, spinor components are transformed via each other as (binary transformations):

$$\xi'^{1} = \alpha\xi^{1} + \beta\xi^{2}\,, \quad \xi'^{2} = \gamma\xi^{1} + \delta\xi^{2}\,, \tag{3.170}$$

where complex coefficients $\alpha, \beta, \gamma, \delta$ are determined by the rotation angles of four-dimensional coordinate system and satisfy the condition

$$\alpha\delta - \beta\gamma = 1\,, \tag{3.171}$$

so that the determinant of transformation (3.170) is equal to 1. Thus, there is a limitation, determined by two equations for four complex coefficients, so that there remain $8 - 2 = 6$ independent real transformation parameters, corresponding to the number of rotation angles of a coordinate system in four-dimensional space-time (rotations in six coordinate planes).

Due to (3.171) transformations (3.170) leave invariant the following bilinear form:

$$\xi^1\Xi^2 - \xi^2\Xi^1\,, \tag{3.172}$$

which is constructed from the components of two spinors ξ^α and Ξ^α, which corresponds to a scalar particle with spin $s = 0$, composed of two particles with spin $s = 1/2$.

Besides contravariant spinors ξ^α, we can also introduce covariant spinors ξ_α as

$$\xi_\alpha = g_{\alpha\beta}\xi^\beta\,, \tag{3.173}$$

where the "metric tensor" $g_{\alpha\beta}$ has the same form as (3.155):

$$g_{\alpha\beta} = g^{\lambda\mu} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{3.174}$$

so that

$$\xi_1 = \xi^2\,, \quad \xi_2 = -\xi^1\,, \tag{3.175}$$

$$\xi^1\Xi^2 - \xi^2\Xi^1 = \xi^\alpha\Xi_\alpha = -\xi_\alpha\Xi^\alpha\,. \tag{3.176}$$

Up to now all the expressions are the same as in nonrelativistic theory. The difference appears when we consider complex conjugate spinors. In nonrelativistic theory the sum $\psi^1\psi^{1*} + \psi^2\psi^{2*}$, determining the probability density of particle localization in space, is scalar. Thus, the components $\psi^{\alpha*}$ are to be transformed as covariant components of a spinor. The corresponding transformation (3.141), as we have seen, is unitary. In relativistic theory particle density is not a scalar, but the time-component of a 4-vector, so that there are no limitations on coefficients of transformation (3.170), except (3.171). Thus, in relativistic theory, complex conjugate transformations of spinors are essentially different. Correspondingly, here we are dealing with two types of spinors. The indices of the spinors, transformed by complex conjugate expressions (3.170), will be supplied by additional dots (dotted indices).

By definition we have $\eta^{\dot\alpha} \sim \xi^{\alpha*}$ (here $\sim$ means "transformed as") and the rule of transformation for spinors with dotted indices is written as

$$\eta'^{\dot1} = \alpha^*\eta^{\dot1} + \beta^*\eta^{\dot2}\,, \quad \eta'^{\dot2} = \gamma^*\eta^{\dot1} + \delta^*\eta^{\dot2}\,. \tag{3.177}$$

Operations of the lowering and lifting of indices are written as usual:

$$\eta_1 = \eta^2\,, \quad \eta_{\dot 2} = -\eta^{\dot 1}\,. \tag{3.178}$$

With respect to three-dimensional rotations, 4-spinors behave as three-dimensional spinors; as we already noted, that rotation group is a subgroup of the Lorentz group. However, for three-dimensional spinors $\psi^*_\alpha \sim \psi^\alpha$. Thus, $\eta_{\dot\alpha}$ under rotations behaves as a contravariant 3-spinor ψ^α.

Spinors of higher rank are defined as objects which are transformed as products of the components of several spinors of rank 1. For example, we can introduce three types of second-rank spinors:

$$\xi^{\alpha\beta} \sim \xi^\alpha \Xi^\beta\,, \quad \zeta^{\alpha\dot\beta} \sim \xi^\alpha \eta^{\dot\beta}\,, \quad \eta^{\dot\alpha\dot\beta} \sim \eta^{\dot\alpha} H^{\dot\beta}\,. \tag{3.179}$$

Accordingly, the rank of a spinor in relativistic theory is denoted by the pair of numbers (k, l), i. e., the number of nondotted and dotted indices.

Contraction of spinors can be performed only over pairs of indices of a similar type (two dotted or two nondotted), as summation over the pair of indices of different types is not an invariant operation. Thus, taking the spinor

$$\zeta^{\alpha_1\alpha_2\ldots\alpha_k\dot\beta_1\dot\beta_2\ldots\dot\beta_l}\,, \tag{3.180}$$

symmetric over all k-dotted and l-nondotted indices, we can not obtain the spinor of the lower rank (contraction over the pair of indices, with respect to which the spinor is symmetric, gives zero, with the account of (3.176)). Thus, symmetric spinors realize irreducible representations of the Lorentz group, and each of these representations is characterized by the pair of numbers (k, l). As each of the spinor indices takes two values, we have $k + 1$ essentially[7] different sets of numbers $\alpha_1, \alpha_2, \ldots, \alpha_k$ in (3.180) (containing $0, 1, 2, \ldots k$ values equal to 1 and $k, k-1, \ldots, 0$ values equal to 2) and $l + 1$ sets of numbers $\dot\beta_1, \dot\beta_2, \ldots, \dot\beta_l$. Accordingly, the symmetric spinor of rank (k, l) has $(k + 1)(l + 1)$ independent components, which defines the dimensionality of the corresponding irreducible representation.

The relationship between spinors and 4-vectors

Spinor $\zeta^{\alpha\dot\beta}$ has $2{\cdot}2 = 4$ components, the same number as the 4-vector a^μ. Both realize the same irreducible representation of the proper Lorentz group, and we have the following linear relations between their components:

$$\begin{aligned} a^1 &= \frac{1}{2}(\zeta^{1\dot 2} + \zeta^{2\dot 1})\,, \qquad a^2 = \frac{i}{2}(\zeta^{1\dot 2} - \zeta^{2\dot 1})\,, \\ a^3 &= \frac{1}{2}(\zeta^{1\dot 1} - \zeta^{2\dot 2})\,, \qquad a^0 = \frac{1}{2}(\zeta^{1\dot 1} + \zeta^{2\dot 2})\,. \end{aligned} \tag{3.181}$$

[7] $\zeta^{\alpha\dot\beta}$ and $\zeta^{\dot\beta\alpha}$ are just the same, as transformations (3.170) and (3.177) are independent.

For spatial components these relations are the same as in the case of the three-dimensional rotation group, taking into account the substitution $\psi_{\beta}^{\alpha} \to \zeta^{\alpha\dot{\beta}}$. The expression for a^0 is obvious from the previous discussion on the probability density of particle localization as a time-component of a 4-vector. The inverse relation has the form

$$\begin{aligned} \zeta^{1\dot{1}} &= \zeta_{22} = a^3 + a^0\,, & \zeta^{2\dot{2}} &= \zeta_{11} = a^0 - a^3\,, \\ \zeta^{1\dot{2}} &= -\zeta_{21} = a^1 - ia^2\,, & \zeta^{2\dot{1}} &= -\zeta_{12} = a^1 + ia^2\,. \end{aligned} \tag{3.182}$$

The coefficients in these expressions are specially chosen for the scalar product to be written as

$$a^2 = \frac{1}{2}\zeta_{\alpha\dot{\beta}}\zeta^{\alpha\dot{\beta}}\,, \quad ab = \frac{1}{2}\zeta_{\alpha\dot{\beta}}\xi^{\alpha\dot{\beta}}\,. \tag{3.183}$$

The correspondence between $\zeta^{\alpha\dot{\beta}}$ and the 4-vector a^{μ} is the special case of the general rule that any symmetric spinor of the rank (k,k) is equivalent to the symmetric irreducible (i. e., becoming zero after contraction over any pair of indices) 4-tensor of the rank k.

Relations between a spinor of rank $(1,1)$ and the 4-vector (3.181), (3.182) can be written in compact form using Pauli matrices:

$$\mathbf{a} = \frac{1}{2}Sp(\hat{\zeta}\sigma) \quad a^0 = \frac{1}{2}Sp\hat{\zeta}\,, \tag{3.184}$$

$$\hat{\zeta} = \mathbf{a}\cdot\sigma + a^0\hat{1}\,, \tag{3.185}$$

where $\hat{\zeta}$ is the $\zeta^{\alpha\dot{\beta}}$ matrix, $\hat{1}$ is the unit matrix.

Let us write the spinor ξ^{α} transformation as

$$\xi'^{\alpha} = (B\xi)^{\alpha}\,, \quad \text{where} \quad B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}. \tag{3.186}$$

Then[8]

$$\eta'^{\dot{\beta}} = (B^*\eta)^{\dot{\beta}} = (\eta B^+)^{\dot{\beta}}\,. \tag{3.187}$$

Then, transformation of the spinor of the rank $(1,1)$ is written as

$$\zeta' = B\zeta B^+\,. \tag{3.188}$$

For infinitesimal transformation we can write $B = 1 + \lambda$, where λ is an infinitesimal matrix. Then, from (3.188) we have

$$\zeta' = \zeta + (\lambda\zeta + \zeta\lambda^+)\,. \tag{3.189}$$

Consider now a Lorentz transformation to the coordinate system, moving with infinitesimal velocity $\delta\mathbf{v}$ (with no change of direction in the spatial axes). Under this

[8] For covariant components we have $\xi'_{\alpha} = (\tilde{B}^{-1}\xi)_{\alpha} = (\xi B^{-1})_{\alpha}$, $\eta'_{\dot{\alpha}} = (\eta B^{*-1})_{\alpha}$, so that the scalar product of spinors remains invariant.

transformation 4-vector $a^\mu = (a^0, \mathbf{a})$ is transformed as

$$\mathbf{a}' = \mathbf{a} - a^0 \delta\mathbf{v}\,, \quad a'^0 = a^0 - \mathbf{a}\cdot\delta\mathbf{v}\,. \tag{3.190}$$

Let us now use equation (3.184). First of all,

$$a'^0 = a^0 - \mathbf{a}\delta\mathbf{v} = a^0 - \frac{1}{2}Sp(\zeta\sigma\delta\mathbf{v})\,, \tag{3.191}$$

On the other hand,

$$a'^0 = \frac{1}{2}Sp\zeta' = a^0 + \frac{1}{2}Sp(\lambda\zeta + \zeta\lambda^+) = a^0 + \frac{1}{2}Sp\zeta(\lambda + \lambda^+)\,. \tag{3.192}$$

Comparing (3.191) and (3.192) we get

$$\lambda + \lambda^+ = -\sigma\delta\mathbf{v}\,. \tag{3.193}$$

Similarly, considering the transformation of $\mathbf{a}$, we obtain

$$\sigma\lambda + \lambda^+\sigma = -\delta\mathbf{v}\,. \tag{3.194}$$

Now, equations (3.193), (3.194) give

$$\lambda = \lambda^+ = -\frac{1}{2}\sigma\cdot\delta\mathbf{v}\,, \tag{3.195}$$

so that the infinitesimal Lorentz transformation of spinor ξ^α is done by the matrix:

$$B = 1 - \frac{1}{2}(\sigma\cdot\mathbf{n})\delta v \tag{3.196}$$

where $\mathbf{n} = \delta\mathbf{v}/\delta v$. Now we can consider finite transformations. Lorentz transformation (to the coordinate system, moving with velocity $\mathbf{v}$) has the geometrical meaning of a rotation of a four-dimensional coordinate system by angle φ in the $(t, \mathbf{n})$-plane, where φ is determined by velocity v: $v = th\varphi$ [33]. Infinitesimal transformation corresponds to the angle $\delta\varphi = \delta v$, and rotation by the finite angle can be achieved by making the $\delta\varphi$-rotation $\varphi/\delta\varphi$ times. Raising (3.196) to the power $\varphi/\delta\varphi$ and going to the limit $\delta\varphi \to 0$, we obtain

$$B = \exp\left(-\frac{\varphi}{2}\mathbf{n}\cdot\sigma\right)\,. \tag{3.197}$$

Taking into account that even powers of $\mathbf{n}\cdot\sigma$ are equal to 1, while odd powers reduce to $\mathbf{n}\cdot\sigma$, we get

$$B = ch\frac{\varphi}{2} - \mathbf{n}\cdot\sigma sh\frac{\varphi}{2}\,, \quad th\varphi = v\,. \tag{3.198}$$

Note that the transformation matrix $B = B^+$ is Hermitian. Equation (3.198) finally determines the Lorentz transformation of a 4-spinor of first rank.

Consider now an infinitesimal rotation of some vector in three-dimensional space:

$$\mathbf{a}' = \mathbf{a} - [\delta\boldsymbol{\theta}\times\mathbf{a}]\,. \tag{3.199}$$

In this case we obtain

$$B = 1 + \frac{i}{2}\sigma \cdot \delta\boldsymbol{\theta}\,, \tag{3.200}$$

while for the finite angle rotation

$$B = \exp\left(i\frac{\theta}{2}\mathbf{n}\cdot\sigma\right) = \cos\frac{\theta}{2} + i\,\mathbf{n}\cdot\sigma\,\sin\frac{\theta}{2}\,, \tag{3.201}$$

where $\mathbf{n}$ determines the direction of the rotation axis. This matrix is unitary $B^+ = B^{-1}$, as it should be for spatial rotation.

Inversion of spinors (P-reflection)

In nonrelativistic quantum mechanics, spatial inversion does not change the sign of an axial vector such as spin. Thus, its s^z-projection also does not change. It follows then that under inversion each component of the three-dimensional spinor ψ^α transforms only via itself:

$$\psi^\alpha \to P\psi^\alpha \tag{3.202}$$

Making inversion twice, we return to the initial coordinate system. In the case of spinors, the return to the initial coordinates can be understood as a rotation by angle 0, or like a rotation by angle 2π. However, we have seen that for spinors these two operations are not the same, since according to (3.161) spinor components ψ^α change sign under rotation by 2π. Thus, we obtain two alternatives:

$$P^2 = 1\,, \quad \text{i.e.,} \quad P = \pm 1\,, \tag{3.203}$$

$$P^2 = -1\,, \quad \text{i.e.,} \quad P = \pm i\,. \tag{3.204}$$

Consider now 4-spinors. Inversion commutes with spatial rotations, as it only changes the signs of x, y, z in x, y, z, t, but does not commute with transformations dealing with the t-axis. Consider the Lorentz transformation L to a system, moving with velocity $\mathbf{v}$, then $PL = L'P$, where L' is the transformation to system, moving with velocity $-\mathbf{v}$. Thus, under the inversion components of a 4-spinor ξ^α can not transform via each other, and inversion transforms ξ^α via some other objects, which may be only $\eta^{\dot\alpha}$. As inversion does not change the sign of s_z, the components ξ^1 and ξ^2 can be transformed only into η_1 and η_2, corresponding to the same values of $s_z = +1/2$ and $s_z = -1/2$. Understanding inversion as an operation giving 1 being applied twice, we define it by

$$\begin{aligned} &\xi^\alpha \to \eta_{\dot\alpha}\,, \qquad \eta_{\dot\alpha} \to \xi^\alpha\,, \\ &\xi_\alpha \to -\eta^{\dot\alpha}\,, \quad \eta^{\dot\alpha} \to -\xi^\alpha \end{aligned} \tag{3.205}$$

for the case of $P^2 = 1$. For the alternative variant of $P^2 = -1$ we can write:

$$\begin{aligned} &\xi^\alpha \to i\,\eta_{\dot\alpha} \qquad \eta_{\dot\alpha} \to i\,\xi^\alpha \\ &\xi_\alpha \to -i\,\eta^{\dot\alpha} \quad \eta^{\dot\alpha} \to -i\,\xi^\alpha \end{aligned} \tag{3.206}$$

The different sign in the second row of these expressions is connected with the fact that the lowering or raising of the same index, according to (3.175), (3.178), is performed with different signs. Below, for definiteness, we shall use the definition (3.206).

With respect to the subgroup of rotations, as we have seen above, ξ^α and $\eta_{\dot\alpha}$ are transformed in the same way. Let us construct the following combinations:

$$\xi^\alpha \pm \eta_{\dot\alpha}\,. \tag{3.207}$$

It is easily seen that these combinations are transformed under inversion via each other, as (3.202) with $P = \pm i$. However, these combinations do not behave as spinors with respect to all transformations of the Lorentz group.

Thus, the inclusion of inversion into our group of symmetry requires the simultaneous consideration of the pair of spinors $(\xi^\alpha, \eta_{\dot\alpha})$ – the so-called *bispinor*. Four components of the bispinor realize one of the irreducible representations of the extended Lorentz group. The scalar product of two bispinors can be constructed in two different ways. The value of

$$\xi^\alpha \Xi_\alpha + \eta_{\dot\alpha} H^{\dot\alpha} \tag{3.208}$$

does not change under inversion and defines the true scalar. The value of

$$\xi^\alpha \Xi_\alpha - \eta_{\dot\alpha} H^{\dot\alpha} \tag{3.209}$$

is also invariant with respect to rotations of the four-dimensional coordinate system, but it changes its sign under inversion, defining the *pseudoscalar*.

Also in two ways, from the components of two bispinors we can define the corresponding spinor of the second rank $\zeta^{\alpha\dot\beta}$. Defining it as

$$\zeta^{\alpha\dot\beta} \sim \xi^\alpha H^{\dot\beta} + \Xi^\alpha \eta^{\dot\beta}\,, \tag{3.210}$$

we obtain the object, transforming under inversion as $\zeta^{\alpha\dot\beta} \to \zeta_{\dot\alpha\beta}$, so that the 4-vector equivalent to this spinor is transformed as $(a^0, \mathbf{a}) \to (a^0, -\mathbf{a})$ and represents the true 4-vector (here $\mathbf{a}$ is the polar vector). But we can also define $\zeta^{\alpha\dot\beta}$ in another form:

$$\zeta^{\alpha\dot\beta} \sim \xi^\alpha H^{\dot\beta} - \Xi^\alpha \eta^{\dot\beta}\,. \tag{3.211}$$

Then, under inversion $\zeta^{\alpha\dot\beta} \to -\zeta_{\dot\alpha\beta}$, and this spinor corresponds to the 4-vector, transformed under inversion as $(a^0, \mathbf{a}) \to (-a^0, \mathbf{a})$, i. e., 4-pseudovector (here $\mathbf{a}$ is the axial vector).

3.3.3 The Dirac equation

A particle with spin 1/2 in the rest system is described by a two-component wave function, a three-dimensional spinor. By it four-dimensional "origin" this may be both nondotted or nondotted 4-spinor: ξ^α or $\eta_{\dot\alpha}$. The only operator entering the wave equation is $p_\mu = i\,\partial_\mu$, which in spinor representation is expressed via $p_{\alpha\dot\beta}$:

$$\begin{aligned} p^{1\dot1} &= p_{2\dot2} = p_z + p_0\,, \qquad p^{2\dot2} = p_{1\dot1} = p_0 - p_z\,, \\ p^{1\dot2} &= -p_{2\dot1} = p_x - i p_y\,, \quad p^{2\dot1} = -p_{1\dot2} = p_x + i p_y\,. \end{aligned} \tag{3.212}$$

From the requirement of relativistic invariance we can immediately write the following system of first order differential equations:

$$\begin{aligned} p^{\alpha\dot\beta}\eta_{\dot\beta} &= m\xi^\alpha\,, \\ p_{\dot\beta\alpha}\xi^\alpha &= m\eta_{\dot\beta}\,, \end{aligned} \tag{3.213}$$

which is the system of Dirac equations in spinor representation.

Substituting $\eta_{\dot\beta}$ from the second equation of (3.213) into the first, we get

$$p^{\alpha\dot\beta}\eta_{\dot\beta} = \frac{1}{m}p^{\alpha\dot\beta}p_{\gamma\dot\beta}\xi^\gamma = m\xi^\alpha\,. \tag{3.214}$$

Taking into account $p^{\alpha\dot\beta}p_{\gamma\dot\beta} = p^2\delta^\alpha_\gamma$, we obtain from (3.214)

$$(p^2 - m^2)\xi^\gamma = 0 \tag{3.215}$$

i. e., a Klein–Gordon equation for each of the spinor components. It is clear now that parameter m is just the particle mass. Note that only the presence of mass requires the simultaneous introduction of *two* spinors ξ^α and $\eta_{\dot\beta}$, i. e., the bispinor, or we would not be able to construct relativistically invariant equations containing the dimensional parameter m. As a result, our wave equation is automatically invariant with respect to spatial inversion, if we define it by (cf. (3.206))

$$P: \qquad \xi^\alpha \to i\eta_{\dot\alpha}\,, \quad \eta_{\dot\alpha} \to i\xi^\alpha\,. \tag{3.216}$$

Simultaneously, $p^{\dot\alpha\beta} \to p_{\alpha\dot\beta}$ in equations (3.213).

With the help of (3.185) and (3.182), equations (3.213) can be written as

$$\begin{aligned} (p_0 + \mathbf{p}\boldsymbol{\sigma})\eta &= m\xi\,, \\ (p_0 - \mathbf{p}\boldsymbol{\sigma})\xi &= m\eta\,, \end{aligned} \tag{3.217}$$

where we have introduced the following columns:

$$\xi = \begin{pmatrix} \xi^1 \\ \xi^2 \end{pmatrix}\,, \quad \eta = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}\,. \tag{3.218}$$

For complex conjugate equations it is convenient to introduce rows:

$$\xi^* = (\xi^{1*}, \xi^{2*}), \qquad \eta^* = (\eta_1^*, \eta_2^*) \tag{3.219}$$

and write (taking into account $p_\mu^* = -p_\mu$)

$$\begin{aligned} \eta^*(p_0 + \mathbf{p}\boldsymbol{\sigma}) &= -m\xi^*, \\ \xi^*(p_0 - \mathbf{p}\boldsymbol{\sigma}) &= -m\eta^*. \end{aligned} \tag{3.220}$$

The inversion for complex conjugate spinors can be written as:

$$P: \quad \xi^{\alpha *} \to -i\eta_{\dot\alpha}^*, \quad \eta_{\dot\alpha}^* \to -i\xi^{\alpha *}. \tag{3.221}$$

In the literature, it is more common to use (instead of (3.213) or (3.217)) the so-called symmetric form of the Dirac equation. To obtain it, we introduce the four-component Dirac bispinor, which is constructed from the columns of (3.218):

$$\psi = \begin{pmatrix} \xi \\ \eta \end{pmatrix}. \tag{3.222}$$

Then, the system of equations (3.217) can be written as

$$p_\mu \gamma_{ik}^\mu \psi_k = m\psi_i \tag{3.223}$$

or, lowering bispinor indices, as

$$(\gamma^\mu p_\mu - m)\psi = 0, \qquad \text{i.e.,} \qquad (i\gamma^\mu \partial_\mu - m)\psi = 0, \tag{3.224}$$

where $\gamma^\mu p_\mu = p_0\gamma^0 - \mathbf{p}\cdot\boldsymbol{\gamma} = i\gamma^0\partial_0 + i\boldsymbol{\gamma}\cdot\nabla$, and we have introduced matrices 4×4 (Dirac matrices):

$$\gamma^0 = \begin{pmatrix} 0 & \hat 1 \\ \hat 1 & 0 \end{pmatrix}, \quad \boldsymbol{\gamma} = \begin{pmatrix} 0 & -\hat\sigma \\ \hat\sigma & 0 \end{pmatrix}. \tag{3.225}$$

In fact, equation (3.217) can be written as

$$\begin{pmatrix} 0 & p_0 + \mathbf{p}\boldsymbol{\sigma} \\ p_0 - \mathbf{p}\boldsymbol{\sigma} & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = m \begin{pmatrix} \xi \\ \eta \end{pmatrix}, \tag{3.226}$$

which coincides with (3.224) if we take γ-matrices as in (3.225).

In the general case γ-matrices should satisfy conditions guaranteeing the identity $p^2 = m^2$. To derive these conditions we multiply (3.224) on the left side by $\gamma^\nu p_\nu$. Then,

$$(\gamma^\nu p_\nu)(\gamma^\mu p_\mu)\psi = m(\gamma^\mu p_\mu)\psi = m^2\psi. \tag{3.227}$$

As $p_\mu p_\nu$ is a symmetric tensor (momentum components commute!), equation (3.227) can be rewritten as

$$\frac{1}{2} p_\mu p_\nu (\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu)\psi = m^2\psi, \tag{3.228}$$

so that the necessary condition is satisfied if

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}\,. \tag{3.229}$$

Thus, the pairs of different matrices γ^μ anticommute, while their squares are

$$(\gamma^1)^2 = (\gamma^2)^2 = (\gamma^3)^2 = -1\,, \quad (\gamma^0)^2 = 1\,. \tag{3.230}$$

Under an arbitrary unitary transformation of bispinor $\psi' = U\psi$ (where U is the unitary matrix 4×4) γ-matrices transform as

$$\gamma' = U\gamma U^{-1} = U\gamma U^+\,, \tag{3.231}$$

so that $(\gamma_\mu p^\mu - m)\psi = 0$ is transformed into $(\gamma'^\mu p_\mu - m)\psi' = 0$. Under this transformation, as it is obvious from (3.225), the following properties are conserved:

$$\boldsymbol{\gamma}^+ = -\boldsymbol{\gamma}\,, \qquad \gamma^{0+} = \gamma^0\,. \tag{3.232}$$

The complex conjugate of equation (3.224) can be written as

$$(-p_0\tilde{\gamma}_0 - \mathbf{p}\tilde{\boldsymbol{\gamma}} - m)\psi^* = 0\,. \tag{3.233}$$

Using $\tilde{\gamma}^\mu\psi^* = \psi^*\gamma^\mu$ and multiplying this equation from the right side by γ^0 (and taking into account $\boldsymbol{\gamma}\gamma^0 = -\gamma^0\boldsymbol{\gamma}$), we obtain a conjugate Dirac's equation as

$$\bar{\psi}(\gamma^\mu p_\mu + m) = 0\,, \tag{3.234}$$

where we have introduced

$$\bar{\psi} = \psi^*\gamma^0, \quad \psi^* = \bar{\psi}\gamma^0 \tag{3.235}$$

– Dirac's conjugate of bispinor ψ.

It is easy to see that the Dirac's equation (3.224)

$$(i\gamma^\mu\partial_\mu - m)\psi = 0 \tag{3.236}$$

can be obtained from the Euler–Lagrange equation

$$\frac{\partial\mathcal{L}}{\partial\bar{\psi}} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\bar{\psi})}\right) = 0 \tag{3.237}$$

using the following Lagrangian of Dirac's field:

$$\mathcal{L} = \frac{i}{2}[\bar{\psi}\gamma^\mu(\partial_\mu\psi) - (\partial_\mu\bar{\psi})\gamma^\mu\psi] - m\bar{\psi}\psi \equiv i\bar{\psi}\gamma^\mu \overleftrightarrow{\partial}_\mu \psi - m\bar{\psi}\psi\,, \tag{3.238}$$

where $\overleftrightarrow{\partial}_\mu$ denotes differentiation "to the right" and "to the left", defined by the given identity. In Euler–Lagrange equations, $\bar{\psi}$ and ψ are considered as independent fields.

The conjugate Dirac equation (3.234) is obtained from equation (3.237) after the replacement $\bar{\psi} \to \psi$. Then we immediately find the canonical momentum $\pi(x)$ of Dirac's field as

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\psi}(x)} = i\psi^+(x)\,. \tag{3.239}$$

Then the Hamiltonian density of Dirac's field is written as

$$\mathcal{H} = \pi\dot{\psi} - \mathcal{L} = \psi^+\gamma^0(-i\gamma^i\partial_i + m)\psi = \psi^+\gamma^0(i\gamma^0\partial_0\psi) = \psi^+ i\frac{\partial\psi}{\partial t}\,, \tag{3.240}$$

where in the second equality we used the Dirac equation (3.224).

Remark on dimensionalities

Using the explicit form of Dirac's field Lagrangian (3.238) and the standard dimensionalities $[\mathcal{L}] = l^{-4}$, $[m] = l^{-1}$, $[\partial] = l^{-1}$, we immediately determine the dimensionality of Dirac's field as

$$[\psi] = [\bar{\psi}] = l^{-3/2} \tag{3.241}$$

This result will be used below.

The inversion (3.216) for ψ can be written as

$$P: \qquad \psi \to i\gamma^0\psi\,, \quad \bar{\psi} \to -i\bar{\psi}\gamma^0\,. \tag{3.242}$$

The invariance of the Dirac equation with respect to (3.242) is obvious. Replacing $\mathbf{p} \to -\mathbf{p}$ and $\psi \to i\gamma^0\psi$, we get $(p_0\gamma^0 + \mathbf{p}\boldsymbol{\gamma} - m)\gamma^0\psi = 0$, so that multiplying this equation on the left side by γ^0 and taking into account anticommutativity of γ^0 and $\boldsymbol{\gamma}$ we return to initial equation.

Let us multiply $(\gamma^\mu p_\mu - m)\psi = 0$ on the left side by $\bar{\psi}$, and $\bar{\psi}(\gamma^\mu p_\mu + m) = 0$ on the right by ψ, then make the sum of both and obtain

$$\bar{\psi}\gamma^\mu(p_\mu\psi) + (p_\mu\bar{\psi})\gamma^\mu\psi = p_\mu(\bar{\psi}\gamma^\mu\psi) = 0 \tag{3.243}$$

which is the continuity equation for 4-current of Dirac's particles:

$$\partial_\mu j^\mu = 0\,, \qquad j^\mu = \bar{\psi}\gamma^\mu\psi = (\psi^*\psi, \psi^*\gamma^0\boldsymbol{\gamma}\psi)\,, \tag{3.244}$$

describing the charge conservation, with the charge density given by $j^0 = \psi^*\psi > 0$.

The Dirac equation can be written in the form of a Schroedinger equation:

$$i\frac{\partial\psi}{\partial t} = H\psi\,, \tag{3.245}$$

where the Hamiltonian H has the form

$$H = \boldsymbol{\alpha}\mathbf{p} + \beta m\,, \tag{3.246}$$

with Dirac's matrices $\boldsymbol{\alpha}$ and β:

$$\boldsymbol{\alpha} = \gamma^0 \boldsymbol{\gamma}\,, \qquad \beta = \gamma^0\,, \tag{3.247}$$

so that (3.246) coincides with (3.240) introduced above. The matrices (3.247) satisfy the commutation relations

$$\alpha_i \alpha_k + \alpha_k \alpha_i = 2\delta_{ik}\,, \quad \beta\boldsymbol{\alpha} + \boldsymbol{\alpha}\beta = 0\,, \quad \beta^2 = 1\,, \tag{3.248}$$

and in explicit form

$$\boldsymbol{\alpha} = \begin{pmatrix} \sigma & 0 \\ 0 & -\sigma \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.249}$$

Consider the nonrelativistic limit. Performing in equation (3.217) the limit of $\mathbf{p} \to 0$, $\varepsilon \to m$, we get $\xi = \eta$, so that both spinors of the bispinor coincide, and all four components of the bispinor are nonzero. At the same time, it is clear that only two components are independent. It is convenient to transform to the so-called standard representation, when in the nonrelativistic limit two components of the bispinor will be zero. Let us introduce

$$\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}, \quad \varphi = \frac{1}{\sqrt{2}}(\xi + \eta)\,, \quad \chi = \frac{1}{\sqrt{2}}(\xi - \eta)\,. \tag{3.250}$$

For the particle at rest we obviously have $\chi = 0$. Adding and subtracting equations (3.217) we obtain

$$\begin{aligned} p_0\varphi - \mathbf{p}\sigma\chi &= m\varphi\,, \\ -p_0\chi + \mathbf{p}\sigma\varphi &= m\chi\,, \end{aligned} \tag{3.251}$$

which allows to find the explicit form of γ-matrices in the standard representation [6]. Note that in equation (3.250) we separately sum the first and the second components of spinors ξ and η. Accordingly, in standard representation, as in the spinor representation considered above, ψ_1, ψ_3 corresponds to spin projection $s^z = +1/2$, while ψ_2, ψ_4 to projection $s^z = -1/2$. Matrix

$$\frac{1}{2}\Sigma = \frac{1}{2}\begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix} \tag{3.252}$$

gives the three-dimensional operator of the spin in standard representation.

Helicity

In relativistic theory, the orbital moment $\mathbf{l}$ and spin $\mathbf{s}$ of a moving particle are not conserved separately. Only the total angular moment $\mathbf{j} = \mathbf{l} + \mathbf{s}$ is conserved. Accordingly, the projection of the spin on some direction (z-axis) is also not conserved and cannot be used to classify polarization (spin) states of a moving particles. However, we may introduce the *helicity* of a particle, i. e., the projection of its spin on the direction of motion (momentum). In fact, $\mathbf{l} = [\mathbf{r} \times \mathbf{p}]$ and

the product $\mathbf{s}\cdot\mathbf{n}$, where $\mathbf{n} = \frac{\mathbf{p}}{|\mathbf{p}|}$, coincides with the conserving product $\mathbf{j}\cdot\mathbf{n}$. The eigenvalues of these spin projections are obviously given by $\lambda = -s, \ldots, +s$. Accordingly, the wave functions of a free particle with momentum $\mathbf{p}$ are characterized by helicity: $\psi_{\mathbf{p}\lambda}$. In the rest system the state of a particle is characterized, as usual, by its spin (projection on z-axis).

For a particle with zero mass there is no rest system of coordinates; this particle moves with the speed of light in any coordinate system. However, for such a particle there is always the special direction in space – the direction of momentum $\mathbf{p}$. In this case there is no symmetry with respect to arbitrary three-dimensional rotations, but only the axial symmetry to rotations around this preferred direction. Accordingly, we have only helicity conservation. If we require symmetry with respect to reflections in planes, passing through the $\mathbf{p}$-axis, the states differing by the sign of λ will be degenerate, and for $\lambda \neq 0$ we have double degeneracy. Thus, in the limit of $m \to 0$ the system of equations for a particle with spin s splits into independent equations for particles with different helicities $\pm s, \pm(s-1), \ldots$: for example, in the case of photon $\lambda = \pm 1$, which corresponds to right and left polarizations of light.

3.3.4 The algebra of Dirac's matrices

For practical calculations it is important to understand algebraic properties of γ-matrices. Here we present a short summary of the main formulas and definitions which will be extensively used in the rest of the book. All the algebraic properties of Dirac's matrices are derived from two basic relations:

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}\,, \tag{3.253}$$

$$g_{\mu\nu}\gamma^\mu\gamma^\nu = \gamma_\mu\gamma^\mu = 4 \quad \text{or} \quad \gamma_0^2 - \gamma_1^2 - \gamma_2^2 - \gamma_3^2 = 4\,, \tag{3.254}$$

i. e., from the main anticommutation relation and the scalar product.

If γ_μ and γ^μ in the matrix product are separated by several other γ-matrices, γ_μ and γ^μ can be placed in neighboring positions with the help of (3.253), after which we can perform the summation over μ using (3.254). In this way we can get the following relations:

$$\begin{aligned}
\gamma_\mu\gamma^\nu\gamma^\mu &= -2\gamma^\nu\,,\\
\gamma_\mu\gamma^\lambda\gamma^\nu\gamma^\mu &= 4g^{\lambda\nu}\,,\\
\gamma_\mu\gamma^\lambda\gamma^\nu\gamma^\rho\gamma^\mu &= -2\gamma^\rho\gamma^\nu\gamma^\lambda\,,\\
\gamma_\mu\gamma^\lambda\gamma^\nu\gamma^\rho\gamma^\sigma\gamma^\mu &= 2(\gamma^\sigma\gamma^\lambda\gamma^\nu\gamma^\rho + \gamma^\rho\gamma^\nu\gamma^\lambda\gamma^\sigma)\,.
\end{aligned} \tag{3.255}$$

Rather often γ^μ appears in combination with 4-vectors. Let us introduce the standard notation:

$$\hat{a} \equiv \gamma^\mu a_\mu\,. \tag{3.256}$$

Then, from (3.253) we obtain

$$\hat{a}\hat{b} + \hat{b}\hat{a} = 2a_\mu b^\mu\,, \quad \hat{a}\hat{a} = a^2\,, \tag{3.257}$$

and from (3.255) we get

$$\begin{aligned}
\gamma_\mu \hat{a}\gamma^\mu &= -2\hat{a}\,,\\
\gamma_\mu \hat{a}\hat{b}\gamma^\mu &= 4a_\mu b^\mu\,,\\
\gamma_\mu \hat{a}\hat{b}\hat{c}\gamma^\mu &= -2\hat{c}\hat{b}\hat{a}\,,\\
\gamma_\mu \hat{a}\hat{b}\hat{c}\hat{d}\gamma^\mu &= 2\big(\hat{d}\hat{a}\hat{b}\hat{c} + \hat{c}\hat{b}\hat{a}\hat{d}\big)\,.
\end{aligned} \tag{3.258}$$

Traces of γ-matrices are widely used. In particular,

$$Sp\gamma^\mu = 0\,. \tag{3.259}$$

Introducing

$$T^{\mu\nu} = \frac{1}{4}Sp(\gamma^\mu\gamma^\nu) \tag{3.260}$$

and calculating the trace of (3.253), we find

$$T^{\mu\nu} = g^{\mu\nu} \tag{3.261}$$

and, respectively,

$$\frac{1}{4}Sp(\hat{a}\hat{b}) = a^\mu b_\mu\,. \tag{3.262}$$

The special case is the matrix γ^5 defined as

$$\gamma^5 = -i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{3.263}$$

It is easy to see that

$$\gamma^5\gamma^\mu + \gamma^\mu\gamma^5 = 0\,, \quad (\gamma^5)^2 = 1\,, \tag{3.264}$$

so that γ^5 anticommutes with all other γ-matrices. As to $\boldsymbol{\alpha}$ and β matrices we have

$$\boldsymbol{\alpha}\gamma^5 - \gamma^5\boldsymbol{\alpha} = 0\,, \quad \beta\gamma^5 + \gamma^5\beta = 0\,. \tag{3.265}$$

The γ^5-matrix is Hermitian:

$$\gamma^{5+} = i\gamma^{3+}\gamma^{2+}\gamma^{1+}\gamma^{0+} = -i\gamma^3\gamma^2\gamma^1\gamma^0 = \gamma^5\,, \tag{3.266}$$

as we can transform the index order 3210 to 0123 by even permutation of γ-matrices.

In spinor representation the explicit form of γ^5 is given by

$$\gamma^5 = \begin{pmatrix} -\hat{1} & 0 \\ 0 & \hat{1} \end{pmatrix}, \tag{3.267}$$

while in standard representation:

$$\gamma^5 = \begin{pmatrix} 0 & -\hat{1} \\ -\hat{1} & 0 \end{pmatrix}, \tag{3.268}$$

and we see that

$$Sp\gamma^5 = 0\,, \tag{3.269}$$

which is obviously independent of the choice of representation.

The set of 16 matrices

$$\{\gamma^A\} = \{\hat{1}, \gamma^5, \gamma^\mu, i\gamma^\mu\gamma^5, i\sigma^{\mu\nu}\}\,, \tag{3.270}$$

where

$$\sigma^{\mu\nu} = \frac{1}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu) \tag{3.271}$$

represents the "complete set", over which we can "expand" any 4×4-matrix. In fact, these matrices have the property:

$$\begin{gathered} Sp\gamma^A = 0 \qquad (A \neq 1)\,, \\ \gamma^A\gamma_A = 1\,, \qquad \frac{1}{4}Sp\gamma^A\gamma_B = \delta^A_B\,. \end{gathered} \tag{3.272}$$

Accordingly, all γ^A-matrices are linearly independent, and any 4×4-matrix can be represented in the form

$$\Gamma = \sum_A c_A\gamma^A\,, \quad c_A = \frac{1}{4}Sp\gamma_A\Gamma\,. \tag{3.273}$$

3.3.5 Plane waves

The state of a free particle with fixed momentum is described by the plane wave, which can be written as

$$\psi_p = \frac{1}{\sqrt{2\varepsilon_{\mathbf{p}}}}u_p e^{-ipx}\,, \tag{3.274}$$

where u_p is a normalized bispinor. For the wave function with "negative frequency" (also changing the sign of $\mathbf{p}$) we have

$$\psi_p = \frac{1}{\sqrt{2\varepsilon_{\mathbf{p}}}}u_{-p} e^{ipx}\,. \tag{3.275}$$

In both cases we write $\varepsilon_{\mathbf{p}} = +\sqrt{\mathbf{p}^2 + m^2}$. The bispinor components u_p and u_{-p} satisfy the following equations, which are derived by substitution of (3.274) and (3.275) into the Dirac equation:

$$(\hat{p} - m)u_p = 0\,, \quad (\hat{p} + m)u_{-p} = 0\,. \tag{3.276}$$

For conjugated bispinors $\bar{u}_p = u_p^*\gamma^0$ we have

$$\bar{u}_p(\hat{p} - m) = 0\,, \quad \bar{u}_{-p}(\hat{p} + m) = 0\,. \tag{3.277}$$

Let us assume the invariant normalization

$$\bar{u}_p u_p = 2m\,, \quad \bar{u}_{-p}u_{-p} = -2m\,. \tag{3.278}$$

Multiplying (3.276) from the left by $\bar{u}_{\pm p}$ we obtain $(\bar{u}_{\pm p}\gamma^\mu u_{\pm p})p_\mu = 2m^2 = 2p^2$, i. e.,

$$\bar{u}_p\gamma^\mu u_p = \bar{u}_{-p}\gamma^\mu u_{-p} = 2p^\mu\,, \tag{3.279}$$

so that the 4-vector of the current density for plane waves (3.274), (3.275) is equal to

$$j^\mu = \bar{\psi}_{\pm p}\gamma^\mu\psi_{\pm p} = \frac{1}{2\varepsilon_{\mathbf{p}}}\bar{u}_{\pm p}\gamma^\mu u_{\pm p} = \frac{p^\mu}{\varepsilon_{\mathbf{p}}}\,, \tag{3.280}$$

i. e., $j^\mu = (1, \mathbf{v})$, where $\mathbf{v} = \frac{\mathbf{p}}{\varepsilon_{\mathbf{p}}}$ is the particle velocity. We see that our choice of normalization corresponds to "one particle in volume $V = 1$".

In standard representation, from (3.251) we get the following system of homogeneous linear equations:

$$\begin{aligned}(\varepsilon_{\mathbf{p}} - m)\varphi - \mathbf{p}\sigma\chi = 0\,,\\ (\varepsilon_{\mathbf{p}} + m)\chi - \mathbf{p}\sigma\varphi = 0\,.\end{aligned} \tag{3.281}$$

Then,

$$\varphi = \frac{\mathbf{p}\sigma}{\varepsilon_{\mathbf{p}} - m}\chi\,, \qquad \chi = \frac{\mathbf{p}\sigma}{\varepsilon_{\mathbf{p}} + m}\varphi\,. \tag{3.282}$$

The common multiplier before φ and χ (which is arbitrary until we are simply dealing with solutions of homogeneous equations) should be chosen from the normalization condition (3.278). Accordingly, in standard representation, spinors u_p and u_{-p} take the form

$$u_p = \begin{pmatrix} \sqrt{\varepsilon_{\mathbf{p}} + m}\, w \\ \sqrt{\varepsilon_{\mathbf{p}} - m}\,(\mathbf{n}\sigma)w \end{pmatrix}, \quad u_{-p} = \begin{pmatrix} \sqrt{\varepsilon_{\mathbf{p}} - m}\,(\mathbf{n}\sigma)w' \\ \sqrt{\varepsilon_{\mathbf{p}} + m}\, w' \end{pmatrix}, \tag{3.283}$$

where $\mathbf{n} = \frac{\mathbf{p}}{|\mathbf{p}|}$, and w is an arbitrary two-component spinor, satisfying the normalization condition

$$w^* w = 1\,. \tag{3.284}$$

The second expression in (3.283) is obtained from the first one by changing the sign before m and replacing $w \to (\mathbf{n}\sigma)w'$. Similarly, we can obtain [6]

$$\begin{aligned}\bar{u}_p &= \left(\sqrt{\varepsilon_{\mathbf{p}} + m}\, w^*,\ -\sqrt{\varepsilon_{\mathbf{p}} - m}\, w^*(\mathbf{n}\sigma)\right),\\ \bar{u}_{-p} &= \left(\sqrt{\varepsilon_{\mathbf{p}} - m}\, w'^*(\mathbf{n}\sigma),\ -\sqrt{\varepsilon_{\mathbf{p}} + m}\, w'^*\right).\end{aligned} \tag{3.285}$$

Direct multiplication gives

$$\bar{u}_{\pm p}u_{\pm p} = \pm 2m\,.$$

In the rest frame, i. e., for $\varepsilon_{\mathbf{p}} = m$, we have

$$u_p = \sqrt{2m}\begin{pmatrix} w \\ 0 \end{pmatrix}, \quad u_{-p} = \sqrt{2m}\begin{pmatrix} 0 \\ w' \end{pmatrix}, \tag{3.286}$$

so that w is the three-dimensional spinor to which our plane waves reduce in the non-relativistic limit:

$$w^{\sigma=1/2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad w^{\sigma=-1/2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.287}$$

For fixed momentum there exist two independent states, corresponding to the two possible values of spin projection. In accordance with the general statements made above, we are speaking of particle helicity λ, i. e., spin projection on the direction of $\mathbf{p}$. The helicity states correspond to the plane waves with spinor $w = w^{(\lambda)}(\mathbf{n})$, which is an eigenfunction of operator $\mathbf{n}\sigma$:

$$\frac{1}{2}(\mathbf{n}\sigma)w^{(\lambda)} = \lambda w^{(\lambda)}. \tag{3.288}$$

3.3.6 Spin and statistics

A second quantization of the Dirac field (of particles with spin $s = 1/2$) can be done as follows. Let us introduce an expansion of an arbitrary Dirac field over plane waves:

$$\begin{aligned} \psi &= \sum_{\mathbf{p}\sigma} \frac{1}{\sqrt{2\varepsilon_{\mathbf{p}}}} \left(a_{\mathbf{p}\sigma} u_{p\sigma} e^{-ipx} + b^{+}_{\mathbf{p}\sigma} u_{-p-\sigma} e^{ipx} \right), \\ \bar{\psi} \equiv \psi^{+}\gamma^{0} &= \sum_{\mathbf{p}\sigma} \frac{1}{\sqrt{2\varepsilon_{\mathbf{p}}}} \left(a^{+}_{\mathbf{p}\sigma} \bar{u}_{p\sigma} e^{ipx} + b_{\mathbf{p}\sigma} \bar{u}_{-p-\sigma} e^{-ipx} \right). \end{aligned} \tag{3.289}$$

We know Dirac's Hamiltonian, so we do not need an energy-momentum tensor. Using (3.240), (3.245), we find the average energy of the Dirac particle in the state with wave function ψ:

$$E = \int d^3\mathbf{r}\psi^* H\psi = i \int d^3\mathbf{r}\psi^* \frac{\partial\psi}{\partial t} = i \int d^3\mathbf{r}\bar{\psi}\gamma^0 \frac{\partial\psi}{\partial t}. \tag{3.290}$$

Substituting here (3.289), taking into account the orthogonality of of the functions with different $\mathbf{p}, \sigma$ and $\bar{u}_{\pm p\sigma}\gamma^0 u_{\pm p,\sigma} = 2\varepsilon_{\mathbf{p}}$ (ñð. (3.280)), we obtain

$$H = \sum_{\mathbf{p}\sigma} \varepsilon_{\mathbf{p}} \left(a^{+}_{\mathbf{p}\sigma} a_{\mathbf{p}\sigma} - b_{\mathbf{p}\sigma} b^{+}_{\mathbf{p}\sigma} \right). \tag{3.291}$$

This expression is obviously the direct consequence of the transformation properties of the Dirac field and the requirement of relativistic invariance. Now it is clear that we must quantize our field using Fermion rules, i. e., introducing *anticommutators*:

$$\{a_{\mathbf{p}\sigma}, a^{+}_{\mathbf{p}\sigma}\} = 1, \quad \{b_{\mathbf{p}\sigma}, b^{+}_{\mathbf{p}\sigma}\} = 1. \tag{3.292}$$

For different indices, as well as for the pairs of "noncrossed" and "crossed" operators, the corresponding anticommutators are equal to zero. As a result, (3.291) is rewritten

in the form

$$H = \sum_{\mathbf{p}\sigma} \varepsilon_{\mathbf{p}} (a^{+}_{\mathbf{p}\sigma} a_{\mathbf{p}\sigma} + b^{+}_{\mathbf{p}\sigma} b_{\mathbf{p}\sigma} - 1)\,, \tag{3.293}$$

so that the eigenvalues of energy (minus the infinite energy of a vacuum) are equal to

$$E = \sum_{\mathbf{p}\sigma} \varepsilon_{\mathbf{p}} (N_{\mathbf{p}\sigma} + \bar{N}_{\mathbf{p}\sigma}) \tag{3.294}$$

and are positive definite. If we had instead used Boson quantization, we would have obtained $E = \sum_{\mathbf{p}} \varepsilon_{\mathbf{p}\sigma} (N_{\mathbf{p}\sigma} - \bar{N}_{\mathbf{p}\sigma})$, i. e., a nonpositive definite expression[9].

For the momentum operator, in a similar way, from $\int d^3\mathbf{r}\psi^* \hat{\mathbf{p}}\psi$ we get

$$\mathbf{P} = \sum_{\mathbf{p}\sigma} \mathbf{p} (N_{\mathbf{p}\sigma} + \bar{N}_{\mathbf{p}\sigma})\,. \tag{3.295}$$

The 4-current density operator $j^{\mu} = \bar{\psi}\gamma^{\mu}\psi$ in a second quantized form defines the charge operator as

$$Q = \int d^3\mathbf{r}\bar{\psi}\gamma^0\psi = \sum_{\mathbf{p}\sigma} \left(a^{+}_{\mathbf{p}\sigma} a_{\mathbf{p}\sigma} + b_{\mathbf{p}\sigma} b^{+}_{\mathbf{p}\sigma}\right) = \sum_{\mathbf{p}\sigma} \left(a^{+}_{\mathbf{p}\sigma} a_{\mathbf{p}\sigma} - b^{+}_{\mathbf{p}\sigma} b_{\mathbf{p}\sigma} + 1\right). \tag{3.296}$$

This gives the eigenvalues of charge:

$$Q = \sum_{\mathbf{p}\sigma} (N_{\mathbf{p}\sigma} - \bar{N}_{\mathbf{p}\sigma})\,, \tag{3.297}$$

leading to the opposite charges of particles and antiparticles.

Fermion anticommutation rules for creation and annihilation operators immediately lead [35] to the validity of the Pauli principle, so that the eigenvalues of the particle number operator in a given state $N_{\mathbf{p}\sigma}$ can only be 0 or 1: now we see that for particles with spin $1/2$ this directly follows from the general requirements of relativistic invariance and positive definiteness of energy. Then, we are coming to the general theorem on spin and statistics: *all particles with a half-integer spin are fermions, while particles with integer spin are bosons*. This becomes obvious if we take into account that any particle with spin s can be assumed to be "composed" of $2s$ particles with spin $1/2$. For the half-integer s the number $2s$ is odd, while for the integer s this number is even. "Composite" particles, consisting of an even number of fermions, is a boson, and those consisting of an odd number of fermions are a fermion. To understand this it is sufficient to consider permutations of such "composite" particles. It is understood, that all particles with the same spin obey the same statistics. If we could have had fermions with spin $s = 0$, then such fermions as well as fermions with spin $1/2$ could be used to "compose" a particle with spin $1/2$, which would be a boson, in contradiction with

[9] Here we are using the standard notations, assuming that the general properties of the creation and annihilation operators are known to the reader [35]

the general results for $s = 1/2$ obtained above. This remarkable theorem is among the most general statements of relativistic quantum field theory and was first proved by Pauli[10].

3.3.7 C, P, T transformations for fermions

Factors of $\psi_{p\sigma} = u_{p\sigma}\exp(-ipx)$, entering (3.289) accompanying operators $a_{\mathbf{p}\sigma}$, represent wave functions of free particles (e. g., electrons) with momentum $\mathbf{p}$ and polarization σ: $\psi^{(e)} = \psi_{p\sigma}$. Factors of $\bar{\psi}_{-p-\sigma}$ accompanying operators $b_{\mathbf{p}\sigma}$ are to be considered as wave functions of the corresponding antiparticles (e. g., positrons) with the same $\mathbf{p}$ and σ. However, $\psi_{p\sigma}$ and $\bar{\psi}_{-p-\sigma}$ differ by their transformation properties, and their components satisfy different systems of equations. To overcome this deficiency we have to perform a certain unitary transformation of $\bar{\psi}_{-p-\sigma}$, such that the new wave function satisfies the same equation as $\psi_{p\sigma}$. We shall call this new wave function the wave function of antiparticles (positron) with momentum $\mathbf{p}$ and polarization σ. Thus, we can write

$$\psi^{(p)}_{p\sigma} = U_C\,\bar{\psi}_{-p-\sigma}\,. \tag{3.298}$$

This operation is called charge conjugation C. It is not limited to plane waves only; in the general case we can write

$$C\,\psi(t,\mathbf{r}) = U_C\,\bar{\psi}(t,\mathbf{r})\,. \tag{3.299}$$

Dropping the details of derivation, which can be found in [6], we quote only the final result:

$$U_C = \gamma^2\gamma^0\,. \tag{3.300}$$

From $\bar{\psi} = \psi^*\gamma^0 = \tilde{\gamma}^0\psi^* = \gamma^0\psi^*$ we obtain

$$C\,\psi = \gamma^2\gamma^0\bar{\psi} = \gamma^2\psi^*\,. \tag{3.301}$$

For plane wave solutions we can easily see that

$$C\,\psi_{-p-\sigma} = \psi_{p\sigma}\,, \tag{3.302}$$

so that both electrons and positrons are described by identical wave functions $\psi^{(e)} = \psi^{(p)} = \psi_{p\sigma}$, as it should be, because these functions carry information only on momentum and polarization of particles.

In a similar way, we may introduce the operation of time inversion. Changing the sign of time should be accompanied by the complex conjugation of the wave function [35]. To obtain the fermion wave function "reversed in time" $T\,\psi$ in the same representation as initial ψ, we again have to perform some unitary transformation of ψ^* (or $\bar{\psi}$):

$$T\,\psi(\mathbf{r},t) = U_T\,\bar{\psi}(\mathbf{r},-t)\,. \tag{3.303}$$

[10] We would like to stress that in quantum field theory this theorem is really *proved*, starting from most general requirements of relativistic invariance (transformation properties of fields) and positive definiteness of energy, i. e., the stability of the ground state of a particle system, and not postulated, as it is done in nonrelativistic quantum mechanics.

It can be shown [6] that

$$U_T = i\gamma^3\gamma^1\gamma^0 , \tag{3.304}$$

so that

$$T\psi(t,\mathbf{r}) = i\gamma^3\gamma^1\gamma^0\bar{\psi}(-t,\mathbf{r}) = i\gamma^3\gamma^1\psi^*(-t,\mathbf{r}) . \tag{3.305}$$

Spatial inversion of (bi)spinors P was defined above in (3.242):

$$P\psi = i\gamma^0\psi , \qquad P\bar{\psi} = -i\bar{\psi}\gamma^0 . \tag{3.306}$$

Let us make T, P, C transformations of Dirac field ψ one after another:

$$\begin{aligned} T\psi(t,\mathbf{r}) &= -i\gamma^1\gamma^3\psi^*(-t,\mathbf{r}) , \\ PT\psi(t,\mathbf{r}) &= i\gamma^0(T\psi) = \gamma^0\gamma^1\gamma^3\psi^*(-t,-\mathbf{r}) , \\ CPT\psi(t,\mathbf{r}) &= \gamma^2(\gamma^0\gamma^1\gamma^3\psi^*)^* = \gamma^2\gamma^0\gamma^1\gamma^3\psi(-t,-\mathbf{r}) \end{aligned} \tag{3.307}$$

or

$$CPT\psi(t,\mathbf{r}) = i\gamma^5\psi(-t,-\mathbf{r}) . \tag{3.308}$$

Applying these operations to (3.289), we can find the following transformation rules for creation and annihilation operators [6]:

$$\begin{aligned} a^C_{\mathbf{p}\sigma} &= b_{\mathbf{p}\sigma} , \quad & b^C_{\mathbf{p}\sigma} &= a_{\mathbf{p}\sigma} , \\ a^P_{-\mathbf{p}\sigma} &= i a_{\mathbf{p}\sigma} , \quad & b^P_{-\mathbf{p}\sigma} &= i b_{\mathbf{p}\sigma} , \\ a^T_{-\mathbf{p}-\sigma} &= 2\sigma i a^+_{\mathbf{p}\sigma} , \quad & b^T_{-\mathbf{p}-\sigma} &= 2\sigma i b^+_{\mathbf{p}\sigma} . \end{aligned} \tag{3.309}$$

3.3.8 Bilinear forms

As bispinors ψ and ψ^* have 4 components each, their multiplication produces $4\times 4 = 16$ independent bilinear combinations. In symmetric form these combinations can be written as (cf. (3.270), (3.271)

$$\begin{aligned} S = \bar{\psi}\psi , \quad V^\mu &= \bar{\psi}\gamma^\mu\psi , \\ P = i\bar{\psi}\gamma^5\psi , \quad A^\mu &= \bar{\psi}\gamma^\mu\gamma^5\psi , \\ T^{\mu\nu} &= i\bar{\psi}\sigma^{\mu\nu}\psi , \end{aligned} \tag{3.310}$$

where

$$\sigma^{\mu\nu} = \frac{1}{2}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu) . \tag{3.311}$$

These bilinear forms reduce to one scalar S, one pseudoscalar P, the 4-vector V^μ, the 4-pseudovector A^μ, and the antisymmetric tensor $T^{\mu\nu}$.

The scalar nature of S and the pseudoscalar of P are obvious from their spinor representations (cf. (3.208) and (3.209)):

$$S = \xi^*\eta + \eta^*\xi , \quad P = i(\xi^*\eta - \eta^*\xi) . \tag{3.312}$$

The vector nature of V^μ is now clear from the Dirac equation $p_\mu \gamma^\mu \psi = m\psi$, which gives $(\bar{\psi}\, p_\mu \gamma^\mu \psi) = m\bar{\psi}\psi$, where scalars are standing in both sides.

The rule of construction of bilinear forms (3.310) is obvious: they are composed in such a way that γ^μ represents a 4-vector, γ^5 is a pseudovector, while $\bar{\psi}$ and ψ, standing on both sides produce a scalar. The absence of bilinear forms reducing to a symmetric 4-tensor is clear from the fact that the symmetric combination $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$, so that the corresponding bilinear form reduces to $g^{\mu\nu}\bar{\psi}\psi$. In practice, bilinear forms (3.310) are widely used during construction of different interaction Lagrangians of spinor fields. The transformation rules of the bilinear forms under discrete transformations C, P, T can be found in [6].

3.3.9 The neutrino

We have seen above that the necessity to describe a particle with spin $s = 1/2$ by two spinors ξ and η is directly related to the finite mass of a particle. There is no such demand if the particle mass is zero[11]. A wave equation describing such a particle can be constructed using only one spinor, e. g., η:

$$p^{\alpha\dot{\beta}}\eta_{\dot{\beta}} = 0 \tag{3.313}$$

or, which is the same,

$$(p_0 + \mathbf{p}\sigma)\eta = 0\,. \tag{3.314}$$

This is the so-called Weyl equation.

We have noted before, that the wave equation with mass m is automatically invariant with respect to spatial inversion (transformation $\xi \leftrightarrow \eta$ (3.216)). However, if we describe our particle by one spinor, this symmetry is lost.

The energy and momentum of a particle with $m = 0$ are related by $\varepsilon = |\mathbf{p}|$. Thus, for a plane wave $\eta_p \sim e^{-ipx}$ equation (3.314) gives

$$(\mathbf{n}\cdot\sigma)\eta_p = -\eta_p\,, \tag{3.315}$$

where $\mathbf{n} = \frac{\mathbf{p}}{|\mathbf{p}|}$. The same equation holds for the wave with "negative frequency" $\eta_{-p} \sim e^{ipx}$:

$$(\mathbf{n}\cdot\sigma)\eta_{-p} = -\eta_{-p}\,. \tag{3.316}$$

[11] Among all the known fermions only the neutrino possesses mass, which is zero or very small: the experimental limitation is that its mass $m_\nu < 2\,\text{eV}$ [67]. However, experimentally observed neutrino oscillations definitely show, that neutrino mass is definitely finite [67]. Still, zero mass is a very good approximation to describe most of neutrino physics.

The second quantized operators of the field η are represented by

$$\eta = \sum_{\mathbf{p}} (\eta_p a_{\mathbf{p}} + \eta_{-p} b^+_{\mathbf{p}}) ,$$
$$\eta^+ = \sum_{\mathbf{p}} (\eta^*_p a^+_{\mathbf{p}} + \eta^*_{-p} b_{\mathbf{p}}) . \tag{3.317}$$

From here, as usual, we see that η^*_{-p} is the wave function of an antiparticle. A neutrino is electrically neutral, but in this formalism it is not a truly neutral particle!

From the definition of operators $p^{\alpha\dot\beta}$ (3.212) it can be seen that $p^{\alpha\dot\beta *} = -p^{\dot\alpha\beta}$. Then, the complex conjugate spinor η^* satisfies the equation $p^{\dot\alpha\beta}\eta^*_{\dot\beta} = 0$, or, which is the same,

$$p_{\dot\alpha\beta}\eta^{\dot\beta *} = 0 . \tag{3.318}$$

Let us denote $\eta^{\dot\beta *} = \xi^\beta$, as complex conjugation transforms dotted spinors into non-dotted ones. Thus, the wave functions of antiparticles satisfy the equation

$$p_{\dot\alpha\beta}\xi^\beta = 0 \tag{3.319}$$

or

$$(p_0 - \mathbf{p}\boldsymbol{\sigma})\xi = 0 . \tag{3.320}$$

For the plane wave we have

$$(\mathbf{n}\cdot\sigma)\xi_p = \xi_p . \tag{3.321}$$

Note that $\frac{1}{2}(\mathbf{n}\cdot\sigma)$ represents the operator of spin projection on the particle momentum (helicity). Thus, equations (3.315) and (3.321) automatically describe the particles with fixed helicities – spin projection is always oriented parallel to the momentum. The spin of the particle is opposite to the momentum (helicity is equal to $-1/2$, "left screw"), while the spin of an antiparticle is oriented along the momentum (helicity $+1/2$, "right screw"). Accordingly, for neutrinos and antineutrinos there is no symmetry towards reflections in the plane, orthogonal to momentum, as shown in Figure 3.2. This corresponds to the experimentally observed breaking of spatial parity in weak interactions. However, symmetry towards CP operation persists, which corresponds to the conservations of the so-called combined parity[12]. This scheme represents the two-component theory of neutrinos first proposed by Landau.

With one spinor η (or ξ) we can construct only four bilinear combinations, which together form a 4-vector:

$$j^\mu = (\eta^*\eta, \eta^*\sigma\eta) . \tag{3.322}$$

[12] In fact, weak interactions weakly break also the CP-invariance, which is mainly observed in decays of K-mesons. This obviously corresponds to the weak breaking of T-invariance. The physical nature of the breaking of CP-invariance has not been well established up to now, and our description of neutrinos simply neglect this small effect.

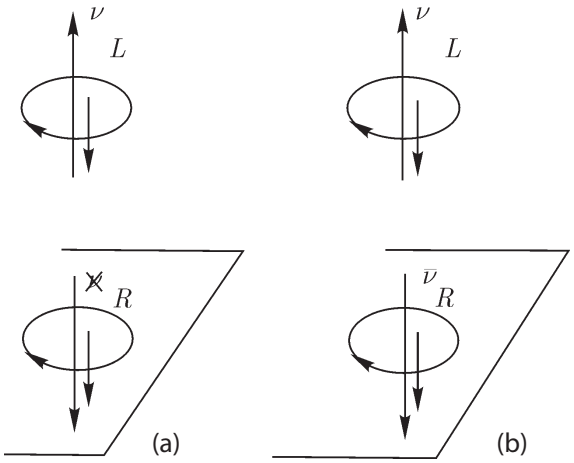


Figure 3.2. Under mirror reflection (spatial inversion) the left-hand neutrino is transformed into a nonexistent right-hand neutrino (a). A real physical state is obtained with simultaneous transformation from particles to antiparticles (charge conjugation), when the left-hand neutrino is transformed into the right-hand antineutrino (b).

Using $(p_0+\mathbf{p}\boldsymbol{\sigma})\eta = 0$ and $\eta^*(p_0-\mathbf{p}\}\boldsymbol{\sigma}) = 0$ we get the continuity equation $\partial_\mu j^\mu = 0$, so that j^μ represents the 4-vector of the neutrino current density.

Neutrino plane waves are conveniently written as

$$\eta_p = \frac{1}{\sqrt{2\varepsilon}} u_p e^{-ipx}\,, \qquad \eta_{-p} = \frac{1}{\sqrt{2\varepsilon}} u_{-p} e^{ipx}\,, \tag{3.323}$$

and spinor amplitudes are normalized by the invariant condition:

$$u^*_{\pm p}(1,\sigma)u_{\pm p} = 2(\varepsilon,\mathbf{p})\,. \tag{3.324}$$

Then particle density and current density are equal to $j^0 = 1, \mathbf{j} = \frac{\mathbf{p}}{\varepsilon} = \mathbf{n}$.

To describe neutrino interactions with other particles, it is convenient to use common notations and introduce the neutrino "bispinor", with two components just equal to zero: $\psi = \begin{pmatrix} 0 \\ \eta \end{pmatrix}$. However, such a form of ψ, in general, changes after the transformation to another (nonspinor) representation. We can overcome this difficulty, noting that in spinor representation

$$\frac{1}{2}(1+\gamma^5) = \frac{1}{2}\left\{\begin{pmatrix} \hat{1} & 0 \\ 0 & \hat{1} \end{pmatrix} + \begin{pmatrix} -\hat{1} & 0 \\ 0 & \hat{1} \end{pmatrix}\right\} = \begin{pmatrix} 0 & 0 \\ 0 & \hat{1} \end{pmatrix},$$
$$\frac{1}{2}(1-\gamma^5) = \begin{pmatrix} \hat{1} & 0 \\ 0 & 0 \end{pmatrix}, \tag{3.325}$$

so that we can write the following identities:

$$\frac{1}{2}(1+\gamma^5)\begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} 0 \\ \eta \end{pmatrix}, \quad (\eta^*,\xi^*)\frac{1}{2}(1-\gamma^5) = (\eta^*,0)\,, \tag{3.326}$$

where ξ is an arbitrary "dummy" spinor. Then the condition of the true two-component nature of a neutrino will also be satisfied in its description by the 4-component bispinor ψ in arbitrary representation, if ψ is understood to be the solution of Dirac's equation with $m = 0$:

$$\widehat{p}\psi = 0 \tag{3.327}$$

with an additional condition (γ^5-invariance)

$$\frac{1}{2}(1+\gamma^5)\psi = \psi \qquad \text{or} \qquad \gamma^5\psi = \psi\,. \tag{3.328}$$

This condition can be taken into account automatically, if in all expressions we replace neutrino bispinors with the following rule:

$$\psi \to \frac{1}{2}(1+\gamma^5)\psi\,, \qquad \bar{\psi} \to \bar{\psi}\frac{1}{2}(1-\gamma^5)\,. \tag{3.329}$$

For example, the 4-vector of the current density is written as

$$j^\mu = \frac{1}{4}\bar{\psi}(1-\gamma^5)\gamma^\mu(1+\gamma^5)\psi = \frac{1}{2}\bar{\psi}\gamma^\mu(1+\gamma^5)\psi\,. \tag{3.330}$$

From the previous discussion of the helicity of massless fermions, it is clear that in the general case we can introduce "right-hand" and "left-hand" fields as

$$\psi^R = \frac{1}{2}(1+\gamma^5)\psi\,, \quad \psi^L = \frac{1}{2}(1-\gamma^5)\psi\,, \quad \psi = \psi^R + \psi^L\,. \tag{3.331}$$

These notations are often used not only for neutrinos, but also for any other fermions with spin $s = 1/2$, while discussing problems where we can neglect fermion masses.

In recent years, because of the indirect experimental indications of neutrino mass finiteness (such as neutrino oscillations [67]), there was a revival of interest in a truly neutral model, the so-called *Majorana neutrino*, which is transformed to itself under charge conjugation and possesses a finite mass which is in some sense different from the usual Dirac's mass. We have seen that mass term of the Dirac type in the Lagrangian mixes L and R components of the same field:

$$\mathcal{L}_D = D(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L) = D\bar{\psi}\psi\,, \tag{3.332}$$

where D denotes Dirac's mass. The mass term of the Majorana type mixes L and R components of *charge conjugate* fields, so that the corresponding contributions to the Lagrangian can be written as [13]

$$\begin{aligned}\mathcal{L}_{MA} &= A(\bar{\psi}^C_L\psi_L + \bar{\psi}_L\psi^C_L) = A\bar{\chi}\chi\,,\\ \mathcal{L}_{MB} &= B(\bar{\psi}^C_R\psi_R + \bar{\psi}_R\psi^C_R) = B\bar{\omega}\omega\,,\end{aligned} \tag{3.333}$$

where the index C denotes the charge conjugation and the Hermitian (truly neutral or Majorana) fields we introduced:

$$\begin{aligned}\chi &= \psi_L + \psi^C_L\,, \quad \chi^C = \chi\,,\\ \omega &= \psi_R + \psi^C_R\,, \quad \omega^C = \omega\,.\end{aligned} \tag{3.334}$$

The inverse relations have the form

$$\psi_L = \frac{1}{2}(1-\gamma^5)\chi\,, \quad \psi_L^C = \frac{1}{2}(1+\gamma^5)\chi\,,$$
$$\psi_R = \frac{1}{2}(1+\gamma^5)\omega\,, \quad \psi_R^C = \frac{1}{2}(1-\gamma^5)\omega\,. \tag{3.335}$$

When both Dirac and Majorana terms are present in the Lagrangian, we have

$$\begin{aligned}\mathcal{L}_{DM} &= D\bar{\psi}_L\psi_R + A\bar{\psi}_L^C\psi_L + B\bar{\psi}_R^C\psi_R + h.c \\ &= \frac{1}{2}D(\bar{\chi}\omega + \bar{\omega}\chi) + A\bar{\chi}\chi + B\bar{\omega}\omega = (\bar{\chi},\bar{\omega})\begin{pmatrix} A & \frac{1}{2}D \\ \frac{1}{2}D & B\end{pmatrix}\begin{pmatrix}\chi\\ \omega\end{pmatrix}.\end{aligned} \tag{3.336}$$

The mass matrix appearing here is easily diagonalized, and its eigenvalues give

$$m_{1,2} = \frac{1}{2}(A+B) \pm \frac{1}{2}\sqrt{(A-B)^2 + D^2}\,. \tag{3.337}$$

Thus, the most general mass term (3.336) with 4-component fermion fields in fact describes two Majorana particles with different masses. The corresponding fields are represented by the following (diagonalizing (3.336)) linear combinations of initial fields:

$$\phi_1 = \chi\cos\theta - \omega\sin\theta\,, \quad \phi_2 = \chi\sin\theta + \omega\cos\theta\,, \tag{3.338}$$

where

$$tg2\theta = \frac{D}{B-A}\,. \tag{3.339}$$

It is easy to see that in the case of $A = B = 0$ (i. e., for zero Majorana masses) we obtain the usual formalism of a 4-component Dirac's field, so that the Dirac's fermion corresponds to the "degenerate" limit $A = B = 0$ of two Majorana particles. Majorana mass terms in the Lagrangian (3.333) obviously lead to nonconservation of additive quantum numbers carried by the field ψ, e. g., electric charge, so that all the known elementary fermions, except neutrinos, being charged should have $A = B = 0$ and should be Dirac's particles. For the neutrino there is no such limitation, and it can be described within the more general Majorana formalism. If the neutrino mass is precisely zero, Majorana neutrinos become equivalent to two-component Weyl neutrinos, considered above. In case of finite neutrino masses, the Majorana theory leads to a number of specific predictions which we shall not discuss here.

Particles with spin $s = 3/2$

Particles with spin $s = 3/2$ in their rest system are described by a symmetric three-dimensional spinor of the third rank, possessing $2s + 1 = 4$ independent components. Accordingly, in the arbitrary Lorentz system of coordinates, the description of such particles should be done using spinors $\xi^{\alpha\dot{\beta}\dot{\gamma}}$, $\eta_{\dot{\alpha}\beta\gamma}$, $\zeta^{\alpha\beta\gamma}$, and $\chi_{\dot{\alpha}\dot{\beta}\dot{\gamma}}$, each of which is symmetric over all indices of identical nature (i. e., dotted or nondotted). Note that the last pair of spinors do not add anything new to the equations obtained with the help of the first pair. There exist several equivalent formulations of wave equations for this problem, but we shall limit ourselves to a short review of only one [6].

We have seen above that to a pair of spinor indices $\alpha\dot{\beta}$ we can associate a single 4-vector index μ. Thus, we can associate $\xi^{\alpha\dot{\beta}\dot{\gamma}} \to \psi_\mu^{\dot{\gamma}}$ and $\eta^{\dot{\beta}\alpha\gamma} \to \psi_\mu^{\gamma}$, i. e., introduce the "mixed" spinor-tensors. Then we can introduce the "vector" bispinor ψ_μ (where we have dropped the

bispinor indices). The wave equation is written as a "Dirac equation" for each of vector components ψ_μ:

$$(\widehat{p} - m)\psi_\mu = 0 \tag{3.340}$$

with an additional condition

$$\gamma^\mu \psi_\mu = 0\,. \tag{3.341}$$

Multiplying (3.340) by γ^μ, taking into account (3.341), we obtain $\gamma^\mu \gamma^\nu p_\nu \psi_\mu = 0$, or, due to commutation rules for γ^μ: $2g^{\mu\nu} p_\nu \psi_\mu - \gamma^\nu \gamma^\mu \psi_\mu = 0$, where the second term gives zero due to (3.341). Then we have

$$p^\mu \psi_\mu = 0\,, \tag{3.342}$$

which guarantees that four-dimensional spinors are transformed to the "right" three-dimensional spinors in the rest system.

The problem of the correct account of additional conditions for wave equations creates significant difficulties during the quantization procedure. Note, however, that, as in the case of particles with higher integer spins, there are no fundamental fermions with $s \geq 3/2$ in the Standard Model.

Chapter 4

The Feynman theory of positron and elementary quantum electrodynamics

4.1 Nonrelativistic theory. Green's functions

In this chapter we shall present an elementary introduction in quantum electrodynamics (QED), which is understood as the theory of the electromagnetic interactions of elementary leptons, i. e., mostly of electrons and positrons. Here we mainly follow the original papers of Feynman; similar presentations can be found in [18, 60]. However, we will begin with nonrelativistic quantum mechanics, so as to introduce some basic notions which are usually not included in traditional courses like [35].

Consider the time-dependent Schroedinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi\,. \tag{4.1}$$

The standard approach to solve this differential equation starts with the wave function at some initial moment of time $\psi(t_1)$, the calculation of its change during a small time interval Δt, to find $\psi(t_1+\Delta t)$, with further continuation of this process of integration. Feynman proposed writing the solution of equation (4.1) in integral form, so that at time $t_2 > t_1$ the wave function at space-time point $(t_2, \mathbf{x}_2)$ is expressed via the wave function at space-time point $(t_1, \mathbf{x}_1)$ as follows:

$$\psi(\mathbf{x}_2, t_2) = \int d^3\mathbf{x}_1 K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1)\psi(\mathbf{x}_1 t_1)\,, \qquad t_2 \geq t_1\,. \tag{4.2}$$

Here the integral kernel $K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1)$ represents the *propagator* (Green's function), corresponding to the linear differential equation (4.1). The physical meaning of propagator is clear from the form of equation (4.2) – this is the quantum mechanical probability amplitude of particle transition from point $\mathbf{x}_1$ at time t_1 to point $\mathbf{x}_2$ at time t_2.

For simplicity, we consider the case of time-independent Hamiltonian H. In accordance with the superposition principle of quantum mechanics, we can expand $\psi(\mathbf{x}_1 t_1)$ in a series over the full set of orthonormalized eigenfunctions $u_n(\mathbf{x})$ of operator H with eigenvalues E_n:

$$\begin{gathered} Hu_n = E_n u_n\,, \\ \int d^3\mathbf{x} u_n^*(\mathbf{x}) u_m(\mathbf{x}) \equiv (u_n, u_m) = \delta_{nm}\,, \\ \sum_n u_n(\mathbf{x}) u_n^*(\mathbf{x}') = \delta(\mathbf{x}-\mathbf{x}')\,. \end{gathered} \tag{4.3}$$

Then,

$$\psi(\mathbf{x}_1 t_1) = \sum_n c_n u_n(\mathbf{x}_1) e^{-iE_n t_1/\hbar} \,. \tag{4.4}$$

The coefficients c_n can be obtained by multiplying (4.4) by $u_n^*(\mathbf{x}_1)$ and performing integration over the whole three-dimensional space:

$$c_n = \int d^3\mathbf{x}_1 u_n^*(\mathbf{x}_1)\psi(\mathbf{x}_1 t_1) e^{iE_n t_1/\hbar} \,. \tag{4.5}$$

The wave function at time t_2 can be written as

$$\psi(\mathbf{x}_2 t_2) = \sum_n c_n u_n(\mathbf{x}_2) e^{-iE_n t_2/\hbar} \,. \tag{4.6}$$

Substituting (4.5) into (4.6), changing the order of summation and integration, and comparing with (4.2), we get

$$K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = \sum_n u_n(\mathbf{x}_2) u_n^*(\mathbf{x}_1) e^{-iE_n(t_2 - t_1)/\hbar} \,, \tag{4.7}$$

and introducing the notation

$$\chi_n(\mathbf{x}, t) = u_n(\mathbf{x}) e^{-iE_n t/\hbar} \,, \tag{4.8}$$

we obtain a shorter expression:

$$K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = \sum_n \chi_n(\mathbf{x}_2 t_2) \,. \chi_n^*(\mathbf{x}_1 t_1) \tag{4.9}$$

At coinciding times $t_1 = t_2 = t$ from (4.7) we obtain

$$K(\mathbf{x}_2 t; \mathbf{x}_1 t) = \sum_n u_n(\mathbf{x}_2) u_n^*(\mathbf{x}_1) = \delta(\mathbf{x}_2 - \mathbf{x}_1) \,. \tag{4.10}$$

This obviously transforms (4.2) for $t_1 = t_2$ into the identity. We are certainly interested in times $t_2 > t_1$; thus it is convenient to put $K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = 0$ for $t_2 < t_1$, which guarantees *causality*, and we define

$$K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = \theta(t_2 - t_1) \sum_n \chi_n(\mathbf{x}_2 t_2) \chi^*(\mathbf{x}_1 t_1) \,, \tag{4.11}$$

where we have introduced the step-function

$$\theta(t) = \begin{cases} 1 & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases} \,. \tag{4.12}$$

The derivative of the θ-function is given by

$$\frac{d\theta(t)}{dt} = \delta(t) \,. \tag{4.13}$$

Now we can write the differential equation for Green's function (propagator) $K(\mathbf{r}_2 t_2; \mathbf{r}_1 t_1)$. As χ_n represent the solutions of the Schroedinger equation (4.1), (4.3), using (4.10), (4.11), and (4.13), we have

$$\left[i\hbar \frac{\partial}{\partial t_2} - H(\mathbf{x}_2) \right] K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = i\hbar \sum_n \chi_n(\mathbf{x}_2 t_2) \chi^*(\mathbf{x}_1 t_1) \frac{\partial}{\partial t_2} \theta(t_2 - t_1)$$
$$= i\hbar \sum_n u_n(\mathbf{x}_2) u_n^*(\mathbf{x}_1) e^{-iE_n(t_2 - t_1)/\hbar} \delta(t_2 - t_1)$$
$$= i\hbar \delta(t_2 - t_1) \sum_n u_n(\mathbf{x}_2) u_n^*(\mathbf{x}_1) = i\hbar \delta(t_2 - t_1) \delta(\mathbf{x}_2 - \mathbf{x}_1) \,. \tag{4.14}$$

Thus, in the general case (even for time-dependent H) Green's function (propagator) $K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1)$ is defined as the solution of the inhomogeneous (with δ-source in the right-hand side) differential equation[1]

$$\left[i\hbar \frac{\partial}{\partial t_2} - H(\mathbf{x}_2 t_2) \right] K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = i\hbar \delta(t_2 - t_1) \delta(\mathbf{x}_2 - \mathbf{x_1}) \,, \tag{4.15}$$

with the boundary condition

$$K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = 0 \quad \text{for} \quad t_2 < t_1 \,. \tag{4.16}$$

For $t_2 \neq t_1$, equation (4.15) reduces to

$$\left[i\hbar \frac{\partial}{\partial t_2} - H(\mathbf{x}_2 t_2) \right] K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = 0 \,. \tag{4.17}$$

If we integrate equation (4.15) over small time interval from $t_2 = t_1 - \varepsilon$ to $t_2 = t_1 + \varepsilon$, we get

$$K(\mathbf{x}_2 t_1 + \varepsilon; \mathbf{x_1} t_1) - K(\mathbf{x}_2 t_1 - \varepsilon; \mathbf{x}_1 t_1) = \delta(\mathbf{x}_2 - \mathbf{x}_1) \,. \tag{4.18}$$

The contribution of the second term in the left-hand side of equation (4.15) drops out as $\varepsilon \to 0$ for finite H. Now take into account that $K(\mathbf{x}_2 t_1 - \varepsilon, \mathbf{x}_1 t_1) = 0$ due to (4.16) and $t_1 - \varepsilon < t_1$. Then,

$$\lim_{\varepsilon \to 0} K(\mathbf{x}_2 t_1 + \varepsilon; \mathbf{x}_1 t_1) = K(\mathbf{x}_2 t_1; \mathbf{x}_1 t_1) = \delta(\mathbf{x}_2 - \mathbf{x}_1) \,, \tag{4.19}$$

which coincides with (4.10).

Thus, the use of equation (4.2) is equivalent to the standard quantum mechanical description. Consider the problem of when we can write $H = H_0 + V$ and if for $H = H_0$ there is an exact solution. Then we can try to construct the perturbation theory over potential V. Let us denote $K_0(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1)$ as the Green's function of a "free" particle, moving in the absence of perturbation V. We can easily convince ourselves

[1] Actually, our definition of Green's function of the Schroedinger equation coincides with its definition in mathematical physics [70].

that differential equation (4.15) and boundary condition (4.16) can be unified in a single *integral equation*:

$$K(2,1) = K_0(2,1) - \frac{i}{\hbar}\int d^4x_3 K_0(2,3)V(3)K(3,1)\,, \tag{4.20}$$

where the numbers denote space-time points, e. g., $(2) = (\mathbf{x}_1, t_2)$ etc., and we have introduced the four-dimensional integration variable $x_3 = (\mathbf{x}_3, t_3)$, where integration over time t_3 is formally done within infinite limits (the presence of θ-function in the definition of propagator automatically guarantees the correct finite limits of integration). To check the validity of equation (4.20) let us act on both of its sides by the operator $[i\hbar\frac{\partial}{\partial t_2} - H_0(2)]$, so that using equation (4.15) for K_0 (i. e., for $V = 0$), we obtain the differential equation

$$\left[i\hbar\frac{\partial}{\partial t_2} - H_0(2)\right]K(2,1) = i\hbar\delta(2,1) + V(2)K(2,1)\,,$$

which, after moving the second term in the right-hand side to the left-hand side, just coincides with equation (4.15). As $K_0(2,1) = 0$ for $t_2 < t_1$, we also have $K(2,1) = 0$ for $t_2 < t_1$.

The advantage of integral equation (4.20) is that it is conveniently solved by iterations, so that we obtain the following perturbation series for the propagator:

$$\begin{aligned} K(2,1) = K_0(2,1) &- \frac{i}{\hbar}\int d^4x_3 K_0(2,3)V(3)K_0(3,1) \\ &+ \left(\frac{-i}{\hbar}\right)^2 \int d^4x_3 d^4x_4 K_0(2,3)V(3)K_0(3,4)V(4)K_0(4,1) + \cdots\,. \end{aligned} \tag{4.21}$$

The terms of this expansion have an obvious and graphic interpretation: the first term describes the propagation of a free particle from point 1 to point 2, and the second describes the propagation of a free particle from point 1 to point 3, where it is scattered by potential V and the propagates as a free particle from 3 to 2. Obviously, point 3 is arbitrary, so that we have to integrate over its coordinates. This process continues in higher orders, so that the series describes all the multiple scattering processes up to an infinite order in V. Such a perturbation theory can be effectively used for the solution of concrete problems, and we shall return to it later.

4.2 Relativistic theory

Now we are going to construct a similar formalism in relativistic theory. Dirac's equation for a *free* particle is written as

$$(i\widehat{\nabla} - m)\psi = 0\,, \quad \widehat{\nabla} = \gamma^\mu\partial_\mu = \gamma^0\partial_0 + \boldsymbol{\gamma}\cdot\nabla\,, \tag{4.22}$$

where, in contrast to the previous section, we returned to a "natural" system of units with $\hbar = c = 1$. A 4-component wave function (Dirac's bispinor) $\psi(\mathbf{x}_2 t_2)$ can be obtained from the "initial" one $\psi(\mathbf{x}_1 t_1)$ with the help of the propagator (Green's function) $K_0(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1)$, which is a 4×4-matrix. This matrix should satisfy the Dirac's equation with the right-hand side, similar to equation (4.15):

$$(i\widehat{\nabla}_2 - m) K_0(2,1) = i\delta(2,1)\,, \tag{4.23}$$

where, as before, we use the obvious number notations for space-time points. In analogy with (4.4), (4.6), the wave function ψ can be expanded over the complete set of eigenfunctions u_n, corresponding both to positive and negative energies. Instead of u_n^*, it is convenient to use conjugate spinors $\bar{u}_n = u^\dagger \gamma^0 = u^\dagger \beta$ (where we use Dirac's matrices $\gamma^0 = \beta$, $\beta^2 = 1$). Repeating the steps used to derive (4.7), we find the desirable propagator as

$$\begin{aligned} &K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = \\ &\quad \sum_{E_n>0} u_n(\mathbf{x}_2)\bar{u}_n(\mathbf{x}_1) e^{-iE_n(t_2-t_1)} + \sum_{E_n<0} u_n(\mathbf{x}_2)\bar{u}_n(\mathbf{x}_1) e^{-iE_n(t_2-t_1)} \quad \text{for } t_2 > t_1\,, \\ &\quad K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = 0 \quad \text{for } t_2 < t_1\,. \end{aligned} \tag{4.24}$$

It is necessary to do the expansion of ψ over the *complete* set of eigenfunctions, including those corresponding to negative energies. This may seem bad for physics, as the introduction of some external perturbation (potential) can induce quantum transition of a particle (e. g., an electron) from positive energy states to the states with negative energies, which leads to instability of the system (absence of the ground state). It is well-known that Dirac solved this problem in the following way. Let us assume that all states with negative energies in the physical ground state (vacuum) are already filled by electrons, so that the Pauli principle prevents an electron, moving above this vacuum, from making a transition to the filled states with negative energies. In our formalism this reduces to the requirement that for $t_2 > t_1$ propagator $K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1)$ should be the sum over the eigenfunctions, corresponding to positive energies only. Mathematically we should write

$$K(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = \sum_{E_n>0} u_n(\mathbf{x}_2)\bar{u}_n(\mathbf{x}_1) e^{-iE_n(t_2-t_1)} \quad \text{for } t_2 > t_1\,, \tag{4.25}$$

so that from (4.24) we should subtract the sum of terms like

$$u_n(\mathbf{x}_2)\bar{u}_n(\mathbf{x}_1) e^{-iE_n(t_2-t_1)} \tag{4.26}$$

over the states with negative energies for *all* moments of time. This can be done, as such a sum represents the solution of homogeneous (i. e., without the right-hand side of) equation (4.23). As a result, this sum is cancelled by the second half of the solution

(4.24), and we obtain the following Green's function for a free particle:

$$K_+(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = \sum_{E_n > 0} u_n(\mathbf{x}_2)\bar{u}_n(\mathbf{x}_1) e^{-iE_n(t_2 - t_1)} \quad \text{for } t_2 > t_1 \,. \tag{4.27}$$

However, for this function

$$K_+(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = -\sum_{E_n < 0} u_n(\mathbf{x}_2)\bar{u}_n(\mathbf{x}_1) e^{-iE_n(t_2 - t_1)} \quad \text{for } t_2 < t_1 \,. \tag{4.28}$$

is obviously nonzero for $t_2 < t_1$. Not the minus sign appearing here! This last expression is conveniently written as

$$K_+(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) = -\sum_{E_n < 0} u_n(\mathbf{x}_2)\bar{u}_n(\mathbf{x}_1) e^{-i|E_n||t_2 - t_1|} \quad \text{for } t_2 < t_1 \,, \tag{4.29}$$

so that in the exponent we have only positive energies, and negative energies seem to disappear.

In the presence of an external potential we can again write an integral equation similar to equation (4.20) and its perturbation expansion of equation (4.21), but with K_0 replaced everywhere by K_+, and the potential V should be considered as a 4×4-matrix. The terms of this series can be graphically understood by drawing space-time *Feynman diagrams*. The first term of the series (4.21) $K_+(2, 1)$ describes free particle propagation from point 1 to point 2 (Figure 4.1). The second term (Figure 4.2) has the form

$$(-i) \int d^4 x_3 K_+(2, 3) V(3) K_+(3, 1) \tag{4.30}$$

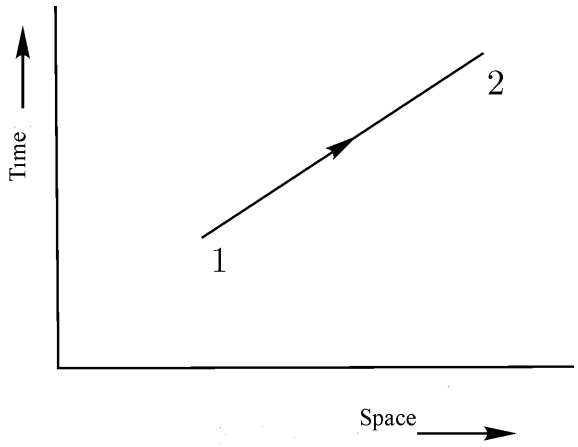


Figure 4.1

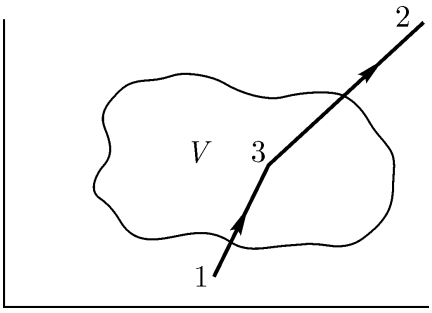


Figure 4.2

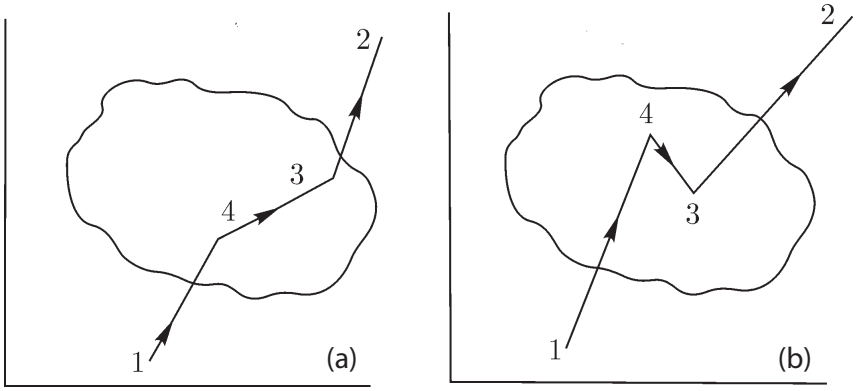


Figure 4.3

and describes single scattering. In the figure the closed curve denotes the region, where the potential V is different from zero. The third term (Figure 4.3)

$$(-i)^2 \int d^4x_3 d^4x_4 K_+(2,3)V(3)K_+(3,4)V(4)K_+(4,1) \tag{4.31}$$

describes two scattering processes. Diagrams in Figure 4.3(a) and Figure 4.3(b) illustrate two variants of such a scattering:

- *Case (a)*. From point 1 to point 2 the electron propagates in such a way that time grows along its world line, expression for K_+ contains only sums like (4.27), i. e., only states with positive particle energies are accounted for. These re the usual second order scatterings of an electron with positive energy, as in nonrelativistic theory.

- *Case (b)*. From point 4 to point 3 the electron propagates *backwards* in time, then expression for K_+ contains only sums over negative energy states like (4.29). In the Feynman interpretation this corresponds to a *positron* (i. e., antiparticle) propagation

from point 3 to point 4. If time always grows, we can interpret this sequence of events in the following way: An electron–positron pair is created at point 3, with the electron propagating in the direction of point 2, while the positron propagates to point 4, where it annihilates with the initial electron, coming there from 1.

Thus, according to Feynman, *a positron is an electron propagating backwards in time*.

This interpretation can also be illustrated from the classical point of view: in equations of motion of a classical particle in electromagnetic field [33]

$$m\frac{d^2x^\mu}{ds^2} = e\frac{dx_\nu}{ds}F^{\mu\nu}\,. \tag{4.32}$$

the change of proper time direction s is equivalent to the change of the charge e sign.

Note that for the process shown in Figure 4.3 we certainly have to integrate over all the possible values of times t_3 and t_4, so that both cases are described by the single term of the series (4.31), which can be represented by the single diagram of Figure 4.3(a), while the diagram of Figure 4.3(b) is just identical. The scattering process shown in Figure 4.3(b) is in fact in accordance with Dirac's theory: an electron with negative energy goes to the state 2 with positive energy (final state), i. e., an electron–positron pair is created, while the hole is filled by the electron coming from 1, which annihilates. As a result, the electron is scattered from state 1 to state 2, and the electron with positive energy is replaced by one of the electrons from the negative "background" (vacuum). Thus, we have an *exchange* of identical particles, and the corresponding matrix element acquires a negative sign, as it should be for fermions. However, we never used the Pauli principle! The appearance of a negative sign in (4.28) was due to a method of construction of propagator K_+, thus guaranteeing the correct statistics! The generalization of these arguments to all higher orders of perturbation theory produces an alternative proof of the spin-statistics theorem [18].

4.3 Momentum representation

In practice, all calculations are most conveniently done in the momentum representation. Propagator K_+ is determined by the equation

$$(i\widehat{\nabla} - m)K_+(2,1) = i\delta(2,1)\,. \tag{4.33}$$

Let us introduce the Fourier transform of K_+, which we shall denote as $S_+(p)$, so that

$$K_+(2,1) = \int d^4p e^{-ip(x_2-x_1)}S_+(p)\,, \tag{4.34}$$

where $d^4p = dp_0 d^3\mathbf{p}$. The operator $(i\widehat{\nabla} - m)$ can be moved under the integral as $(\hat{p} - m)$, while for the δ-function we can write

$$\delta(x_2 - x_1) = \int \frac{d^4p}{(2\pi)^4} e^{-ip(x_2 - x_1)} . \tag{4.35}$$

Then it is easy to find the following equation for $S_+(p)$:

$$S_+(p) = \frac{i}{(2\pi)^4} \frac{1}{\hat{p} - m} . \tag{4.36}$$

This expression can be rewritten as

$$S_+(p) = \frac{i}{(2\pi)^4} \frac{\hat{p} + m}{\hat{p} + m} \frac{1}{\hat{p} - m} = \frac{i}{(2\pi)^4} \frac{\hat{p} + m}{p^2 - m^2} , \tag{4.37}$$

where we have taken into account that $\hat{p}^2 = p_\mu p^\mu = p_0^2 - \mathbf{p}^2 = p^2$, so that the denominator in (4.37) contains no matrices. Then,

$$K_+(2,1) = \frac{i}{(2\pi)^4} \int d^4p e^{-ip(x_2 - x_1)} \frac{\hat{p} + m}{p^2 - m^2} . \tag{4.38}$$

Let us introduce, by definition, the following integral:

$$I_+(2,1) = \frac{1}{(2\pi)^4} \int d^4p \frac{e^{-ip(x_2 - x_1)}}{p^2 - m^2} . \tag{4.39}$$

Then (4.38) can be written as

$$K_+(2,1) = i(i\widehat{\nabla}_2 + m) I_+(2,1) . \tag{4.40}$$

Substituting (4.40) into (4.33) we find that I_+ satisfies the following equation:

$$(\Box + m^2) I_+(2,1) = -\delta(2,1) \tag{4.41}$$

i. e., it represents, in fact, Green's function of the Klein–Gordon equation. Separating the spatial coordinates and time, we rewrite I_+ as

$$I_+(x_2 - x_1) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} dp_0 \int d^3\mathbf{p} \frac{e^{-ip_0(t_2 - t_1)} e^{i\mathbf{p}(\mathbf{r}_2 - \mathbf{r}_1)}}{p_0^2 - \mathbf{p}^2 - m^2} . \tag{4.42}$$

Here we have a problem: the integrand contains poles at $p_0^2 - \mathbf{p}^2 - m^2 = 0$, i. e., for $p_0 = \pm\sqrt{\mathbf{p}^2 + m^2} \equiv \pm E_\mathbf{p}$. Thus, we have to introduce a certain integration path going around these poles (the rule to encircle poles). Feynman's rule is to replace

$$m \to m - i\delta , \quad \delta > 0 , \quad \delta \to +0 . \tag{4.43}$$

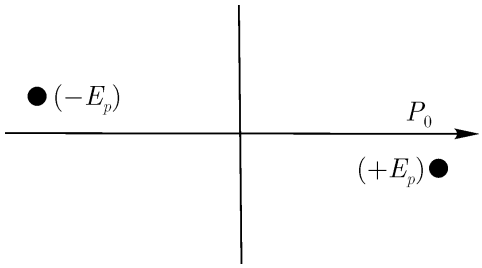


Figure 4.4

Then our integral is written as

$$I_+(x_2 - x_1) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} e^{i\mathbf{p}(\mathbf{r}_2 - \mathbf{r}_1)} \int_{-\infty}^{\infty} \frac{dp_0}{2\pi} \frac{e^{-ip_0(t_2 - t_1)}}{(p_0 - E_\mathbf{p} + i\varepsilon)(p_0 + E_\mathbf{p} - i\varepsilon)}, \quad (4.44)$$

because the replacement (4.43) adds to $E_\mathbf{p}$ an infinitesimal imaginary part, which we denoted $i\varepsilon$. Consider now the integrand as a function of a complex variable p_0. The pole $p_0 = -E_\mathbf{p}$ is now slightly above the real axis of p_0 (which is the integration path in (4.44)), while the pole $p_0 = +E_\mathbf{p}$ is slightly below it, as shown in Figure 4.4. Let us integrate (4.44), assuming that $t_2 - t_1 > 0$. In this case, integration is easily done using the Cauchy theorem and closing the integration contour (path) in the lower half-plane of p_0. The integral over the semicircle at infinity gives zero, due to fast damping of the exponential factor in the integrand, and what remains is just the integral along the real axis we need. But the integral over the closed contour is simply determined by the residue at the pole $+E_\mathbf{p}$, which is inside the contour (and is encircled clockwise). Finally we get

$$-\frac{2\pi i}{2E_\mathbf{p}} e^{-iE_\mathbf{p}(t_2 - t_1)}. \quad (4.45)$$

In the case of $t_2 - t_1 < 0$, to make zero the contribution of a semicircle at infinity we have to close the integration contour in the upper half-plane. Then inside the integration contour we have only the pole at $-E_\mathbf{p}$, which is encircled counterclockwise, so that our integral is equal to

$$-\frac{2\pi i}{2E_\mathbf{p}} e^{+iE_\mathbf{p}(t_2 - t_1)}. \quad (4.46)$$

Note that $E_\mathbf{p}$ is assumed to be positive by definition, so that the arguments of exponential factors in both (4.45) and (4.46) are positive (up to a factor of $-i$). Thus, both the integral I_+ and the propagator K_+ behave in way similar to (4.27) and (4.28); for $t_2 - t_1 > 0$, only positive energies contribute, while for $t_2 - t_1 < 0$ – only negative ones! In fact, the replacement (4.43) guarantees the equivalence to our previous definition of propagator K_+.

We could define integration (4.42) in another way and instead of (4.43) just add an infinitesimal imaginary part to p_0:

$$p_0 \to p_0 + i\delta\,, \quad \delta \to +0\,. \tag{4.47}$$

Then, in the integrand of (4.42) two poles appear, both in the lower half-plane of p_0. Then for $t_2 - t_1 > 0$, when the integration contour is naturally closed in the lower half-plane, both positive and negative energies contribute. At the same time, for $t_2 - t_1$ we close the contour in the upper half-plane, where there are no poles at all, so that the integral is just zero. This definition of propagator K gives, in fact, equation (4.24) ("retarded" Green's function), i. e., Dirac's theory for electrons only. Feynman's rule also has an obvious advantage of the imaginary part being introduced into the relativistic invariant m, so that all expressions remain covariant, while in the theory of "electrons only" the imaginary part in p_0 makes it different from other components of 4-momentum.

We shall return to the discussion of these rather fine details of the definition of analytical properties of Green's functions several times later on, while here we just note the most general property: poles of the propagators (Green's functions) in momentum representation determine, in fact, the energy spectrum of the corresponding particles. In our discussion above, $E_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2}$ represents the relativistic spectrum of an electron (positron). This property of Green's functions is also widely used in the modern condensed matter theory. In particular, it is the basis of the whole concept of quasiparticles – the elementary excitations in many particle systems [1].

4.4 The electron in an external electromagnetic field

Consider an electron interacting with an external electromagnetic field. This interaction is described by the expression $ej^\mu A_\mu = e\bar{\psi}\gamma^\mu\psi A_\mu$, so that the interaction "potential" is conveniently denoted as $e\gamma^\mu A_\mu \equiv e\widehat{A}$ (e is the electron charge). Dirac's equation, taking into account interaction with the electromagnetic field, has an obvious form:

$$(i\widehat{\nabla} - e\widehat{A} - m)\psi = 0\,, \tag{4.48}$$

where we introduced the appropriate covariant derivative of electrodynamics.

Accordingly, the particle propagator in the external field K_+^A is defined by the equation

$$(i\widehat{\nabla}_2 - e\widehat{A}_2 - m)K_+^A(2,1) = i\delta(2,1)\,. \tag{4.49}$$

Dirac's equation (4.48) can also be rewritten in the form of a Schroedinger equation with the appropriate Hamiltonian:

$$i\frac{\partial\psi}{\partial t} = H\psi = [\boldsymbol{\alpha}\cdot(\mathbf{p} - e\mathbf{A}) + e\varphi + m\beta]\,\psi\,, \tag{4.50}$$

where we have taken into account that $A_\mu = (\varphi, -\mathbf{A})$. Then the propagator is defined as the solution of the following equation:

$$\left[i\frac{\partial}{\partial t_2} - e\varphi(2) - \boldsymbol{\alpha}\cdot(-i\nabla_2 - e\mathbf{A}_2) - m\beta \right] K_+^A(2,1) = i\beta\delta(2,1)\,, \tag{4.51}$$

where the appearance of the $\beta = \gamma^0$ matrix in the is related to the use of the Dirac's conjugated spinors in equation (4.24) and guarantees the relativistic invariance. Multiplying equation (4.51) by the matrix β, we write it as

$$(i\hat{\nabla}_2 - e\hat{A}_2 - m)K_+^A(2,1) = i\,\delta(2,1) \tag{4.52}$$

which coincides with (4.49).

The solution of equation (4.49) satisfies the integral equation, similar to equation (4.20):

$$K_+^A(2,1) = K_+(2,1) - ie\int d^4x_3 K_+(2,3)\widehat{A}(3)K_+^A(3,1)\,, \tag{4.53}$$

which produces perturbation expansion (by iteration) similar to equation (4.21):

$$\begin{aligned} K_+^A(2,1) = {} & K_+(2,1) - ie\int d^4x_3 K_+(2,3)\widehat{A}(3)K_+(3,1) \\ & + (-ie)^2\int d^4x_3 d^4x_4 K_+(2,3)\widehat{A}(3)K_+(3,4)\widehat{A}(4)K_+(4,1) + \cdots\,. \end{aligned} \tag{4.54}$$

It seems that in the relativistic case, the relation between wave functions $\psi(2)$ at point x_2 and $\psi(1)$ at point x_1 can be written in analogy with (4.2) as

$$\psi(2) = \int d^3\mathbf{x}_1 K_+^A(2,1)\beta\psi(1)\,, \tag{4.55}$$

where $d^3\mathbf{x}_1$ is the volume element of three-dimensional space at fixed time moment t_1, which is illustrated in Figure 4.5(a). Waves radiated from the point of the hyperplane $t_1 = \text{const}$ form the wave function at point $\mathbf{x}_2$ at a later moment t_2. However. this is all wrong! The thing is that we defined Green's function (propagator) in relativistic theory in such a way that it describes the propagation of particles with positive energies ahead of time and particles with negative energies *backwards* in time. Thus, the analogue of equation (4.2) should be written as

$$\psi(\mathbf{x}_2 t_2) = \int d^3\mathbf{x}_1 K_+^A(\mathbf{x}_2 t_2, \mathbf{x}_1, t_1)\beta\psi(\mathbf{x}_1 t_1) - \int d^3\mathbf{x}_1 K_+^A(\mathbf{x}_2 t_2, \mathbf{x}_1 t_1')\beta\psi(\mathbf{x}_1 t_1')\,, \tag{4.56}$$

where $t_1 < t_2 < t_1'$! Here, in accordance with Figure 4.5(b), the first term represents the contribution of states with positive energy and depends on the previous moments

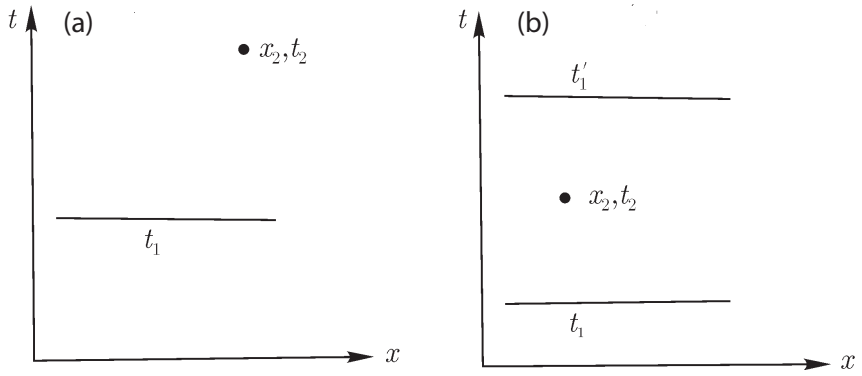


Figure 4.5

of time, while the second term gives the contribution of states with negative energies and depends on future moments of time. The probability amplitude for the particle transition to the point $\mathbf{x}_2, t_2$ is not defined if we only know the probability amplitude to find an electron (or positron) at the earlier moment of time. Even if there was no positron present at an earlier time, an external field could have created an electron–positron pair during the system evolution, which leads to the finite probability amplitude to find a positron in the future. In the Feynman approach, contributions to the propagator corresponding to particles with positive energies are considered as probability amplitudes for an electron with a negative electric charge, while contributions corresponding to particles with negative energies are considered as the probability amplitudes for a positron with energy $-E > 0$. Thus, to determine the wave function of a Dirac's field at some moment of time, we need to know its electronic component at a previous moment of time and the positron component at some *future* moment of time!

Equation (4.56) can be generalized if we note that to define the wave function $\psi(\mathbf{x}_2 t_2)$ we need the knowledge of $\psi(\mathbf{x}_1 t_1)$ on some four-dimensional hypersurface surrounding the point $\mathbf{x}_2, t_2$, as shown in Figure 4.6:

$$\psi(\mathbf{x}_2 t_2) = \int d\sigma(x_1) K_+^A(2,1) \widehat{N}(1) \psi(1)\,, \tag{4.57}$$

where $\widehat{N} = N_\mu \gamma^\mu$, with N_μ being the vector normal to the hypersurface, surrounding $\mathbf{x}_2, t_2$. Integration into equation (4.57) goes over this hypersurface. Then we can say that the form of equation (4.55) assumes precisely this. Thus, in the future, for brevity we shall use this simplest formulation. We only need to remember, that spatial integration in equation (4.55) should be done over the correctly chosen hypersurface in four-dimensional space-time.

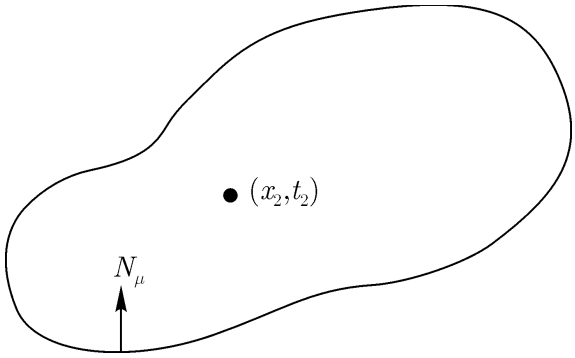


Figure 4.6

The formal derivation of equation (4.57) can be done as follows. Let us use the four-dimensional Gauss theorem:

$$\int_\Omega d^4x' \frac{\partial F_\mu(x')}{\partial x'_\mu} = \int_S d\sigma(x') F_\mu(x') n^\mu(x') \,, \tag{4.58}$$

where $F_\mu(x')$ is some 4-vector function, defined in space-time volume Ω limited by hypersurface S, $n^\mu(x')$ is an external normal to surface element $d\sigma(x')$ at point x'. Let $\psi(x)$ be the solution of the Dirac's equation $i\gamma^\mu \frac{\partial \psi(x)}{\partial x^\mu} - m\psi(x) = 0$. Let us choose $F(x') = iK_+(x - x')\gamma^\mu \psi(x')$, where $x, x' \in \Omega$. Then we have

$$\begin{aligned} \frac{\partial F_\mu(x')}{\partial x'_\mu} &= i \frac{\partial}{\partial x'_\mu} [K_+(x - x')\gamma_\mu \psi(x')] \\ &= \left[i \frac{\partial K_+(x - x')}{\partial x'_\mu} \gamma_\mu + m K_+(x - x') \right] \psi(x') \\ &\quad + K_+(x - x') \left[i\gamma_\mu \frac{\partial \psi(x')}{\partial x'_\mu} - m\psi(x') \right] . \end{aligned} \tag{4.59}$$

The second term here is equal to zero in accordance with Dirac's equation. Let us substitute this expression into the left-hand side of equation (4.58) and take into account that

$$i \frac{\partial K_+(x - x')}{\partial x'_\mu} \gamma_\mu + m K_+(x - x') = -i\delta(x - x') \,. \tag{4.60}$$

Then we have

$$\psi(x) = -\int_S d\sigma(x') K_+(x - x')\gamma_\mu \psi(x') n^\mu(x') \,, \tag{4.61}$$

where n^μ is an external normal. If we introduce internal normal N^μ, we get

$$\psi(x) = \int_S d\sigma(x') K_+(x - x')\gamma_\mu \psi(x') N^\mu(x') \,, \tag{4.62}$$

which completes the proof. Equation (4.56) now follows if we choose hypersurface S consisting of two space-like hyperplanes t_1 and t_2, and neglect the "side" contributions, taking into account that these parts of S are moved to infinity. Note that in this case $N^0\gamma_0 = \beta$.

The probability amplitude for an *electron* transition from some state with wave function $\psi_0(\mathbf{x}_1 t_1)$ of a particle with positive energy at the moment t_1 to the state with wave function $\varphi_0(\mathbf{x}_2 t_2)$, also corresponding to states with positive energies, at the moment $t_2 > t_1$ is given by the following expression:

$$\begin{aligned} M &= \int d^3\mathbf{x}_1 d^3\mathbf{x}_2 \varphi_0^*(\mathbf{x}_2 t_2) K_+(\mathbf{x}_2 t_2; \mathbf{x}_1 t_1) \beta \psi_0(\mathbf{x}_1 t_1) \\ &= \int d^3\mathbf{x}_1 d^3\mathbf{x}_2 \bar{\varphi}_0(2) \beta K_+(2,1) \beta \psi_0(1) \,. \end{aligned} \tag{4.63}$$

If between moments t_1 and t_2 the potential $e\widehat{A}$ acts, the function K_+ is replaced by K_+^A. In the first order, the transition amplitude, according to (4.54), is equal to

$$M_1 = -ie \int d^3\mathbf{x}_1 d^3\mathbf{x}_2 d^4 x_3 \bar{\varphi}_0(2) \beta K_+(2,3) \widehat{A}(3) K_+(3,1) \beta \psi_0(1) \,. \tag{4.64}$$

With the help of (4.55) we can perform integration over $\mathbf{x}_1$ and $\mathbf{x}_2$, introducing

$$\psi_0(3) = \int d^3\mathbf{x}_1 K_+(3,1) \beta \psi_0(1) \,, \tag{4.65a}$$

$$\bar{\varphi}_0(3) = \int d^3\mathbf{x}_2 \bar{\varphi}_0(2) \beta K_+(2,3) \,, \tag{4.65b}$$

so that (4.64) becomes

$$M_1 = -ie \int d^4 x \, \bar{\varphi}_0(x) \widehat{A}(x) \psi_0(x) \,. \tag{4.66}$$

Let the initial wave function correspond to an electron with 4-momentum p_1, while the final one corresponds to an electron with 4-momentum p_2:

$$\psi_0(x) = u(p_1) e^{-ip_1 x} \,, \quad \bar{\varphi}_0(x) = \bar{u}(p_2) e^{ip_2 x} \,, \tag{4.67}$$

where u are spinors, corresponding to free particles with positive energy. Introducing the Fourier transform of $A_\mu(x)$,

$$A_\mu(x) = \int d^4 k e^{-ikx} a_\mu(k) \tag{4.68}$$

and substituting (4.67) and (4.68) into (4.66), we perform integration over x_3 and get

$$\begin{aligned} M_1 &= -ie(2\pi)^4 \int d^4 k \, \delta(p_2 - k - p_1) \bar{u}(p_2) \hat{a}(k) u(p_1) \\ &= -ie(2\pi)^4 \bar{u}(p_2) \hat{a}(p_2 - p_1) u(p_1) \,, \end{aligned} \tag{4.69}$$

which is depicted by the diagram shown in Figure 4.7. Similarly, we can write the

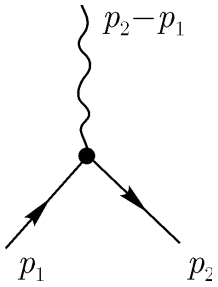


Figure 4.7

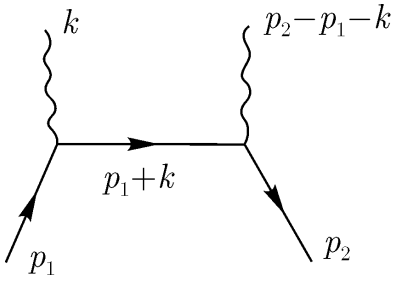


Figure 4.8

second-order matrix element for transition from the state with p_1 into the state with p_2 as

$$M_2 = (-ie)^2 \int d^4x \int d^4y \, \bar{\varphi}_0(x)\widehat{A}(x)K_+(x,y)\widehat{A}(y)\psi_0(y)\,. \tag{4.70}$$

Then, after the substitution of the Fourier transforms of $\widehat{A}$ and K_+ from (4.68), (4.38), and (4.67) we obtain

$$\begin{aligned} M_2 = -ie^2(2\pi)^4 \int d^4p \int d^4k_1 \int d^4k_2 \delta(p_2 - k_1 - p) \\ \times \delta(p_1 + k_2 - p)\bar{u}(p_2)\hat{a}(k_1)\frac{1}{\hat{p} - m}\hat{a}(k_2)u(p_1) \end{aligned} \tag{4.71}$$

or

$$M_2 = -ie^2(2\pi)^4 \int d^4k \, \bar{u}(p_2)\hat{a}(p_2 - p_1 - k)\frac{1}{\hat{p}_1 + \hat{k} - m}\hat{a}(k)u(p_1)\,, \tag{4.72}$$

which can be associated with the diagram in Figure 4.8. It is clear that in a similar way we can write all terms of the higher orders of perturbation theory: as a result we obtain the following *diagrammatic rules* to describe electron scattering by the potential of an external electromagnetic field:

1. the atrix element of transition has the form $M = \bar{u}_2 N u_1$;
2. to each *virtual* electron state (internal electron line) with momentum p corresponds in N by a factor $\frac{i}{\hat{p}-m}$;
3. to each photon (wavy line) with momentum q corresponds in N by a factor of $-ie\hat{a}(q)$;
4. over all momenta q_i, which are not fixed by conservation laws, and which are obeyed in interaction *vertices*, we should perform integration $\frac{d^4q_i}{(2\pi)^4}$;

While calculating integrals, the integration contour for the time component of the momentum should br chosen according to the Feynman rule of pole encirclement: mass m in the integrand is replaced by $m \to m - i\delta \quad (\delta \to +0)$.

As a simple example of concrete calculating we shall briefly consider the scattering of an electron by the Coulomb field of an atomic nuclei (Rutherford scattering) with electric charge Ze. The potential of nuclei is given by

$$A_0 = V(r) = \frac{Ze}{r}\,. \tag{4.73}$$

Then,

$$\hat{a}(q) = \frac{1}{(2\pi)^4}\int d^4x\, e^{iqx}\gamma^\mu A_\mu = \frac{1}{(2\pi)^3}\delta(q_0)\gamma^0\int d^3\mathbf{x}\, e^{-i\mathbf{qr}}V(\mathbf{r}) = \frac{Ze}{2\pi^2\mathbf{q}^2}\delta(q_0)\gamma^0\,. \tag{4.74}$$

The first-order transition amplitude (4.69) has the form

$$M_1 = -2\pi i\,\delta(E_1 - E_2)\left[\bar{u}(\mathbf{p}_2)\frac{4\pi Ze^2}{|\mathbf{p}_2 - \mathbf{p}_1|^2}\gamma^0 u(\mathbf{p}_1)\right], \tag{4.75}$$

where E_1 and E_2 are the initial and final energies of an electron. From (4.75) we can see that $E_1 = E_2 = E$, i. e., we are dealing with elastic scattering (static potential!). The probability of scattering is determined by

$$|M_1|^2 = (2\pi)^2\left|\bar{u}(\mathbf{p}_2)\frac{4\pi Ze^2}{|\mathbf{p}_2 - \mathbf{p}_1|^2}\gamma^0 u(\mathbf{p}_1)\right|^2\delta(E_1 - E_2)\delta(0)\,. \tag{4.76}$$

Here we have written $[\delta(E_1 - E_2)]^2 = \delta(E_1 - E_2)\delta(0)$, which creates obvious problems. However, $\delta(0)$ should be interpreted, according to the well-known Fermi recipe, as

$$\delta(0) = \lim_{T\to\infty}\lim_{x\to 0}\frac{1}{2\pi}\int\limits_{-T/2}^{T/2} dt\, e^{ixt} = \lim_{T\to\infty}\frac{T}{2\pi}\,, \tag{4.77}$$

where T is the interaction time. Then we can define the transition probability per unit of time $w_{1\to 2}$ as:

$$w_{1\to 2} = 2\pi\left|\bar{u}(\mathbf{p}_2)\frac{4\pi Ze^2}{|\mathbf{p}_2 - \mathbf{p}_1|^2}\gamma^0 u(\mathbf{p}_1)\right|^2\delta(E_1 - E_2)\,. \tag{4.78}$$

Further calculations (assuming the nonpolarized nature of the beam of initial electrons) require the averaging over both initial spin polarizations of electrons and the summation over the final polarizations. To perform this averaging and summation there exists a certain well-developed mathematical apparatus which uses the explicit form of spinors $u(\mathbf{p})$ and the properties of Dirac's matrices. We pass over these details, which have been well-described in [60] and [6]. Finally, from (4.78) it is possible to obtain a relativistic version of the Rutherford formula (Mott formula) for a differential scattering crossection to the element of solid angle $d\Omega$ [60]:

$$\frac{d\sigma}{d\Omega} = \frac{Z^2e^4}{4\mathbf{p}^2v^2\sin^4\frac{\theta}{2}}\left(1 - v^2\sin^2\frac{\theta}{2}\right), \tag{4.79}$$

taking into account that $|\mathbf{p}_1 - \mathbf{p}_2| = 2|\mathbf{p}|\sin\theta/2$, where θ is the scattering angle, and we introduced the velocity $v = |\mathbf{p}|/E$.

4.5 The two-particle problem

As we have seen above, in the Lorentz gauge the Maxwell equation for potentials acquires the form

$$\Box A_\mu = 4\pi j_\mu \,. \tag{4.80}$$

This equation is easily solved with the help of Green's function D_+, which is defined by the following equation:

$$\Box_2 D_+(2,1) = 4\pi\delta(2,1)\,. \tag{4.81}$$

Making some obvious Fourier transformations, we have

$$D_+(2,1) = -\frac{4\pi}{(2\pi)^4}\int d^4k\, e^{-ik(x_2-x_1)}\frac{1}{k^2+i\delta}\,, \qquad \delta\to+0\,. \tag{4.82}$$

In fact, up to a constant and sign, this expression coincides with the integral I_+ from (4.39) if we put $m = 0$ and use the rule (4.43). Now we can write the solution of (4.80) in the almost obvious form

$$A_\mu(2) = \int d^4x_1 D(2,1) j_\mu(1)\,. \tag{4.83}$$

Here the possibly inhomogeneous term is absent, which corresponds to the boundary condition of the absence of free electromagnetic radiation at $t = \pm\infty$ (i. e., there are no solutions $A^{(0)}_\mu$ equation $\Box A_\mu = 0$, which can always be added to the right-hand side of equation (4.83)).

Consider now the case of two charged (interacting!) fermions. Each of the particles is the source of an electromagnetic field which acts upon the motion of the other particle. As a result of this interaction, each particle is scattered by the other. Let us write an expression for the current, corresponding to the transition of electron "a" from the state $u_a(p_1)e^{-ip_1x}$ to the state $u_a(p_2)e^{-ip_2x}$:

$$j^\mu(x) = e\bar{u}_a(p_2)\gamma^\mu_a u_a(p_1)e^{i(p_2-p_1)x}\,. \tag{4.84}$$

In accordance with equation (4.83) this current creates electromagnetic potential at the space-time point x:

$$\begin{aligned} A^\mu(x) &= e\int d^4x' D_+(x-x')e^{i(p_2-p_1)x'}\bar{u}_a(p_2)\gamma^\mu_a u_a(p_1) \\ &= -4\pi e\int d^4k\frac{1}{k^2+i\delta}e^{-ikx}\delta(k+p_2-p_1)\bar{u}_a(p_2)\gamma^\mu_a u_a(p_1)\,. \end{aligned} \tag{4.85}$$

This potential acts upon the motion of the second electron "b". According to equation (4.66), the first-order matrix element, corresponding to the transition of electron "b" from the state with 4-momentum q_1 into the state with 4-momentum q_2, induced

by potential (4.85), has the following form:

$$M = -ie\int d^4x\bar{u}_b(q_2)e^{iq_2x}\gamma_b^\mu A_\mu(x)u_b(q_1)e^{-iq_1x}$$
$$= 4\pi ie^2(2\pi)^4\frac{\bar{u}_b(q_2)\gamma_b^\mu u_b(q_1)\bar{u}_a(p_2)\gamma_{a\mu}u_a(p_1)}{|p_1-p_2|^2}\delta(p_1+q_1-p_2-q_2)\,. \tag{4.86}$$

Consider now the propagator for a system of two particles (two-particle Green's function). In a nonrelativistic approximation the system of two particles is described by the Schroedinger wave function $\psi(\mathbf{x}_a,\mathbf{x}_b,t)$, and, as in the previous case of a single particle, we can define the propagator $K(\mathbf{x}_a,\mathbf{x}_b,t;\mathbf{x}'_a,\mathbf{x}'_b,t')$, which determines the probability amplitude of particle the "a" transition from point $\mathbf{x}'_a$ at the time moment t' to point $\mathbf{x}_a$ at the moment t, while particle "b" propagates from point $\mathbf{x}'_b$ at the moment t' to point $\mathbf{x}_b$ at the moment t. If the particles do not interact, we obviously have

$$K(\mathbf{x}_a,\mathbf{x}_b,t;\mathbf{x}'_a,\mathbf{x}'_b,t') = K_{0a}(\mathbf{x}_at;\mathbf{x}'_at')K_{0b}(\mathbf{x}_bt;\mathbf{x}'_bt')\,, \tag{4.87}$$

where K_{0a} and K_{0b} are the propagators of the free particles "a" and "b". In the absence of any interaction we can also define a more general two-particle Green's function with different time moments for particles in initial and final states:

$$K_0(3,4;1,2) = K_{0a}(3,1)K_{0b}(4,2)\,. \tag{4.88}$$

Equation (4.86) can now be considered as the matrix element, appearing due to the first-order correction $K^{(1)}$ to the propagator of two free particles, which is written as

$$K_+(3,4;1,2) =$$
$$-ie^2\int d^4x_5\int d^4x_6K_{+a}(3,5)\gamma_a^\mu K_{+a}(5,1)D_+(5,6)K_{+b}(4,6)\gamma_{b\mu}K_{+b}(6,2) \tag{4.89}$$

and can be represented by the Feynman diagram shown in Figure 4.9. In this expression, D_+ can be considered as the propagator of the *virtual* photon. In fact, our derivation is not completely satisfactory, because we did not quantize the electromagnetic field itself. However, we shall see later that the same result is reproduced in the rigorous theory.

In momentum representation we can rewrite the previous expressions in a more transparent way. Assuming the validity of the Lorentz condition $\frac{\partial A_\mu}{\partial x^\mu} = 0$, and differentiating (4.85) over x_μ, we get

$$\bar{u}_a(p_2)\gamma_a^\mu k_\mu u_a(p_1) = \bar{u}_a(p_2)(\gamma_a^0k^0 - \boldsymbol{\gamma}\cdot\mathbf{k})u_a(p_1) = 0\,. \tag{4.90}$$

This relation is satisfied, because, due to the presence of δ-function in (4.85), we have $\hat{k} = \hat{p}_1 - \hat{p}_2$, while $u(p_2)$ and $u(p_1)$ are free-particle spinors, so that

$$\bar{u}_1(p_2)(\hat{p}_1-\hat{p}_2)u(p_1) = \bar{u}(p_2)[(\hat{p}_1-m)-(\hat{p}_2-m)]u(p_1) = 0\,. \tag{4.91}$$

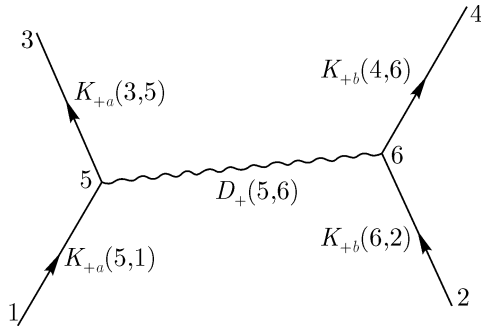


Figure 4.9

Thus, everywhere in (4.86) where the γ_a^0 matrix is present we can use the relation

$$\gamma_a^0 \cdot k_0 - \boldsymbol{\gamma}_a \cdot \mathbf{k} = 0 \tag{4.92}$$

and express γ_a^0 as

$$\gamma_a^0 = \gamma_{al} \left(\frac{|\mathbf{k}|}{k_0} \right), \tag{4.93}$$

where γ_l is the γ-matrix "projected on the direction of the propagation" $\mathbf{k}$ (let us stress this the for virtual photon $k_0 \neq |\mathbf{k}|$). Thus, denoting the "transverse" components of γ as γ_t^i, we can rewrite $\frac{\gamma_a^\mu \gamma_b^\mu}{k^2}$ in (4.86) as

$$\frac{\gamma_a^\mu \gamma_b^\mu}{k^2} = \frac{\gamma_a^0 \gamma_b^0 - \gamma_a^l \gamma_b^l - \sum_{i=1}^2 \gamma_{at}^i \gamma_{bt}^i}{k_0^2 - \mathbf{k}^2} = \frac{\gamma_a^0 \gamma_b^0 \left(1 - \frac{k_0^2}{|\mathbf{k}|^2}\right) - \sum_{i=1}^2 \gamma_{at}^i \gamma_{bt}^i}{k_0^2 - \mathbf{k}^2}$$
$$= -\frac{\gamma_a^0 \gamma_b^0}{\mathbf{k}^2} - \frac{\sum_{i=1}^2 \gamma_{at}^i \gamma_{bt}^i}{k_0^2 - \mathbf{k}^2}. \tag{4.94}$$

The first term in this expression describes in (4.86) and (4.89) the instantaneous Coulomb interaction of two electrons, while the second one takes into account the transversal quanta responsible for the retarded magnetic interaction of particles. The appearance of instantaneous interaction is connected to the noncovariant separation of initially the covariant interaction (4.94) into two terms: in fact the first term gives the main contribution in the limit of small velocities, while the second produces corrections to the instantaneous Coulomb interaction.

Up to now we have not taken into account that electrons are identical particles complying to the Pauli principle. This can be taken into account requiring antisymmetry of the wave function of the particle system, which can be achieved by introducing the two-particle propagator $K(3,4;1,2) - K(4,3;1,2)$, which describes the transition of two particles from points 1 and 2 to points 3 and 4, including the *exchange* process. Thus, instead of (4.86), we obtain the matrix element for the scattering of two identical

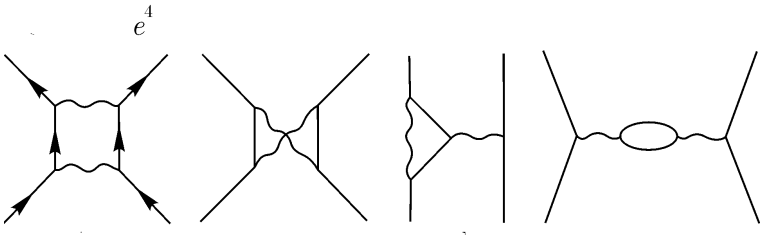


Figure 4.10

particles, in the first order over interaction, in the following form:

$$M = 4\pi i e^2 (2\pi)^4 \left\{ \frac{\bar{u}_b(q_2)\gamma_b^\mu u_b(q_1)\bar{u}_a(p_2)\gamma_{a\mu}u_a(p_1)}{|p_1 - p_2|^2} - \frac{\bar{u}_b(p_2)\gamma_b^\mu u_b(q_1)\bar{u}_a(q_2)\gamma_{a\mu}u_a(p_1)}{|q_1 - p_2|^2} \right\} \delta(p_1 + q_1 - p_2 - q_2)\,,$$

which determines e. g., the cross-section of the so-called Möller scattering.

In higher orders of perturbation theory over interaction, an infinite number of corrections appear which correspond to the exchange of a larger and larger numbers of virtual photons between interacting particles and particle self-interactions. All such processes are described by Feynman diagrams, which correspond to the appropriate mathematical expressions. Examples of diagrams of the order of e^4 are shown in Figure 4.10. Additional Feynman diagram rules for two-particle scattering are formulated as follows:

1. the probability amplitude for the radiation of a virtual photon is given by $e\gamma^\mu$, which is attributed to an interaction point (vertex) on the diagram;
2. the probability amplitude for a photon transition (propagator, wavy line) from point 1 to point 2 is given by $D_+(2,1)$ or, in momentum representation, $-\frac{4\pi}{k^2+i\delta}$.

Let us limit ourselves to the so-called "ladder" diagrams (with no intersections of interaction lines) shown in Figure 4.11. Introducing the probability amplitude $\varphi(x_1, x_2)$ to find two particles at points x_1 and x_2 after the exchange of n virtual photons, we can write the same probability amplitude after the exchange of the next $(n+1)$-th photon as

$$\varphi_{n+1}(1,2) = -ie^2 \int d^4x_3 \int d^4x_4 K_{+a}(1,3)\gamma_a^\mu K_{+b}(2,4)\gamma_{b\mu} D_+(3,4)\varphi_n(3,4)\,. \tag{4.95}$$

Then the total probability amplitude in the ladder approximation can be written as

$$\psi(x_1, x_2) = \sum_{n=0}^{\infty} \varphi_n(x_1, x_2) \tag{4.96}$$

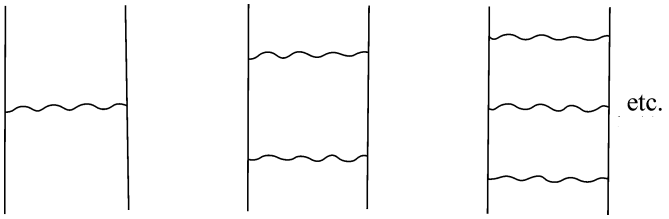


Figure 4.11

and, accordingly,

$$\psi(2,1) = \varphi_0(2,1) - ie^2 \int d^4x_3 \int d^4x_4 K_{+a}(1,3)\gamma_a^\mu K_{+b}(2,4)\gamma_{b\mu} D_+(3,4)\psi(3,4)\,, \tag{4.97}$$

where $\varphi_0(2,1)$ is the wave function, satisfying Dirac's equation for the free particle (by both variables). Applying Dirac's differential operators for the "a" and "b" particles to both sides of equation (4.97), we obtain the differential equation for $\psi(2,1)$

$$(i\hat{\nabla}_a - m)(i\hat{\nabla}_b - m)\psi(2,1) = ie^2\gamma_a^\mu \gamma_{b\mu} D_+(2,1)\psi(2,1)\,. \tag{4.98}$$

This is the so-called Bethe–Salpeter equation (in a ladder approximation), which is the relativistic wave equation for a two-particle system. In principle, it allows the complete analysis of the bound state problem in such a system, e. g., the study of the formation and of the energy spectrum of positronium.

Chapter 5

Scattering matrix

5.1 Scattering amplitude

Most experiments in high-energy physics (physics of elementary particles) are essentially scattering experiments: studies of reactions between particles and their decays. Particles usually interact at very small distances and during very short time intervals (inside a target or at crossings of accelerator beams), while practically free reaction products are registered in detector systems which are placed rather far from the space region where particles interact with each other, producing these reaction products. Thus, we are usually dealing with a rather general scattering problem: knowing the initial state of a system of free particles, we have to find the probability of different final states, which are also the sets of free particles, produced as a result of interactions.

Let $|i\rangle$ be some initial state. The result of an interaction can be represented by a superposition

$$\sum_f |f\rangle\langle f|S|i\rangle\,, \tag{5.1}$$

where the summation is performed over all possible final states $|f\rangle$. The coefficients $S_{fi} = \langle f|S|i\rangle$ form the so-called *scattering matrix* or *S-matrix*[1]. Now $|S_{fi}|^2$ gives the probability of transition $i \to f$. In the absence of interactions, the S-matrix is obviously a unit matrix. Then it is convenient to separate this unit part and write

$$S_{fi} = \delta_{fi} + i(2\pi)^4\delta(P_f - P_i)T_{fi}\,, \tag{5.2}$$

where the δ-functions simply express the conservation of the 4-momentum. For non-diagonal elements we simply have

$$S_{fi} = i(2\pi)^4\delta(P_i - P_f)T_{fi}\,. \tag{5.3}$$

Here T_{fi} is called the *scattering amplitude*.

While calculating the square of (5.2) we encounter the badly defined square of the δ-function, which expresses the 4-momentum conservation law. The correct way to proceed is to introduce the Fourier transform:

$$\delta(P_f - P_i) = \frac{1}{(2\pi)^4}\int d^4x e^{i(P_f - P_i)x}\,. \tag{5.4}$$

[1] The notion of scattering matrix was first introduced by Heisenberg, who suggested considering it as the most fundamental characteristic of elementary particle interactions.

and dealing with the second such integral perform calculations for $P_f = P_i$, but with integration extended to some large, but finite, volume V and time interval T, which gives $VT/(2\pi)^4$. Thus, we write

$$|S_{fi}|^2 = (2\pi)^4 \delta(P_f - P_i)|T_{fi}|^2 VT\,. \tag{5.5}$$

Now we can introduce the well-defined probability of transition in unit time (and in finite volume)[2]:

$$w_{fi} = (2\pi)^4 \delta(P_f - P_i)|T_{fi}|^2 V\,. \tag{5.6}$$

Free particles are described by appropriate plane waves with amplitudes u, representing bispinors for Dirac fermions, 4-vectors for photons, etc. Then we have

$$T_{fi} = u_1^* u_2^* \cdots Q u_1 u_2 \cdots\,, \tag{5.7}$$

where Q is some matrix over indices of the wave function amplitudes of all particles.

Let us consider the most important cases, where there are only one or two particles in the initial state, i. e., decays of single particles or collisions of two particles. Let us start from decays. The single particle can decay into several other particles with momenta $\mathbf{p}'_a$, belonging to an element of phase space $\prod_a d^3\mathbf{p}'_a$ (a numerates here the particles in the final state). The number of states in this elementary phase space volume is $\prod_a \frac{Vd^3\mathbf{p}'_a}{(2\pi)^3}$, and we have to multiply (5.6) by this number to obtain the probability of transition into the final states:

$$dw = (2\pi)^4 \delta(P_f - P_i)|T_{fi}|^2 V \prod_a \frac{Vd^3\mathbf{p}'_a}{(2\pi)^3}\,. \tag{5.8}$$

Everywhere we use normalization by a "single particle in volume V", so that the wave functions of all particles contain the factor of $\frac{1}{\sqrt{2\varepsilon_{\mathbf{p}} V}}$, where $\varepsilon_{\mathbf{p}}$ is the particle energy. It is convenient to move these factors into the scattering amplitude and write the wave functions in the following as

$$\psi = ue^{-ipx}\,, \qquad \bar{u}u = 2m \qquad \text{(electrons)}\,, \tag{5.9}$$

$$A_\mu = \sqrt{4\pi} e_\mu e^{-ikx}\,, \quad e_\mu e^{*\mu} = -1\,, \quad e_\mu k^\mu = 0 \quad \text{(photons)} \tag{5.10}$$

etc., rewriting the scattering amplitude via the new amplitude M_{fi}, defined as

$$T_{fi} = \frac{M_{fi}}{(2\varepsilon_1 V \ldots 2\varepsilon'_1 V \ldots)^{1/2}}\,, \tag{5.11}$$

where the denominator contains one factor of $(2\varepsilon_i V)$ per each initial and final particle. Then, the decay probability is written as

$$dw = (2\pi)^4 \delta(P_f - P_i)|M_{fi}|^2 \frac{1}{2\varepsilon} \prod_a \frac{d^3\mathbf{p}'_a}{(2\pi)^3 2\varepsilon'_a}\,, \tag{5.12}$$

[2] As we already noted above, this recipe was first proposed by Fermi.

where ε is an energy of the decaying particle. As should be expected, all normalization volumes in (5.12) has cancelled. If among final particles we have N identical ones, the phase volume of the final states should be divided by $N!$, to account for their permutations, producing the same state.

Consider in more detail the case of the decay into two particles with momenta $\mathbf{p}'_1$, $\mathbf{p}'_2$ and energies ε'_1, ε'_2. In the rest frame of the decaying particle, $\mathbf{p}'_1 = -\mathbf{p}'_2 \equiv \mathbf{p}'$, $\varepsilon'_1 + \varepsilon'_2 = m$. Then

$$dw = \frac{1}{(2\pi)^2}|M_{fi}|^2 \frac{1}{2m}\frac{1}{4\varepsilon'_1\varepsilon'_2}\delta(\mathbf{p}'_1 + \mathbf{p}'_2)\delta(\varepsilon'_1 + \varepsilon'_2 - m)d^3\mathbf{p}'_1 d^3\mathbf{p}'_2\,. \tag{5.13}$$

First, the δ-function here disappears after integration over $d^3\mathbf{p}'_2$. Then we rewrite $d^3\mathbf{p}'_1$ as

$$d^3\mathbf{p}' = |\mathbf{p}'|^2 d|\mathbf{p}'|d\Omega = |\mathbf{p}'|d\Omega' \frac{\varepsilon'_1\varepsilon'_2 d(\varepsilon'_1 + \varepsilon'_2)}{\varepsilon'_1 + \varepsilon'_2}\,, \tag{5.14}$$

where we have taken into account $\varepsilon'^2_1 - m_1^2 = \varepsilon'^2_2 - m_2^2 = |\mathbf{p}'|^2$. Then, integrating over $d(\varepsilon'_1 + \varepsilon'_2)$, we get rid of the second δ-function in (5.13). Accordingly, for the decay probability into an element of the solid angle $d\Omega'$ we get

$$dw = \frac{1}{32\pi^2 m^2}|M_{fi}|^2|\mathbf{p}'|d\Omega'\,. \tag{5.15}$$

Let us now consider the collision of two particles with momenta $\mathbf{p}_1$, $\mathbf{p}_2$ and energies ε_1, ε_2, producing in the final state some set of particles with momenta $\mathbf{p}'_a$ and energies ε'_a. Then,

$$dw = (2\pi)^4\delta(P_f - P_i)|M_{fi}|^2 \frac{1}{4\varepsilon_1\varepsilon_2 V}\prod_a \frac{d^3\mathbf{p}'_a}{(2\pi)^3 2\varepsilon'_a}\,. \tag{5.16}$$

The invariant (with respect to Lorentz transformations) scattering cross-section is obtained from (5.16) dividing by [33]

$$j = \frac{I}{V\varepsilon_1\varepsilon_2}\,, \quad \text{where} \quad I = \sqrt{(p_1 p_2)^2 - m_1^2 m_2^2}\,. \tag{5.17}$$

In the center of the mass frame we have $\mathbf{p}_1 = -\mathbf{p}_2 \equiv \mathbf{p}$, so that $I = |\mathbf{p}|(\varepsilon_1 + \varepsilon_2)$ and

$$j = \frac{|\mathbf{p}|}{V}\left(\frac{1}{\varepsilon_1} + \frac{1}{\varepsilon_2}\right) = \frac{v_1 + v_2}{V}\,, \tag{5.18}$$

which gives the current density of the colliding particles (v_1, v_2 are the particle velocities). Then finally

$$d\sigma = (2\pi)^4\delta(P_f - P_i)|M_{fi}|^2 \frac{1}{4I}\prod_a \frac{d^3\mathbf{p}'_a}{(2\pi)^3 2\varepsilon'_a}\,. \tag{5.19}$$

Let us drop the δ-functions for the case where there are two particles in the final state. Consider again scattering in the center of the mass frame. Let $\varepsilon = \varepsilon_1 + \varepsilon_2 = \varepsilon'_1 + \varepsilon'_2$ be the total energy

of the colliding particles, while $\mathbf{p}_1 = -\mathbf{p}_2 \equiv \mathbf{p}$ and $\mathbf{p'}_1 = -\mathbf{p'}_2 \equiv \mathbf{p'}$ are the initial and final momenta. Performing calculations similarly to the case of a decaying particle, we obtain

$$d\sigma = \frac{1}{64\pi^2}|M_{fi}|^2\frac{|\mathbf{p'}|}{|\mathbf{p}|\varepsilon^2}d\Omega' \tag{5.20}$$

For the case of elastic scattering we have $|\mathbf{p'}| = |\mathbf{p}|$. Let us introduce the kinematic invariant:

$$t \equiv (p_1 - p'_1)^2 = {m_1}^2 + {m'_1}^2 - 2(p_1 p'_1) = {m_1}^2 + {m'_1}^2 - 2\varepsilon_1\varepsilon'_1 + 2|\mathbf{p}_1||\mathbf{p'}_1|\cos\theta\,, \tag{5.21}$$

where θ is the scattering angle. In the center of the mass reference frame, $|\mathbf{p}_1| \equiv |\mathbf{p}|$ and $|\mathbf{p'}_1| \equiv |\mathbf{p'}|$ are determined only by the total energy ε, so that for its given value we have

$$dt = 2|\mathbf{p}||\mathbf{p'}|d\cos\theta\,. \tag{5.22}$$

Correspondingly, in equation (5.20) we can write

$$d\Omega' = -d\varphi d\cos\theta = \frac{d\varphi d(-t)}{2|\mathbf{p}||\mathbf{p'}|}\,, \tag{5.23}$$

where φ is the asimuthal angle of the vector $\mathbf{p'}_1$ with respect to $\mathbf{p}_1$. Further, for brevity, we write $d(-t)$ as dt and obtain

$$d\sigma = \frac{1}{64\pi}|M_{fi}|^2\frac{dt}{I^2}\frac{d\varphi}{2\pi}\,. \tag{5.24}$$

If the cross section does not depend on the asimuthal angle φ, we get

$$d\sigma = \frac{1}{64\pi}|M_{fi}|^2\frac{dt}{I^2}\,. \tag{5.25}$$

5.2 Kinematic invariants

Consider now the details of the kinematics of two particles scattering into two particles in the final state. The conservation law for a 4-momentum can be written in the form (where we do not predetermine which particles are in initial and which are in final state)

$$p_1 + p_2 + p_3 + p_4 = 0\,. \tag{5.26}$$

The scattering amplitude for this process can be represented by a graph (diagram), as shown in Figure 5.1, where directions of the arrows correspond to the momenta "entering" the amplitude (cf. equation (5.26)). Two of the momenta correspond to initial particles, while two others to particles in the final state (with moments $-p_a$). In these notations, two of p_a possess the time component $p_a^0 > 0$, while two others possess the time component $p_a^0 < 0$. For given types of particles participating in the scattering process, the squares the 4-momenta p_a^2 are determined by their masses: $p_a^2 = m_a^2$ (free particles always belong to their "mass surface"). Depending on the

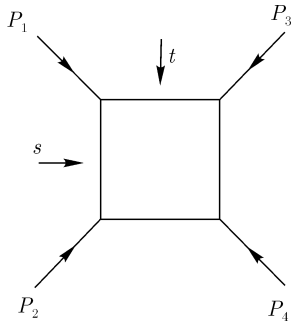


Figure 5.1

values, which acquire time components p_a^0, and also on the values of particles charges, the scattering amplitude in Figure 5.1 can describe three different reactions:

$$\begin{aligned} &(1) \qquad 1+2 \to 3+4 \quad (s\text{-channel}), \\ &(2) \qquad 1+\bar{3} \to \bar{2}+4 \quad (t\text{-channel}), \\ &(3) \qquad 1+\bar{4} \to \bar{2}+3 \quad (u\text{-channel}), \end{aligned} \tag{5.27}$$

where the bars denote the appropriate antiparticles. These scattering processes are called cross-reactions, whcih can be represented graphically as in the diagrams in Figure 5.2. We can also speak of three cross-channels of the same reaction, shown in Figure 5.1. We go from one reaction to another by changing the sign of the appropriate time component of momentum p_a^0 in (5.26):

$$\begin{aligned} &p_1^0 > 0, \quad p_2^0 > 0, \quad p_3^0 < 0, \quad p_4^0 < 0 \quad (s\text{-channel}), \\ &p_1^0 > 0, \quad p_2^0 < 0, \quad p_3^0 > 0, \quad p_4^0 < 0 \quad (t\text{-channel}), \\ &p_1^0 > 0, \quad p_2^0 < 0, \quad p_3^0 < 0, \quad p_4^0 > 0 \quad (u\text{-channel}), \end{aligned} \tag{5.28}$$

and also the signs of the charges. All initial and final states in (5.29) obviously possess positive energy. Transformation to the cross-channel reaction particle momentum in

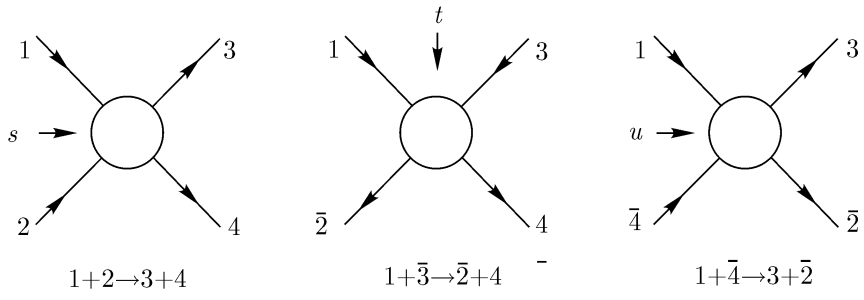


Figure 5.2

the initial state p_a is replaced by the antiparticle momentum $-p_a$ in the final state, with a corresponding change of charge. Due to the CPT-invariance of the theory we can also consider three CPT-conjugate reactions, which are obtained from (5.28) by replacement of all particles by antiparticles and the interchange of the initial and final states. If the theory is invariant with respect to charge conjugation C, we can add to these six reactions six more C-conjugate reactions, where all the particles are replaced by the corresponding antiparticles.

From the four 4-momenta entering the reaction we can construct two independent invariants. Due to (5.26) there are only three independent 4-vectors p_a; let these be p_1, p_2, p_3. From these we can construct six invariants: $p_1^2, p_2^2, p_3^2, p_1p_2, p_1p_3, p_2p_3$. The first three reduce to the corresponding squares of masses: m_1^2, m_2^2, m_3^2. The other three are connected by one relation, following from $(p_1 + p_2 + p_3)^2 = p_4^2 = m_4^2$. Usually the following symmetric notations are used, introducing three kinematic invariants:

$$\begin{aligned} s &= (p_1 + p_2)^2 = (p_3 + p_4)^2 , \\ t &= (p_1 + p_3)^2 = (p_2 + p_4)^2 , \\ u &= (p_1 + p_4)^2 = (p_2 + p_3)^2 , \end{aligned} \tag{5.29}$$

which are called *Mandelstam variables*. It is easily checked that

$$s + t + u = h \equiv m_1^2 + m_2^2 + m_3^2 + m_4^2 . \tag{5.30}$$

In channel (1), the invariant s represents the square of the total energy of colliding particles 1 and 2 in their center of the mass reference frame. In fact, for $\mathbf{p}_1 + \mathbf{p}_2 = 0$ we immediately obtain $s = (\varepsilon_1 + \varepsilon_2)^2$. In channel (2), a similar role is played by invariant t, while in channel (3) it is played by invariant u. Correspondingly, we are speaking about the s, t, and u reaction channels.

Let us consider in more detail the s-channel. Let

$$\begin{aligned} p_1 &= (\varepsilon_1, \mathbf{p}_s) , & p_2 &= (\varepsilon_2, -\mathbf{p}_s) , \\ p_3 &= (-\varepsilon_3, -\mathbf{p}'_s) , & p_4 &= (-\varepsilon_4, \mathbf{p}'_s) . \end{aligned} \tag{5.31}$$

Then it is easy to get

$$s = \varepsilon_s^2 , \quad \text{where} \quad \varepsilon_s = \varepsilon_1 + \varepsilon_2 = \varepsilon_3 + \varepsilon_4 , \tag{5.32}$$

$$\begin{aligned} 4s\mathbf{p}_s^2 &= [s - (m_1 + m_2)^2][s - (m_1 - m_2)^2] , \\ 4s\mathbf{p}'^2_s &= [s - (m_3 + m_4)^2][s - (m_3 - m_4)^2] , \end{aligned} \tag{5.33}$$

$$\begin{aligned} 2t &= h - s + 4\mathbf{p}_s\mathbf{p}'_s - \frac{1}{s}(m_1^2 - m_2^2)(m_3^2 - m_4^2) , \\ 2u &= h - s - 4\mathbf{p}_s\mathbf{p}'_s + \frac{1}{s}(m_1^2 - m_2^2)(m_3^2 - m_4^2) . \end{aligned} \tag{5.34}$$

In the case of elastic scattering ($m_1 = m_3$, $m_2 = m_4$) we have $|\mathbf{p}_s| = |\mathbf{p}'_s|$, so that $\varepsilon_1 = \varepsilon_3$ and $\varepsilon_2 = \varepsilon_4$. Then equations (5.33) simplify to

$$\begin{aligned} t &= -(\mathbf{p}_s - \mathbf{p}'_s)^2 = -2\mathbf{p}_s^2(1 - \cos\theta_s) \\ u &= -2\mathbf{p}_s^2(1 + \cos\theta_s) + (\varepsilon_1 - \varepsilon_2)^2 \end{aligned} \tag{5.35}$$

where θ_s is an angle between $\mathbf{p}_s$ and $\mathbf{p}'_s$, i. e., the scattering angle. Thus, in this case, invariant $-t$ represents the square of the transferred 3-momentum.

Similar expressions for other channels are obtained by the obvious changes of notation. The transformation to the t-channel is achieved by the replacements $s \leftrightarrow t$, $2 \leftrightarrow 3$, while the transformation to the u-channel is made by $s \leftrightarrow u$, $2 \leftrightarrow 4$.

If colliding particles are spinless, the scattering amplitude depends only on the kinematic invariants s, t, u, and in fact reduces to the single function

$$M_{fi} = f(s,t)\,. \tag{5.36}$$

For particles with spin, besides s, t, u there exist invariants which can be constructed from the amplitudes of the wave functions (bispinors, 4-vectors, 4-tensors etc.). Then the scattering amplitude has the form

$$M_{fi} = \sum_n f_n(s,t) F_n\,, \tag{5.37}$$

where F_n are invariants, linearly dependent on wave function amplitudes of all the colliding particles, as well as on their 4-momenta. The coefficients $f_n(s,t)$ are called *invariant amplitudes*.

5.3 Unitarity

The scattering matrix should be unitary: $SS^+ = 1$, or

$$(SS^+)_{fi} = \sum_n S_{fn} S^*_{in} = \delta_{fi}\,, \tag{5.38}$$

where n enumerates all the possible intermediate states. The unitarity condition (5.38) expresses the conservation of normalization and orthogonality of quantum states in scattering processes. In particular, the diagonal elements of (5.38) represent the sum of all transition probabilities from the fixed initial state to all the possible final states:

$$\sum_n |S_{ni}|^2 = 1\,. \tag{5.39}$$

Using (5.2) we obtain from (5.38)

$$T_{fi} - T^*_{if} = i(2\pi)^4 \sum_n \delta(P_f - P_n) T_{fn} T^*_{in}\,. \tag{5.40}$$

The left-hand side is linear, while the right-hand side is quadratic over T. If the interaction contains a small parameter, the left-hand side is "larger" than the right-hand side, and in the first approximation, neglecting the right-hand side, we can write

$$T_{fi} = T^*_{if}\,, \tag{5.41}$$

so that the T-matrix in this approximation is Hermitian.

Consider the collision of two particles. Only in the case of elastic scatterings are all the intermediate states in (5.40) also two-particle states. The summation over these states reduces to the integration over the intermediate momenta $\mathbf{p''}_1, \mathbf{p''}_2$ and the summation over spins (helicities) of both particles, which we denote λ'':

$$\sum_n = V^2 \int \frac{d^3\mathbf{p''}_1 d^3\mathbf{p''}_2}{(2\pi)^6} \sum_{\lambda''} \tag{5.42}$$

After dropping the δ-function, in a way similar to that used above, we may obtain the "two-particle" unitarity condition as

$$T_{fi} - T^*_{if} = i\frac{V^2}{(2\pi)^2} \sum_{\lambda''} \frac{|\mathbf{p}|}{\varepsilon} \int T_{fn} T^*_{in} \varepsilon''_1 \varepsilon''_2 d\Omega''\,, \tag{5.43}$$

where $\mathbf{p}$ is the momentum and ε is the total energy in the center of the mass reference frame. The normalization volume disappears after transition to amplitudes M_{fi}:

$$M_{fi} - M^*_{if} = \frac{i}{(4\pi)^2} \sum_{\lambda''} \frac{|\mathbf{p}|}{\varepsilon} \int M_{fn} M^*_{in} d\Omega''\,. \tag{5.44}$$

The diagonal element T_{ii} is called the zero-angle scattering amplitude. For this amplitude the unitarity condition takes the form

$$2\,\mathrm{Im}\, T_{ii} = (2\pi)^4 \sum_n |T_{in}|^2 \delta(P_i - P_n)\,. \tag{5.45}$$

The right-hand side here is proportional to the total cross section of all the scattering processes from the fixed initial state i, which we shall denote as σ_{tot}. In fact, summing (5.6) over f and dividing by the particle current density j, we obtain

$$\sigma_{tot} = \frac{(2\pi)^4 V}{j} \sum_n |T_{in}|^2 \delta(P_i - P_n)\,, \tag{5.46}$$

so that

$$\frac{2V}{j}\,\mathrm{Im}\, T_{ii} = \sigma_{tot}\,. \tag{5.47}$$

The normalization volume is cancelled after the transformation to $T_{ii} = M_{ii}/(2\varepsilon_1 V 2\varepsilon_2 V)$ (where $\varepsilon_1, \varepsilon_2$ are particle energies in the center of the mass reference frame) and substitution of j from (5.18) we get

$$\mathrm{Im}\, M_{ii} = 2|\mathbf{p}|\varepsilon\sigma_{tot}\,, \tag{5.48}$$

which is called the *optical theorem*.

Due to the CPT-theorem we have

$$T_{fi} = T_{\bar{i}\bar{f}}\,, \tag{5.49}$$

where $\bar{i}$ and $\bar{f}$ are states obtained from i and f by changing all the particles with their antiparticles. For diagonal elements,

$$T_{ii} = T_{\bar{i}\bar{i}}\,. \tag{5.50}$$

Then it follows from (5.45) and (5.48) that the total cross section of all possible scattering processes (with a fixed initial state) is the same for the reactions between particles and antiparticles. In particular, this means that the total lifetime (decay probability) of a particle and an antiparticle are equal.

During the period of the late 1950s and early 1960 where there was a certain dissatisfaction in quantum field theory, it was proposed to limit the theory of elementary particles to the analysis of the general properties of the S-matrix, such as unitarity and some general analyticity properties related to causality. This was the basis of the so-called analytical theory of the S-matrix [14]. Despite some successes and important theorems which are proved within this approach, it was insufficient for the construction of a complete dynamical theory of elementary particles. At the same time, as we shall see below, modern quantum field theory gives the well-developed formalism for the calculation of S-matrix via the standard perturbation theory approach.

Chapter 6

Invariant perturbation theory

6.1 Schroedinger and Heisenberg representations

Let us proceed to a systematic presentation of mathematical apparatus of perturbation theory over interactions in quantum field theory. It is well known that there exist two main formulations for equations of motion in quantum theory. In *Schroedinger representation* the quantum state at a given moment of time t is represented by the state vector $\Psi_S(t)$, containing the complete set of all possible results of measurements, applied to the system at this moment of time. The further evolution of the system is described by the time dependence of this state vector (wave function), described by the Schroedinger equation

$$i\hbar\frac{\partial\Psi_S}{\partial t} = H_S\Psi_S(t)\,. \tag{6.1}$$

In this representation, the operators of physical variables F_S do not depend on time; for all t they are the same: $dF_S/dt = 0$. At the same time, the average value of an operator

$$\langle F_S\rangle = \langle\Psi_S(t)|F_S|\Psi_S(t)\rangle \tag{6.2}$$

in the general case will depend on time as

$$i\hbar\frac{d}{dt}\langle F_S\rangle = \langle\Psi_S(t)|[F_S,H]|\Psi_S(t)\rangle\,. \tag{6.3}$$

Let us make the following time-dependent unitary transformation of vector $\Psi_S(t)$:

$$\Phi(t) = V(t)\Psi_S(t)\,, \tag{6.4}$$

where

$$V(t)V^+(t) = V^+(t)V(t) = 1\,,\quad V^+(t) = V^{-1}(t)\,. \tag{6.5}$$

Then the new state vector $\Phi(t)$ satisfies the equation[1]

$$i\hbar\frac{\partial\Phi(t)}{\partial t} = \left(i\hbar\frac{\partial V}{\partial t}V^{-1} + VH_SV^{-1}\right)\Phi(t)\,. \tag{6.6}$$

Let us choose $V(t)$ satisfying the equation

$$-i\hbar\frac{\partial V}{\partial t} = (VH_SV^{-1})V = VH_S\,. \tag{6.7}$$

[1] We have $i\hbar\frac{\partial\Phi(t)}{\partial t} = i\hbar\frac{\partial V}{\partial t}\Psi_S(t) + i\hbar V\frac{\partial\Psi_S}{\partial t} = i\hbar\frac{\partial V}{\partial t}V^{-1}\Phi(t) + VH_S\Psi_S = i\hbar\frac{\partial V}{\partial t}V^{-1}\Phi(t) + VH_SV^{-1}\Phi(t)$, which coincides with (6.6).

Then the transformed state vector will not depend on time, which is directly seen from (6.6). Due to the unitarity of $V(t)$, the average value of an operator F_S is expressed as

$$\begin{aligned}\langle F\rangle &= \langle\Psi_S(t)|F_S|\Psi_S(t)\rangle = \langle V(t)\Psi_S(t)|V(t)F_S\Psi_S(t)\rangle \\ &= \langle\Phi_H|V(t)F_SV^{-1}(t)|\Phi_H\rangle\,,\end{aligned} \tag{6.8}$$

where we have defined Φ_H as

$$\Phi_H = V(t)\Psi_S(t)\,, \tag{6.9}$$

and $V(t)$ satisfies equation (6.7). Let us define $F_H(t)$ as

$$F_H(t) = V(t)F_SV^{-1}(t)\,. \tag{6.10}$$

Then the time-dependent operator $F_H(t)$ has the same average value in the state defined by vector Φ_H, which the operator F_S has in the state defined by vector Ψ_S. Differentiating (6.5) over time, we have

$$\frac{dV(t)}{dt}V^+(t) + V(t)\frac{dV^+(t)}{dt} = 0\,. \tag{6.11}$$

Then from (6.7) and (6.10) we obtain for the time dependence of $F_H(t)$[2]

$$\frac{\partial F_H(t)}{\partial t} = \frac{\partial V}{\partial t}V^+F_H(t) + F_H(t)V\frac{\partial V^+}{\partial t} = \frac{i}{\hbar}[VH_SV^+, F_H(t)] = \frac{i}{\hbar}[H_H, F_H(t)]\,, \tag{6.12}$$

which represents the equation of motion for the operator of the physical variable in a *Heisenberg representation*. The Heisenberg-state vector Φ_H does not depend on time:

$$\frac{\partial\Phi_H}{\partial t} = 0\,. \tag{6.13}$$

We can assume that Φ_H just coincides with $\Psi_S(0)$ at $t = 0$.

6.2 Interaction representation

Consider once again the usual Schroedinger equation

$$i\hbar\frac{\partial\Phi(t)}{\partial t} = (H_0 + H_I)\Phi(t)\,, \tag{6.14}$$

where H_0 is the Hamiltonian of noninteracting fields (particles), while H_I is some interaction Hamiltonian. The state vector Φ in the absence of interactions, i. e., for $H_I = 0$, describes the motion of the given number of free particles with fixed momenta and spins. The operator H_I describes interactions of these particles.

[2] To obtain (6.12) we take into account that $\frac{\partial F_H}{\partial t} = \frac{\partial V}{\partial t}F_SV^{-1} + VF_S\frac{\partial V^{-1}}{\partial t} = \frac{\partial V}{\partial t}V^{-1}F_H + F_HV\frac{\partial V^{-1}}{\partial t}$.

Let us introduce the state vector

$$\Psi(t) = \exp\left(\frac{iH_0t}{\hbar}\right)\Phi(t)\,. \tag{6.15}$$

It is easy to see that $\Psi(t)$ satisfies the equation

$$i\hbar\frac{\partial\Psi(t)}{\partial t} = \exp\left(\frac{iH_0t}{\hbar}\right)H_I\exp\left(-\frac{iH_0t}{\hbar}\right)\Psi(t) \tag{6.16}$$

or

$$i\hbar\frac{\partial\Psi(t)}{\partial t} = H_I^{IR}(t)\Psi(t)\,, \tag{6.17}$$

where

$$H_I^{IR}(t) = \exp\left(\frac{iH_0t}{\hbar}\right)H_I\exp\left(-\frac{iH_0t}{\hbar}\right), \tag{6.18}$$

the operator of the interaction energy in this new representation. This operator explicitly depends on time, in contrast to the Schroedinger operator H_I. In general, an arbitrary operator $Q_{IR}(t)$ in this, the so-called *interaction representation*, is related to the Schroedinger operator Q_S as

$$Q_{IR}(t) = \exp\left(\frac{iH_0t}{\hbar}\right)Q_S\exp\left(-\frac{iH_0t}{\hbar}\right). \tag{6.19}$$

Now it immediately follows that in interaction representation the dependence of operators on time is determined by the Hamiltonian of free particles; differentiating (6.19) by t we obtain

$$i\hbar\frac{\partial Q_{IR}(t)}{\partial t} = [Q_{IR}(t), H_0]\,. \tag{6.20}$$

Note that $H_0^{IR} = H_0^S$. Thus, in interaction representation, field operators satisfy equations of motion of *free* fields[3], while the time-dependence of the state vector $\Psi(t)$ is determined, according to (6.17), only by the interaction energy. Interaction representation is quite convenient for the construction of perturbation theory.

Consider as an example the theory of Dirac fermions interacting with a scalar field. In Schroedinger representation the Hamiltonian of free fields has the form

$$H_0 = \int d^3\mathbf{r}\left[\bar\psi(x)\left(-i\boldsymbol{\gamma}\cdot\nabla + m\right)\psi(x) + \frac{1}{2}\left(\frac{\partial\varphi(x)}{\partial t}\right)^2 + \frac{1}{2}(\nabla\varphi(x))^2 + \frac{1}{2}m^2\varphi^2(x)\right], \tag{6.21}$$

and the interaction Hamiltonian (based on simplest principles of relativistic invariance) is written as

$$H_I = g\int d^3\mathbf{r}\bar\psi(x)\psi(x)\varphi(x)\,, \tag{6.22}$$

[3] In particular, this means that commutation relations for these operators are the same for arbitrary moments of time.

where g is the dimensionless coupling constant. After transformation to interaction representation, field operators $\varphi(x)$ and $\psi(x)$ satisfy the equations

$$(i\widehat{\nabla} - m)\psi_{IR}(x) = 0\,, \quad (\Box + m^2)\varphi_{IR}(x) = 0\,, \tag{6.23}$$

and equation (6.17) reduces to

$$i\hbar\frac{\partial\Psi(t)}{\partial t} = g\int_{ct=x_0} d^3\mathbf{r}\bar{\psi}_{IR}(x)\psi_{IR}(x)\varphi_{IR}(x)\Psi(t)\,. \tag{6.24}$$

Equation (6.17) can be generalized to covariant form. This is achieved by introduction of an arbitrary hypersurface in space-time instead of the hyperplane $t = const$. The only condition for this hypersurface to satisfy is that any vector normal to it $n_\mu(x)$ at an arbitrary point $\mathbf{r}$ should be time-like, i. e., $n_\mu(\mathbf{r})n^\mu(\mathbf{r}) > 0$. This means that no points on this hypersurface can be connected by a light signal, or that any two points on it should br separated by a space-like interval. Let us denote such surfaces as σ. At an arbitrary point $\mathbf{r}$ on this surface we can introduce time $t(\mathbf{r})$, which is called the local time. In the limit when this surface becomes just a plane, all points on it possess the same time $t = const$. Now we can generalize $\Psi(t)$ by introducing $\Psi[t(\mathbf{r})]$. The basic equation (6.17)

$$i\hbar\frac{\partial\Psi(t)}{\partial t} = H_I(t)\Psi(t) \tag{6.25}$$

can now be considered to be the result of the summation of an infinite number of equations obtained after the introduction of local time at each point of the space-like hypersurface. If the interaction Hamiltonian is expressed as the sum over small three-dimensional cells ΔV on the space-like hypersurface σ, i. e.,

$$H_I = \sum_\sigma \mathcal{H}_I(x)\Delta V\,, \tag{6.26}$$

the equation in a small cell surrounding the space-time point $\mathbf{r}, t(\mathbf{r})$ can be written as

$$i\hbar\frac{\partial\Psi[t(\mathbf{r})]}{\partial t(\mathbf{r})} = \mathcal{H}_I(x)\Delta V\Psi[t(\mathbf{r})]\,, \tag{6.27}$$

which directly generalizes equation (6.17). The variation of $\Psi(t)$, corresponding to a rigid infinitesimal translation of hypersurface $t = const$ as a whole, is determined by the integral $\int_t \mathcal{H}_I\, d^3\mathbf{r}$, so that it becomes clear that the variation of $\Psi[t(\mathbf{r})]$ relative to point x is determined by the interaction energy in $\mathcal{H}_I(x)\Delta V$ with an infinitesimal volume surrounding x. As the product $\Delta V\Delta t$ is a relativistic invariant, we may introduce the following invariant differentiation procedure. Consider a function defined on the space-like hypersurface $\Psi[t(\mathbf{r})] = \Psi(\sigma)$. Let us compare the values of this function on two space-like hypersurfaces σ and σ', which are infinitesimally different from each other in the vicinity of point x, as shown in Figure 6.1. Now we define the

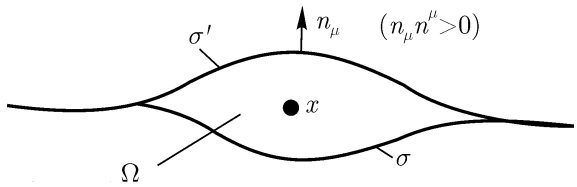


Figure 6.1

invariant operation $\delta/\delta\sigma(x)$ in the following way:

$$\frac{\delta\Psi(\sigma)}{\delta\sigma(x)} = \lim_{\Delta t\,\Delta V\to 0} \frac{\Psi[t(\mathbf{r})+\Delta t(\mathbf{r})]-\Psi[t(\mathbf{r})]}{c\int_{\Delta V} d^3\mathbf{r}\,\Delta t(\mathbf{r})}$$
$$= \lim_{\Delta t\,\Delta V\to 0} \frac{\Psi(\sigma')-\Psi(\sigma)}{c\,\Delta t(\mathbf{r})\,\Delta V} = \lim_{\Omega(x)\to 0} \frac{\Psi(\sigma')-\Psi(\sigma)}{\Omega(x)}\,, \tag{6.28}$$

where $\Omega(x)$ is the 4-volume between σ and σ'. Then, in the limit of $\Omega(x)\to 0$ equation (6.27) can be rewritten in the form of the so-called Tomonaga–Schwinger equation:

$$i\hbar c\frac{\delta\Psi(\sigma)}{\delta\sigma(x)} = \mathcal{H}_I(x)\Psi(\sigma)\,. \tag{6.29}$$

This equation is covariant, as $\mathcal{H}_I(x)$ is relativistic invariant (scalar), and we do not need any specific Lorentz reference frame to define the space-like surface σ. Thus, the Tomonaga–Schwinger equation is written with no reference to any system of coordinates. However, in the following we shall mainly deal with equation (6.17), written in the fixed reference frame.

6.3 S-matrix expansion

The solution of equation of motion in interaction representation (6.17) can be written in integral form as

$$\Psi(t) = \Psi(t_0) - \frac{i}{\hbar}\int_{t_0}^{t} dt' H_I(t')\Psi(t')\,. \tag{6.30}$$

Here we take into account the initial condition: for $t = t_0$ the function Ψ reduces to $\Psi(t_0)$.

Let us write the relation between $\Psi(t)$ and $\Psi(t_0)$ as[4]

$$\begin{aligned}\Psi(t) &= U(t,t_0)\Psi(t_0)\,,\\ \Psi(t_0) &= U^{-1}(t,t_0)\Psi(t)\,,\\ U(t_0,t_0) &= 1\,,\end{aligned} \tag{6.31}$$

[4] The formalism presented below was developed by Dyson.

where $U(t,t_0)$ is a unitary (conserving normalization!) operator of evolution. Then

$$S = U(+\infty,-\infty) \tag{6.32}$$

defines the S-matrix (scattering matrix), which determines all the possible changes of the system states due to interaction:

$$\Psi(+\infty) = S\Psi(-\infty)\,, \tag{6.33}$$

where $\Psi(-\infty)$ and $\Psi(+\infty)$ are asymptotic state vectors of the system, in particular, the asymptotic forms of incoming and scattered waves in a typical scattering process.

The operator $U(t,t_0)$ satisfies the following differential equation, which is obvious from (6.17):

$$i\hbar\frac{\partial U(t,t_0)}{\partial t} = H_I(t)U(t,t_0)\,. \tag{6.34}$$

Similarly,

$$-i\hbar\frac{\partial U^+(t,t_0)}{\partial t} = U^+(t,t_0)H_I(t)\,, \tag{6.35}$$

as $H_I(t)$ is Hermitian. From these equations it immediately follows that

$$\frac{\partial}{\partial t}(U^+(t,t_0)U(t,t_0)) = 0\,, \tag{6.36}$$

which is equivalent to

$$U^+(t,t_0)U(t,t_0) = 1\,. \tag{6.37}$$

To prove unitarity we still have to show that

$$U(t,t_0)U^+(t,t_0) = 1\,. \tag{6.38}$$

The evolution operator satisfies the group property

$$U(t,t_1)U(t_1,t_0) = U(t,t_0)\,. \tag{6.39}$$

In fact, from

$$\Psi(t) = U(t,t_1)\Psi(t_1)\,, \quad \Psi(t_1) = U(t_1,t_0)\Psi(t_0) \tag{6.40}$$

it follows that

$$\Psi(t) = U(t,t_0)\Psi(t_0) = U(t,t_1)U(t_1,t_0)\Psi(t_0)\,, \tag{6.41}$$

which is necessary to satisfy (6.39). If in (6.39) we set $t = t_0$, we get

$$U(t_0,t_1) = U^{-1}(t_1,t_0)\,. \tag{6.42}$$

From $U(t_0,t_1)U(t_1,t_0) = 1$, multiplying it from the left side by $U^+(t_0,t_1)$ and using (6.37), we obtain

$$U(t_1,t_0) = U^+(t_0,t_1) = U^{-1}(t_0,t_1)\,, \tag{6.43}$$

which proves the unitarity of the evolution operator.

It follows directly from group property (6.39) that any transition of the system during the finite time interval can be represented by the multiplication of evolution operators, corresponding to infinitesimal transitions:

$$U(t,t') = U(t,t_1)U(t_1,t_2)\cdots U(t_{n-1},t_n)U(t_n,t')\,, \tag{6.44}$$

where $U(t_j,t_{j+1})$ corresponds to an infinitesimal transformation from time moment t_j to t_{j+1}.

The solution of equation (6.34) can obviously be written also in integral form:

$$U(t,t') = 1 - \frac{i}{\hbar}\int_{t'}^{t} d\tau H_I(\tau)U(\tau,t')\,. \tag{6.45}$$

Thus, for the infinitesimal time difference $t_j - t_{j+1}$ we have:

$$\begin{aligned} U(t_j,t_{j+1}) &= 1 - \frac{i}{\hbar}\int_{t_{j+1}}^{t_j} d\tau H_I(\tau)U(\tau,t_{j+1}) \\ &\approx 1 - \frac{i}{\hbar}\int_{t_{j+1}}^{t_j} dt' H_I(t')U(t_{j+1},t_{j+1}) = 1 - \frac{i}{\hbar}\int_{t_{j+1}}^{t_j} dt' H_I(t')\,. \end{aligned} \tag{6.46}$$

Increasing the number of time intervals (to infinity!) and regrouping the terms in (6.44) we obtain

$$\begin{aligned} U(t,t_0) = 1 &+ \left(\frac{-i}{\hbar}\right)\int_{t_0}^{t} dt_1 H_I(t_1) + \left(\frac{-i}{\hbar}\right)^2 \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 H_I(t_1)H_I(t_2) \\ &+ \left(\frac{-i}{\hbar}\right)^3 \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \int_{t_0}^{t_2} dt_3 H_I(t_1)H_I(t_2)H_I(t_3) + \cdots\,. \end{aligned} \tag{6.47}$$

Consider the integral, determining the n-th order of perturbation theory:

$$\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n H_I(t_1)H_I(t_2)\cdots H_I(t_n)\,. \tag{6.48}$$

Here integration is performed essentially over the whole time interval from t_0 to t, but with a limitation: the time moment t_j is earlier than $t_{j-1}(j \le n)$. Of course, in equation (6.48) we can arbitrarily rename the integration variables $t_1,\ldots,t_n \to t_{p_1},t_{p_2}\ldots t_{p_n}$, and the value of integral will not change. Making all permutations of variables $t_1,\ldots,t_n$, summing all the expressions obtained, and dividing by the number of permutations $n!$, we extend the integration over each of the variables to the whole time interval from t_0 to t. However, it is necessary to guarantee that operators $H_I(t_j)$ under the integral are placed from left to right in the order of the growth of the time arguments. This can be achieved defining the operator T-ordering, which acts on the operators, depending on time, and places them in chronological order, i. e., an operator

with a larger value of time in the product stands to the left of those with smaller times:

$$T\big(H_I(t_1)\cdots H_I(t_k)\big) = H_I(t_i)H_I(t_j)\cdots H_I(t_k) \quad \text{for} \quad t_i > t_j > \cdots > t_k\,, \tag{6.49}$$

which gives the definition of the chronological or T-product of the operators. Then, using the symmetry of the integrand (6.48) mentioned above, we get:

$$\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n H_I(t_1)H_I(t_2)\cdots H_I(t_n) = \frac{1}{n!}\int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_n T\big(H_I(t_1)H_I(t_2)\cdots H_I(t_n)\big)\,. \tag{6.50}$$

Let us consider in more detail the equivalence of these two forms of integral for the case $n = 2$. From the definition of the T-product we have

$$\int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 T\big(H_I(t_1)H_I(t_2)\big) = \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 H_I(t_1)H_I(t_2) + \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2 H_I(t_2)H_I(t_1)\,. \tag{6.51}$$

The integration region of the left-hand side is shown in Figure 6.2 as a square. On the other hand, in the first term in the right-hand side of (6.51), integration is extended over the region I (nondashed triangle), while in the second term integration is performed over the dashed region II. Changing the order of integration in the second integral,

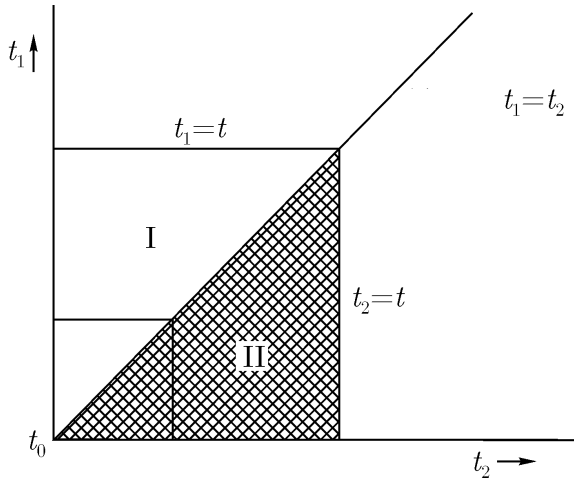


Figure 6.2

we shall first integrate over t_1; then the limits of integration change, and we get

$$\int_{t_0}^{t} dt_2 \int_{t_0}^{t_2} dt_1 H_I(t_2) H_I(t_1)\,. \tag{6.52}$$

Now, if we make the change of variables $t_1 \to t_2$ and $t_2 \to t_1$, equation (6.52) takes the form

$$\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 H_I(t_1) H_I(t_2)\,, \tag{6.53}$$

so that (6.51) reduces to

$$\int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 T\big(H_I(t_1) H_I(t_2)\big) = 2! \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 H_I(t_1) H_I(t_2)\,, \tag{6.54}$$

which proves the validity of (6.50) for the case of $n = 2$.

Thus, expansion (6.47) can be written as

$$\begin{aligned} U(t,t_0) &= 1 + \left(\frac{-i}{\hbar}\right) \int_{t_0}^{t} dt_1 T\big(H_I(t_1)\big) \qquad (6.55)\\ &\quad + \frac{1}{2!}\left(\frac{-i}{\hbar}\right)^2 \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 T\big(H_I(t_1) H_I(t_2)\big)\\ &\quad + \frac{1}{3!}\left(\frac{-i}{\hbar}\right)^3 \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \int_{t_0}^{t} dt_3 T\big(H_I(t_1) H_I(t_2) H_I(t_3)\big) + \cdots\\ &= \sum_{n=0}^{\infty} \frac{1}{n!}\left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_n T\big(H_I(t_1) H_I(t_2) \cdots H_I(t_n)\big)\,, \qquad (6.56) \end{aligned}$$

which can be rewritten as

$$U(t,t_0) = T\left\{\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} H_I(t')dt'\right)\right\}, \tag{6.57}$$

where we have performed the symbolic summation of the series (6.56), which reduces it to the so-called T-exponent.

It can be directly checked that the series (6.56) gives the solution of equation (6.34). Let us differentiate (6.56) by time t:

$$\frac{\partial U(t,t_0)}{\partial t} = \sum_{n=1}^{\infty} \frac{1}{n!}\left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_{n-1} n H_I(t) T\big(H_I(t_1) H_I(t_2) \cdots H_I(t_{n-1})\big). \tag{6.58}$$

While writing the right-hand side of (6.58) we used the symmetry of the integrand and the fact that operator $H_I(t)$ always depends on the time moment t, which is later, than $t_1, \ldots, t_{n-1}$. This allows us to move operator $H_I(t)$ outside the sign of the T-product, putting it to the left

of all the other factors. Then (6.58) can be rewritten as

$$
\begin{aligned}
i\hbar\frac{\partial U(t,t_0)}{\partial t} &= H_I(t)\sum_{n=1}^{\infty}\frac{1}{(n-1)!}\left(-\frac{i}{\hbar}\right)^{n-1}\int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2 \\
&\quad\cdots\int_{t_0}^{t}dt_{n-1}T\left(H_I(t_1)H_I(t_2)\cdots H_I(t_{n-1})\right) \\
&= H_I(t)\sum_{n=0}^{\infty}\frac{1}{n!}\left(-\frac{i}{\hbar}\right)^{n}\int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2\cdots\int_{t_0}^{t}dt_nT\left(H_I(t_1)H_I(t_2)\cdots H_I(t_n)\right) \\
&= H_I(t)U(t,t_0)\,, \qquad (6.59)
\end{aligned}
$$

which proves the desired result!

Recalling that

$$H_I(t) = \int d^3\mathbf{r}\mathcal{H}_I(x)\,, \qquad (6.60)$$

we can rewrite (6.56) in the form, explicitly demonstrating its covariance,

$$
\begin{aligned}
U(t,t_0) &= \sum_{n=0}^{\infty}\frac{1}{n!}\left(-\frac{i}{\hbar}\right)^{n}\int_{t_0}^{t}d^4x_1\int_{t_0}^{t}d^4x_c \\
&\quad\cdots\int_{t_0}^{t}d^4x_nT\left(\mathcal{H}_I(x_1)\mathcal{H}_I(x_2)\cdots\mathcal{H}_I(x_n)\right)\,, \qquad (6.61)
\end{aligned}
$$

where we have used $\int dt\int d^3\mathbf{r} = \int d^4x/c$. We can generalize (6.61), introducing integration limits at space-like surfaces σ and σ', then $U(\sigma,\sigma')$ will be explicitly invariant, as both $\mathcal{H}_I$ and the volume element d^4x are 4-scalars.

The important point in the justification of the above formalism is the so-called *adiabatic hypothesis*. Following the definition of the S-matrix we have to tend the initial moment of time t_0 to $-\infty$ and the final moment t to $+\infty$. However, we have to be cautious: for the n-th order term in expansion (6.56), this can be done in $n!$ ways for each of the limits. Dyson proposed to overcome this difficulty by introducing the convergence factor $e^{-\lambda|t|}$, multiplying the interaction Hamiltonian, with $\lambda \to 0$ at the end of the calculations. This procedure is equivalent to an averaging procedure over all possible $n!$ ways to perform the limit of $t \to \pm\infty$. Assuming the validity of this adiabatic hypothesis, we can consider the initial- and final-state wave functions as eigenstates of the "free" Hamiltonian H_0; these are usually called the wave functions of "bare" particles. Then, any scattering process is considered as consisting of the following stages.

1. At time $t = -\infty$ the system is in a state described by the wave function Φ, which is an eigenstate of the operator H_0. In this state, there is a given number of particles with fixed spins and momenta, and these particles are separated from each other and noninteracting. Vector Φ is constant and independent of the time ($H_I = 0$) vector in interaction representation.

2. The interaction is switched on adiabatically, so that the state with wave function Φ transforms into a state $\Psi(t_0) = U(t_0, -\infty)\Phi$, which is assumed to correspond to the real state of physical particles with the same momenta and spins. At this stage it is still assumed, that particles are well separated and do not interact with each other. However, switching on H_I induces self-interaction, so that the "bare" particles are "dressed" by virtual quanta and the particles become real physical particles which satisfy the condition $p^2 = m^2$, where m is the observable physical mass.
3. Further on, the particles interact with each other, i. e., are scattered, transform into other type of particles, etc. After a long enough time $T = t - t_0$, particles separate again, but their states are now described by the wave function $\Psi(t) = U(t, t_0)\Psi(t_0)$, this state corresponding to the "dressed" (i. e., real physical) particles after scattering.
4. Then interaction is adiabatically switched of,f and the state with wave function $\Psi(t)$ transforms into a state with the wave function Φ', which corresponds to "bare" particles after scattering, and $\Phi' = U(\infty, t)\Psi(t)$.

Thus, the real scattering problem $\Psi(t_0) \to \Psi(t)$ is replaced by an "equivalent" problem which introduces the "bare" particles at $t = \pm\infty$. Consider the relation

$$\Psi(t) = U(t, t_0)\Psi(t_0) \tag{6.62}$$

which can be rewritten as

$$U^{-1}(\infty, t)\Psi' = U(t, t_0)U(t_0, -\infty)\Phi\,. \tag{6.63}$$

Now we have

$$\Phi' = U(\infty, t)U(t, t_0)U(t_0, -\infty)\Phi = U(\infty, -\infty)\Phi = S\Phi\,. \tag{6.64}$$

This means that Φ' at $t = +\infty$ is the wave function of "bare" particles which appear as a result of scattering from the state described by wave function Φ at $t = -\infty$.

The adiabatic hypothesis leads to results which are in excellent agreement with experiments. This may seem strange, as it is clear that interaction between real particles can not be "switched off" (adiabatically or in any other way). In this respect, quantum field theory is rather different from quantum mechanics, where we usually deal with potentials with finite radius (except the Coulomb case, but there we know the exact wave functions), so that in the scattering problem the wave functions of the initial and final states are really corresponding to free particles.

6.4 Feynman diagrams for electron scattering in quantum electrodynamics

In quantum electrodynamics (QED), interaction Hamiltonian density has the form

$$\mathcal{H}_I(x) = j_\mu(x)A^\mu(x)\,, \tag{6.65}$$

where j_μ is the current density of Dirac electrons, while A^μ is the vector-potential of the electromagnetic field. Then the scattering matrix is written as[5]

$$S = T\exp\left\{-ie\int d^4x j_\mu(x)A^\mu(x)\right\}\,, \tag{6.66}$$

where we returned to the system of units with $\hbar = c = 1$.

Let us consider some specific examples of the calculation of matrix elements of a scattering matrix. The current density operator j contains the product of two electronic ψ-operators. Thus, in the first order of perturbation theory only the processes involving three particles – two electrons and one photon – can appear, as shown by the diagram in Figure 6.3, similar to that of Figure 4.7. However, such processes with *free* particles are impossible because of energy and momentum conservation. In fact, if p_1 and p_2 are the 4-momenta of electrons, and k is the 4-momentum of a photon, conservation law is written as $k = p_2 - p_1$ or $k = p_1 + p_2$. However, these equalities are impossible, as for the real photon we always have $k^2 = 0$, while the square $(p_2 \pm p_1)^2$ is easily shown to be nonzero. Let us calculate $(p_2 \pm p_1)^2$ in the rest frame of one of the electrons, e. g., electron 1. Then $(p_2 \pm p_1)^2 = 2(m^2 \pm p_1 p_2) = 2(m^2 \pm \varepsilon_1\varepsilon_2 \mp \mathbf{p}_1\mathbf{p}_2) = 2m(m \pm \varepsilon_2)$, and because of $\varepsilon_2 > m$ we have $(p_2 + p_1)^2 > 0$ or $(p_2 - p_1)^2 < 0$.

Thus, the first nonzero matrix elements of the S-matrix can appear only in the second order of perturbation theory:

$$S^{(2)} = -\frac{e^2}{2!}\int d^4x \int d^4x' T\left(j^\mu(x)A_\mu(x)j^\nu(x')A_\nu(x')\right)\,. \tag{6.67}$$

As electron and photon operators in interaction representation commute with each other, (6.67) can be rewritten as

$$S^{(2)} = -\frac{e^2}{2!}\int d^4x \int d^4x' T\left(j^\mu(x)j^\nu(x')\right)T\left(A_\mu(x)A_\nu(x')\right)\,. \tag{6.68}$$

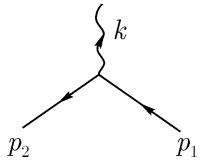


Figure 6.3

[5] In the following, here and in the next chapter, we mainly follow [6].

As a first example we consider the elastic scattering of two electrons. In the initial state we have two electrons with momenta p_1 and p_2, while in the final state we have two electrons with momenta p_3 and p_4. It is assumed that electrons are in some concrete spin states, but the spin indices in the following are dropped for brevity. We have to calculate the matrix element between the initial and final states with appropriate particles. As in both states photons are just absent, the required matrix element of the T-product of the photon operators is simply $\langle 0|\cdots|0\rangle$, where $|0\rangle$ is a photon vacuum. Accordingly, from (6.68) we obtain a tensor:

$$D_{\mu\nu}(x-x') = i\,\langle 0|TA_\mu(x)A_\nu(x')|0\rangle\,, \tag{6.69}$$

which is called the photon propagator or the photon Green's function.

From the T-product of electron operators in (6.68) the following matrix element appears:

$$\langle 34|Tj^\mu(x)j^\nu(x')|12\rangle\,, \tag{6.70}$$

where $|12\rangle$ and $|34\rangle$ denote states with two electrons with the appropriate momenta. This matrix element can also be written in the form of a vacuum average, if we use the relation

$$\langle 2|F|1\rangle = \langle 0|a_2 F a_1^+|0\rangle\,, \tag{6.71}$$

where F is an arbitrary operator, while a_1^+ and a_2 are operators of the creation of the 1-st and annihilation of 2-nd electrons. It is clear that instead of (6.70) we have to calculate

$$\langle 0|a_3 a_4 T(j^\mu(x)j^\nu(x'))a_2^+ a_1^+|0\rangle\,. \tag{6.72}$$

Each of the current operators is written as $j = \bar\psi\gamma\psi$, and the ψ-operators are represented by

$$\psi = \sum_{\mathbf{p}}(a_{\mathbf{p}}\psi_p + b_{\mathbf{p}}^+\psi_{-p})\,,\quad \bar\psi = \sum_{\mathbf{p}}(a_{\mathbf{p}}^+\bar\psi_p + b_{\mathbf{p}}\bar\psi_{-p})\,, \tag{6.73}$$

where ψ_p denotes the appropriate spinors (plane waves). The second terms here contain positron operators, which are irrelevant for the problem under consideration. Taking (6.73) into account, the product $j^\mu(x)j^\nu(x')$ is represented by the sum of the terms, each containing the product of two operators $a_{\mathbf{p}}$ and two $a_{\mathbf{p}}^+$, which are responsible for the annihilation of electrons 1 and 2 and the creation of electrons 3 and 4. It is clear that these should be operators a_1, a_2, a_3^+, a_4^+, which are "paired" (or "contracted") with the "external" operators a_1^+, a_2^+, a_3, a_4 according to an obvious equality:

$$\langle 0|a_p a_p^+|0\rangle = 1\,. \tag{6.74}$$

The operators just disappear, and only the c-numbers remain. Depending on which of ψ-operators provide a_1, a_2, a_3^+, a_4^+ for pairing (contraction) with external a_1^+, a_2^+, a_3,

a_4, equation (6.72) produces four terms like

$$a_3^{\cdot\cdot\cdot} a_4^{\cdot} (\bar\psi^{\cdot} \gamma^\mu \psi^{\cdot\cdot})(\bar\psi'^{\cdot\cdot\cdot} \gamma^\nu \psi'^{\cdot\cdot\cdot\cdot}) a_2^{+\cdot\cdot} a_1^{+\cdot\cdot\cdot\cdot} + a_3^{\cdot} a_4^{\cdot\cdot\cdot} (\bar\psi^{\cdot} \gamma^\mu \psi^{\cdot\cdot})(\bar\psi'^{\cdot\cdot\cdot} \gamma^\nu \psi'^{\cdot\cdot\cdot\cdot}) a_2^{+\cdot\cdot\cdot\cdot} a_1^{+\cdot\cdot}$$
$$+ a_3^{\cdot} a_4^{\cdot\cdot\cdot} (\bar\psi^{\cdot} \gamma^\mu \psi^{\cdot\cdot})(\bar\psi'^{\cdot\cdot\cdot} \gamma^\nu \psi'^{\cdot\cdot\cdot\cdot}) a_2^{+\cdot\cdot} a_1^{+\cdot\cdot\cdot\cdot} + a_3^{\cdot\cdot\cdot} a_4^{\cdot} (\bar\psi^{\cdot} \gamma^\mu \psi^{\cdot\cdot})(\bar\psi'^{\cdot\cdot\cdot} \gamma^\nu \psi'^{\cdot\cdot\cdot\cdot}) a_2^{+\cdot\cdot\cdot\cdot} a_1^{+\cdot\cdot}, \tag{6.75}$$

where $\psi = \psi(x)$ and $\psi' = \psi(x')$, and the same number of dots denote paired (contracted) fermion operators. Now, in each of these terms it is necessary to make permutations of the "paired" operators $a_1, a_2, \ldots$ from ψ, written as (6.73), to put them alongside their external partners $a_1^+, a_2^+, \ldots$, so that we can use (6.74) and obtain the vacuum average as a simple product of averages, corresponding to these pairings (contractions). Taking into account the anticommutativity of these operators (1,2,3,4 are different states!), we find, that the matrix element (6.70) is equal to [6]

$$\langle 34|T j^\mu(x) j^\nu(x')|12\rangle = (\bar\psi_4 \gamma^\mu \psi_2)(\bar\psi_3' \gamma^\nu \psi_1') + (\bar\psi_3 \gamma^\mu \psi_1)(\bar\psi_4' \gamma^\nu \psi_2')$$
$$- (\bar\psi_3 \gamma^\mu \psi_2)(\bar\psi_4' \gamma^\nu \psi_1') - (\bar\psi_4 \gamma^\mu \psi_1)(\bar\psi_3' \gamma^\nu \psi_2'), \tag{6.76}$$

where the ψ are not operators, but are the corresponding spinors (plane waves with momenta 1,2,3,4)! The total sign here is the subject of agreement; it depends on the order of placement of the "external" electron operators. The sign of a matrix element for the scattering of identical particles is, in general, arbitrary. The first and second terms in (6.76) (as well as the third and fourth) differ from each other only by the permutation of indices μ and ν and arguments x and x'. But such permutations do not change the matrix element (6.70), where the order of all factors is determined by the symbol of T-ordering. Thus, after multiplying (6.76) and (6.69), and integration over $d^4x d^4x'$, four terms from (6.76) give

$$S_{fi} = ie^2 \int d^4x d^4x' D_{\mu\nu}(x - x')[(\bar\psi_4 \gamma^\mu \psi_2)(\bar\psi_3' \gamma^\nu \psi_1') - (\bar\psi_4 \gamma^\mu \psi_1)(\bar\psi_3' \gamma^\nu \psi_2')] \tag{6.77}$$

Note that the factor of 2! cancelled! Taking into account that the electronic wave function here are plane waves, we can write the expression in square brackets in (6.77) as

$$(\bar u_4 \gamma^\mu u_2)(\bar u_3 \gamma^\nu u_1) e^{-i(p_2 - p_4)x - i(p_1 - p_3)x'}$$
$$- (\bar u_4 \gamma^\mu u_1)(\bar u_3 \gamma^\nu u_2) e^{-i(p_1 - p_4)x - i(p_2 - p_3)x'} =$$
$$\Big\{(\bar u_4 \gamma^\mu u_2)(\bar u_3 \gamma^\nu u_1) e^{-i[(p_2 - p_4) + (p_3 - p_1)]\xi/2}$$
$$- (\bar u_4 \gamma^\mu u_1)(\bar u_3 \gamma^\nu u_2) e^{-i[(p_1 - p_4) + (p_3 - p_2)]\xi/2}\Big\} e^{-i(p_1 + p_2 - p_3 - p_4)X}, \tag{6.78}$$

where we have introduced $\xi = x - x'$ and $X = \frac{1}{2}(x + x')$. The integration in (6.77) over $d^4x d^4x'$ is now replaced by $d^4\xi d^4X$. The integral over d^4X produces

[6] Due to the anticommutativity of fermion operators, the current operators $j(x)$ and $j(x')$, composed of pairs of these operators, commute, and we can drop the symbol T-product.

$$e^2(\overline{u}_4\gamma^\mu u_2)D_{\mu\nu}(k)(\overline{u}_3\gamma^\nu u_1)=$$

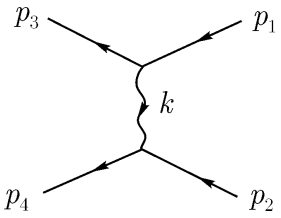


Figure 6.4

$$e^2(\overline{u}_4\gamma^\mu u_1)D_{\mu\nu}(k')(\overline{u}_3\gamma^\nu u_2)=$$

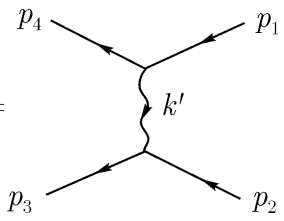


Figure 6.5

$\delta(p_1 + p_2 - p_3 - p_4)$, corresponding to the conservation of the 4-momentum. Transforming from S_{fi} to M_{fi} according to (5.2), (5.3), (5.11), we obtain scattering amplitude M_{fi} as

$$M_{fi} = e^2[(\bar{u}_4\gamma^\mu u_2)D_{\mu\nu}(p_4 - p_2)(\bar{u}_3\gamma^\nu u_1) - (\bar{u}_4\gamma^\mu u_1)D_{\mu\nu}(p_4 - p_1)(\bar{u}_3\gamma^\nu u_2)]\,, \tag{6.79}$$

where

$$D_{\mu\nu}(k) = \int d^4\xi e^{ik\xi} D_{\mu\nu}(\xi) \tag{6.80}$$

is the photon propagator in momentum representation.

Each of the contributions to the scattering amplitude in (6.79) can be represented by an appropriate Feynman diagram. For example, the first term corresponds to the diagram in Figure 6.4, where $k = p_1 - p_3 = p_4 - p_2$. Similarly, the second term is represented by the diagram int Figure 6.5, where $k' = p_1 - p_4 = p_3 - p_2$. The rules of diagram construction are similar to those discussed in Chapter 4:

1. The full lines "entering" the diagram (incoming lines) and directed towards the interaction vertex correspond to the initial electrons and are associated with bispinors u. The full lines "leaving" the diagram (outgoing lines) and directed outside the vertices correspond to the final electrons and are associated with bispinors $\bar{u}$. These factors are written from left to right in the order corresponding to the movement along the full lines against the direction of arrows.
2. To each vertex we associate the factor $(-ie\gamma^\mu)$. The vertices are connected by the photon line, to which we associate the factor $-iD_{\mu\nu}$. The 4-momenta of all particles (lines) in the vertices are conserved. the direction of the photon line is irrelevant;

it only changes the sign of the photon momentum k, but the photon propagator $D_{\mu\nu}(k)$ is an even function of k.

These two diagrams differ from each other because of the *exchange* of two electrons with momenta p_3 and p_4, which corresponds to the exchange of identical particles in the final state, leading to a sign change of the scattering amplitude (Pauli principle!).

Consider now electron–positron scattering. We now denote the initial momenta as p_- and p_+, while the final momenta are denoted as p'_- and p'_+. The operators of the creation and annihilation of positrons enter the field operators (6.73) together with the corresponding creation and annihilation operators of electrons. In the previous case of electron–electron scattering, the annihilation of initial particles was done by the operator ψ, while the creation of the final particle was achieved by the operator $\bar{\psi}$. Now the roles of these operators change: the conjugate function $\bar{\psi}(-p_+)$ describes the initial positron, while the final positron is described by $\psi(-p_+)$. With this difference in mind we can easily write the scattering matrix as

$$M_{fi} = -e^2(\bar{u}(p'_-)\gamma^\mu u(p_-))D_{\mu\nu}(p_- - p'_-)(\bar{u}(-p_+)\gamma^\nu u(-p'_+)) + e^2(\bar{u}(-p_+)\gamma^\mu u(p_-))D_{\mu\nu}(p_- + p_+)(\bar{u}(p'_-)\gamma^\nu u(-p'_+)), \qquad (6.81)$$

which is represented by the diagrams in Figure 6.6. The rules for constructing these diagrams remain the same as before. the incoming full lines are associated with the bispinor u, and the outgoing lines with $\bar{u}$. However, now the incoming lines correspond to the *final* positrons, while the outgoing lines correspond to the *initial* positrons, and their momenta are taken with the opposite sign. This is in agreement with the Feynman interpretation of a positron (discussed in Chapter 4) being an electron propagating backwards in time. In the first diagram of Figure 6.6, in one the vertex cross the lines of the initial and final electrons, while in the other they cross the positron lines, so that this diagram describes electron scattering by positron. In the second diagram in each of the vertices the electron and positron lines meet. In the upper vertex the pair is annihilated and a virtual photon is emitted, while in the lower vertex the pair is created by this photon. This difference is also reflected in the properties of the virtual photons. In the first diagram (scattering channel) the 4-momentum of the virtual photon is equal to the difference of the 4-momenta of two electrons (or positrons), so that $k^2 < 0$

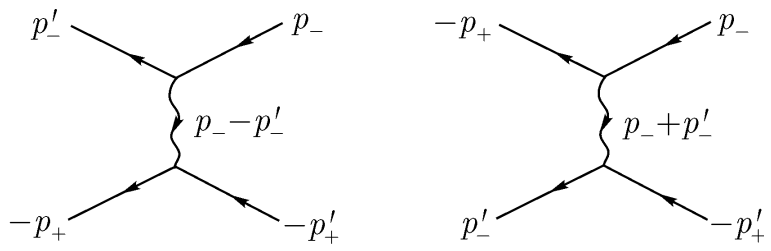


Figure 6.6

(cf. the footnote at the beginning of this section). In the second diagram (annihilation channel), $k' = p_- + p_+$, so that $k'^2 > 0$. Note that for a virtual photon we always have $k^2 \neq 0$, in contrast to a real photon, for which we always have $k^2 = 0$.

6.5 Feynman diagrams for photon scattering

Consider now another effect which appears in the second order of perturbation theory – the photon scattering by electrons (the Compton effect). In the initial state we have a photon and electron with 4-momenta k_1 and p_1 respectively, while in the final state they have momenta k_2 and p_2 (for brevity, we drop the polarization indices). During the calculation of the matrix element $S^{(2)}$ between the initial and final states, the following photon matrix element appears:

$$\langle 2|TA_\mu(x)A_\nu(x)|1\rangle = \langle 0|c_2 TA_\mu(x)A_\nu(x')c_1^+|0\rangle\,, \tag{6.82}$$

where (cf. (3.41))

$$A_\mu = \sum_{\mathbf{k}}(c_{\mathbf{k}}A_{k\mu} + c_{\mathbf{k}}^+ A_{k\mu}^*)\,. \tag{6.83}$$

In (6.82) we are performing all pairings (contractions) of "external" and "internal" photon operators and obtain

$$c_2^{\cdot}A_\mu^{\cdot}A'^{\cdot\cdot}_\nu c_1^{+\cdot\cdot} + c_2^{\cdot}A_\mu^{\cdot\cdot}A'^{\cdot}_\nu c_1^{+\cdot\cdot} = A_{2\mu}^* A'_{1\nu} + A_{1\mu}A'^*_{2\nu}\,. \tag{6.84}$$

Here we have taken into account the commutativity of c_1 and c_2^+, allowing the symbol of T-ordering to be dropped.

Similarly, we can analyze the electronic part of the matrix element:

$$\langle 2|Tj^\mu(x)j^\nu(x')|1\rangle = \langle 0|a_2 T(\bar\psi\gamma^\mu\psi)(\bar\psi'\gamma^\nu\psi')a_1^+|0\rangle\,. \tag{6.85}$$

Here again we are dealing with four ψ-operators. Only two of them annihilate electron 1 and create electron 2, i. e., paired with operators a_1^+ and a_2. These may be operators $\bar\psi'$, ψ or ψ', $\bar\psi$, but not ψ, $\bar\psi$ or ψ', $\bar\psi'$, as the creation or annihilation at the same point x or x' of a pair of real electrons (together with one real photon) obviously produces zero. Making all the contractions, we obtain in the matrix element (6.85) two terms, which we first write for the case of $t > t'$:

$$a_2^{\cdot}(\bar\psi^{\cdot}\gamma^\mu\psi)(\bar\psi'\gamma^\nu\psi'^{\cdot\cdot})a_1^{+\cdot\cdot} + a_2^{\cdot}(\bar\psi\gamma^\mu\psi^{\cdot\cdot})(\bar\psi'^{\cdot}\gamma^\nu\psi')a_1^{+\cdot\cdot}\,. \tag{6.86}$$

Contractions in the first term give

$$a_2\bar\psi \to a_2 a_2^+\bar\psi_2\,, \quad \psi' a_1^+ \to a_1 a_1^+\psi'_1\,. \tag{6.87}$$

Products $a_2 a_2^+$ and $a_1 a_1^+$ are diagonal and can be replaced with their vacuum averages, which, according to (6.74), reduce to unity. For the similar transformation of the

second term in (6.87) we first have to move operator a_2^+ to the left and a_1 to the right, which can be done using the commutation rules, which gives

$$\{a_p, \psi\}_+ = \{a_p^+, \bar{\psi}\}_+ = 0\,,$$
$$\{a_p, \bar{\psi}\}_+ = \bar{\psi}_p\,, \quad \{a_p^+, \psi\}_+ = \psi_p\,, \tag{6.88}$$

where in the right-hand side of the last two expressions spinors appeared, corresponding to plane waves with 4-momentum p (cf. (6.73)). As a results (6.86) is transformed to the form

$$\langle 0|(\bar{\psi}_2\gamma^\mu\psi)(\bar{\psi}'\gamma^\nu\psi_1') - (\bar{\psi}\gamma^\mu\psi_1)(\bar{\psi}_2'\gamma^\nu\psi')|0\rangle \quad \text{for } t > t'\,, \tag{6.89}$$

where ψ without index are operators, while ψ_1, ψ_2 are again just spinors (plane waves) with the appropriate momenta. Similarly, for $t < t'$ we obtain an expression which differs by permutation of the primes and the indices μ and ν:

$$\langle 0| - (\bar{\psi}'\gamma^\nu\psi_1')(\bar{\psi}_2\gamma^\mu\psi) + (\bar{\psi}_2'\gamma^\nu\psi')(\bar{\psi}\gamma^\mu\psi_1)|0\rangle \quad \text{for } t < t'\,. \tag{6.90}$$

Both expressions (6.89) and (6.90) can be written in a unified way, introducing the following definition of chronological (T-ordered) product of Fermion operators:

$$T\psi(x)\bar{\psi}(x') = \begin{cases} \psi(x)\bar{\psi}(x')\,, & t' < t\,, \\ -\bar{\psi}(x')\psi(x)\,, & t' > t\,. \end{cases} \tag{6.91}$$

Then the first and second terms in (6.89) and (6.90) are written as

$$\bar{\psi}_2\gamma^\mu\langle 0|T\psi\bar{\psi}'|0\rangle\gamma^\nu\psi_1' + \bar{\psi}_2'\gamma^\nu\langle 0|T\psi'\bar{\psi}|0\rangle\gamma^\mu\psi_1\,. \tag{6.92}$$

Note that in accordance with definition of (6.91) the products of the operators for $t < t'$ and $t > t'$ are taken with different signs. This is the main difference of the definition of the T-product for fermion operators from those given previously, which is related to the anticommutation of these operators, in contrast to commuting bilinear forms, entering the interaction Hamiltonian.

Let us define the electron propagator (Green's function) as a second rank bispinor of the form

$$G(x - x') = -i\,\langle 0|T\psi(x)\bar{\psi}(x')|0\rangle\,. \tag{6.93}$$

Then the matrix element of interest to us is written as

$$\langle 2|Tj^\mu(x)j^\nu(x')|1\rangle = i\,\bar{\psi}_2\gamma^\mu G(x - x')\gamma^\nu\psi_1' + +i\,\bar{\psi}_2'\gamma^\nu G(x' - x)\gamma^\mu\psi_1\,. \tag{6.94}$$

After multiplication by the photon matrix element (6.82), (6.84), and integration over $d^4x d^4x'$, both terms in (6.94) give the same result, so that

$$S_{fi} = -i\,e^2\int d^4x\int d^4x'\bar{\psi}_2(x)\gamma^\mu G(x - x')\gamma^\nu\psi_1(x')$$
$$\times\left[A_{2\mu}^*(x)A_{1\nu}(x') + A_{2\nu}^*(x')A_{1\mu}(x)\right]. \tag{6.95}$$

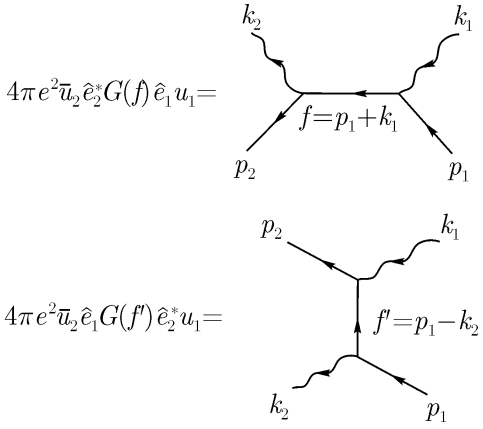


Figure 6.7

Substituting plane waves for the electron and photon wave functions and separating, as in previous examples, the δ-function, corresponding to the 4-momentum conservation law, we obtain the scattering amplitude as

$$M_{fi} = -4\pi e^2 \bar{u}_2 [(\gamma e_2^*) G(p_1 + k_1)(\gamma e_1) + (\gamma e_1) G(p_1 - k_2)(\gamma e_2^*)] u_1 \,, \qquad (6.96)$$

where e_1 and e_2 are 4-vectors of photon polarization and $G(p)$ is the electron propagator in momentum representation. Two terms of this expression are represented by the Feynman diagrams shown in Figure 6.7.

To the incoming line (initial photon) we associated the factor $\sqrt{4\pi}e$, to the outgoing line (final photon) the factor $\sqrt{4\pi}e^*$. The full internal line corresponds to the virtual electron with 4-momentum determined by the 4-momentum conservation law in the vertices. This line is associated with the electron propagator $iG(f)$. In contrast to the 4-momentum of a real particle, the square of the virtual electron 4-momentum does not belong to its mass surface, i. e., it is not equal to m^2. Writing the invariant f^2 in electron rest frame, it is easy to show that

$$f^2 = (p_1 + k_1)^2 > m^2 \,, \qquad f'^2 = (p_1 - k_2)^2 < m^2 \,. \qquad (6.97)$$

6.6 Electron propagator

Let us calculate explicitly the propagators (Green's functions) of free particles. By definition (6.93) the electron propagator is given by

$$G(x - x') = -i \langle 0 | T \psi(x) \bar{\psi}(x') | 0 \rangle \,.$$

Let us act upon it from the right side by the operator $\gamma^\mu p_\mu - m$, where $p_\mu = i\partial_\mu$. As the free field $\psi(x)$ satisfies the Dirac's equation $(\gamma^\mu p_\mu - m)\psi(x) = 0$, we shall get

zero in all points x except those where $t = t'$. Note that $G(x - x')$ tends to different limits as $t \to t' + 0$ and $t \to t' - 0$, and according to definition (6.93) these limits are given by

$$-i\langle 0|\psi(\mathbf{r}t)\bar{\psi}(\mathbf{r}'t)|0\rangle \quad \text{and} \quad +i\langle 0|\bar{\psi}(\mathbf{r}'t)\psi(\mathbf{r}t)|0\rangle\,, \tag{6.98}$$

so that for $t = t'$ the Green's function demonstrates finite discontinuity. This leads to the appearance of an additional term with the δ-function in the derivative $\partial G/\partial t$:

$$\frac{\partial G}{\partial t} = -i\langle 0|T\frac{\partial \psi}{\partial t}\psi(x')|0\rangle + \delta(t - t')[G|_{t\to t'+0} - G|_{t\to t'-0}]\,. \tag{6.99}$$

Note that in $\gamma^\mu p_\mu - m$ the derivative over t enters as $i\gamma^0\partial/\partial t$, so that

$$(\gamma^\mu p_\mu - m)G(x - x') = \delta(t - t')\gamma^0\langle 0|\left\{\psi(\mathbf{r}t), \bar{\psi}(\mathbf{r}'t)\right\}_+|0\rangle\,. \tag{6.100}$$

Now calculate the anticommutator. Multiplying the field operators, which we take in the form of (6.73), and using commutation relations for $a_\mathbf{p}$ and $b_\mathbf{p}$, we obtain

$$\left\{\psi(\mathbf{r}, t), \bar{\psi}(\mathbf{r}'t)\right\}_+ = \sum_\mathbf{p}[\psi_p(\mathbf{r})\bar{\psi}_p(\mathbf{r}') + \psi_{-p}(\mathbf{r})\bar{\psi}_{-p}(\mathbf{r}')]\,, \tag{6.101}$$

where $\psi_{\pm p}(\mathbf{r})$ are plane waves (bispinors) without a time dependent factor. These functions form the full set, so that

$$\sum_\mathbf{p}[\psi_p(\mathbf{r})\psi_p^*(\mathbf{r}') + \psi_{-p}(\mathbf{r})\psi_{-p}^*(\mathbf{r}')] = \delta(\mathbf{r} - \mathbf{r}')\delta_{ik} \tag{6.102}$$

where δ_{ik} is the Kronecker symbol over the spinor indices. The sum in the right-hand side of (6.101) differs from (6.102) by the replacement of ψ^* by $\psi^*\gamma^0$, so that

$$\left\{\psi(\mathbf{r}t), \bar{\psi}(\mathbf{r}'t)\right\}_+ = \gamma^0\delta(\mathbf{r} - \mathbf{r}')\,. \tag{6.103}$$

Substituting (6.103) into (6.100) we finally get

$$(\gamma^\mu p_\mu - m)G(x - x') = \delta(x - x')\,. \tag{6.104}$$

Thus, the electron propagator satisfies the Dirac's equation with a δ-function in the right-hand side, so that it is really the Green's function for this equation[7].

Consider now the Fourier transform of Green's function

$$G(p) = \int d^4\xi e^{-ip\xi}G(\xi)\,. \tag{6.105}$$

Calculating the Fourier transforms of both sides of equation (6.104), we get

$$(\gamma^\mu p_\mu - m)G(p) = 1\,. \tag{6.106}$$

[7] It is easy to see that $iG(x_1 - x_2)$ just coincides with Feynman's function $K_+(2, 1)$ introduced in Chapter 4.

Solving this equation, we obtain the result, which we already know from Chapter 4,

$$G(p) = \frac{\gamma^\mu p_\mu + m}{p^2 - m^2}\,. \tag{6.107}$$

The components of the 4-vector p in $G(p)$ are independent variables and are not restricted by any relations like $p^2 \equiv p_0^2 - \mathbf{p}^2 = m^2$. If we write the denominator of (6.107) as $p_0^2 - (\mathbf{p}^2 + m^2)$, we can see, that $G(p)$ as the function of p_0 for a given value of $\mathbf{p}$ has two poles at $p_0 = \pm\varepsilon$, where $\varepsilon = \sqrt{\mathbf{p}^2 + m^2}$. Then again, during the integration over dp_0 in

$$G(\xi) = \int \frac{d^4p}{(2\pi)^4} e^{-ip\xi} G(p) = \frac{1}{(2\pi)^4} \int d^3\mathbf{p} e^{i\mathbf{pr}} \int dp_0 e^{-ip_0\tau} G(p)\,, \quad (\tau = t - t') \tag{6.108}$$

we meet the problem of encircling these poles which we first discussed in Chapter 4. Again we shall use Feynman's approach. Let us return definition (6.93). Substitute into it ψ-operators in the form given by (6.73). Note that nonzero vacuum averages appear only from the following products of creation and annihilation operators:

$$\langle 0|a_p a_p^+|0\rangle = 1\,, \quad \langle 0|b_p b_p^+|0\rangle = 1\,. \tag{6.109}$$

Then,

$$G(x - x') = -i \sum_{\mathbf{p}} \psi_p(\mathbf{r}t)\bar{\psi}_p(\mathbf{r}'t') = -i \sum_{\mathbf{p}} e^{-i\varepsilon(t-t')} \psi_p(\mathbf{r})\bar{\psi}_p(\mathbf{r}') \tag{6.110}$$

for $t - t' > 0$. Accordingly,

$$G(x - x') = i \sum_{\mathbf{p}} \bar{\psi}_{-p}(\mathbf{r}'t')\psi_{-p}(\mathbf{r}t) = i \sum_{\mathbf{p}} e^{i\varepsilon(t-t')} \psi_{-p}(\mathbf{r})\bar{\psi}_{-p}(\mathbf{r}') \tag{6.111}$$

for $t - t' < 0$. We see that, as in Chapter 4, for $t - t' > 0$ only electrons contribute to G, while for $t - t' < 0$ – only positrons. Comparing (6.110) and (6.111) with (6.108) we see that integral

$$\int dp_0 e^{-ip_0\tau} G(p) \tag{6.112}$$

in equation (6.108) must produce a factor $e^{-i\varepsilon\tau}$ for $\tau > 0$ and $e^{i\varepsilon\tau}$ for $\tau < 0$. This can be achieved if during the calculation of (6.112) we encircle poles $p_0 = \varepsilon$ and $p_0 = -\varepsilon$ in the upper and lower half-planes of complex variable p_0 correspondingly, as shown in Figure 6.8. In fact, for $\tau > 0$ we have (to guarantee convergence!) closed the integration contour to a semicircle at infinity in the lower half-plane of p_0, then the

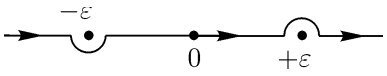


Figure 6.8

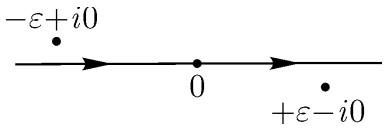


Figure 6.9

value of integral (6.112) will be determined by the residue at the pole $p_0 = +\varepsilon$. For $\tau < 0$ we close the contour in the upper half-plane, and integral is determined by the residue at the pole $p_0 = -\varepsilon$. Thus we achieve the desired result. The Feynman rule to deal with the poles, as we have seen in Chapter 4, can be formulated in another form: integration over p_0 is performed along the real axis, but we an add infinitesimally small negative imaginary part to particle mass m:

$$m \to m - i0\,. \tag{6.113}$$

Then

$$\varepsilon \to \sqrt{\mathbf{p}^2 + (m - i0)^2} = \sqrt{\mathbf{p}^2 + m^2} - i0 = \varepsilon - i0\,. \tag{6.114}$$

Correspondingly, the poles $p_0 = \pm\varepsilon$ are moved up and down from the real axis, as shown in Figure 6.9, so that integration becomes equivalent to integration along the contour shown in Figure 6.8[8].

This rule for dealing with poles is equivalent to the well-known relation

$$\frac{1}{x \pm i0} = P\frac{1}{x} \mp i\pi\delta(x)\,, \tag{6.115}$$

which is understood in the sense that integration with some regular function $f(x)$ is done as follows:

$$\int_{-\infty}^{\infty} dx \frac{f(x)}{x \pm i0} = P\int_{-\infty}^{\infty} dx \frac{f(x)}{x} \mp i\pi f(0)\,, \tag{6.116}$$

where P denotes the principal value of the integral.

Using Feynman's rule we write the electron propagator in momentum representation as

$$G(p) = \frac{\gamma^\mu p_\mu + m}{p^2 - m^2 + i0}\,. \tag{6.117}$$

This Green's function is the product of the bispinor $\gamma^\mu p_\mu + m$ and the scalar

$$G^{(0)}(p) = \frac{1}{p^2 - m^2 + i0}. \tag{6.118}$$

[8] It is useful to note that this rule of pole encirclement corresponds to $G(x-x')$ acquiring an infinitesimal damping over $|\tau|$. If we write the value of p_0 in displaced poles as $-(\varepsilon - i\delta)$ and $+(\varepsilon - i\delta)$, where $\delta \to +0$, the time-dependent exponent in integral (6.112) will be equal to $\exp(-i\varepsilon|\tau| - \delta|\tau|)$.

In coordinate representation $G^{(0)}(\xi)$ satisfies the equation

$$(\Box - m^2)G^{(0)}(x - x') = \delta(x - x'), \tag{6.119}$$

being the Green's function of the Klein-Gordon equation. It is obvious, that it determines the propagator of scalar particles and can be defined via scalar filed as

$$G^{(0)}(x - x') = -i \langle 0|T\varphi(x)\varphi^+(x')|0\rangle , \tag{6.120}$$

where

$$T\varphi(x)\varphi^+(x') = \begin{cases} \varphi(x)\varphi^+(x') , & t' < t , \\ \varphi^+(x')\varphi(x) , & t' > t \end{cases} \tag{6.121}$$

is the definition of the T-product for the Bose field.

6.7 The photon propagator

While analyzing the free electromagnetic field, we used the expansion of the vector-potential over transversal plane waves. This description does not apply in the case of an arbitrary electromagnetic field. This is obvious, as for example in case of electron scattering we have to take into account Coulomb interaction, which is described by the scalar potential and is not reduced to the exchange of transversal virtual photons. Thus, it seems that we still do not have the full definition of operators A_μ and cannot calculate the photon propagator directly using the expression

$$D_{\mu\nu}(x - x') = i \langle 0|TA_\mu(x)A_\nu(x')|0\rangle . \tag{6.122}$$

Besides that, the gauge invariance makes the field operators somehow unphysical. However, below we shall present some general analysis which solves all of these problems [6].

The most general form of the symmetric 4-tensor of the second rank, depending only on the 4-vector $\xi = x - x'$, is given by

$$D_{\mu\nu}(\xi) = g_{\mu\nu}D(\xi^2) - \partial_\mu\partial_\nu D^l(\xi^2) , \tag{6.123}$$

where D and D^l are scalar functions of invariant ξ^2. Then, in momentum representation we have

$$D_{\mu\nu}(k) = g_{\mu\nu}D(k^2) + k_\mu k_\nu D^l(k^2) , \tag{6.124}$$

where $D(k^2)$ and $D^l(k^2)$ are Fourier components of $D(\xi^2)$ and $D^l(\xi^2)$.

The photon Green's function always enters scattering amplitudes multiplied by the matrix elements of the transition currents of a pair of electrons, i. e., in combinations like $j_{21}^\mu D_{\mu\nu} j_{43}^\nu$, which is seen, e. g., from equation (6.79). Current conservation gives $\partial_\mu j^\mu = 0$, so that the matrix elements of the current satisfy the condition of four-dimensional transversality:

$$k_\mu j_{21}^\mu = 0 , \tag{6.125}$$

where $k = p_2 - p_1$. Thus, the physical results do not change after the replacement

$$D_{\mu\nu} \to D_{\mu\nu} + \chi_\mu k_\nu + \chi_\nu k_\mu \,, \tag{6.126}$$

where χ_μ are arbitrary functions of k. This arbitrariness corresponds, in fact, to different choices of the gauge for field operators. Thus, the choice of the function $D^l(k^2)$ in (6.124) is in fact arbitrary[9] and can be done to make calculations more convenient. Thus, the full definition of the Green's function reduces to the choice of a single gauge invariant function $D(k^2)$. Taking the fixed value of k^2 and choosing the z-axis along the direction of $\mathbf{k}$, we can see that transformations (6.126) will not affect $D_{xx} = D_{yy} = -D(k^2)$. Thus, it is sufficient to calculate only one component D_{xx}, using an arbitrary choice of the gauge for potentials.

It is convenient to use, as before, the Coulomb gauge $\operatorname{div} \mathbf{A} = 0$, when the operator $\mathbf{A}$ is given by

$$\mathbf{A} = \sum_{\mathbf{k}\alpha} \sqrt{\frac{2\pi}{\omega}} \Big(c_{\mathbf{k}\alpha} \mathbf{e}^{(\alpha)} e^{-ikx} + c^+_{\mathbf{k}\alpha} \mathbf{e}^{(\alpha)*} e^{ikx} \Big) , \tag{6.127}$$

where $\omega = |\mathbf{k}|$ and $\alpha = 1, 2$ enumerates polarizations. The only nonzero vacuum average of the product of operators c, c^+ is

$$\langle 0 | c_{\mathbf{k}\alpha} c^+_{\mathbf{k}\alpha} | 0 \rangle = 1 \,. \tag{6.128}$$

Then, using the definition (6.122), we obtain

$$D_{ik}(\xi) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{2\pi i}{\omega} \Big(\sum_\alpha e^{(\alpha)}_i e^{(\alpha)*}_k \Big) e^{-i\omega|\tau| + i\mathbf{k}\boldsymbol{\xi}} \,, \tag{6.129}$$

where i, k are three-dimensional vector indices. The presence in the exponent of the modulus of $\tau = t - t'$ reflects the T-ordering of the field operators in (6.122). From (6.129) it is clear that the integrand without the factor of $e^{i\mathbf{k}\boldsymbol{\xi}}$ represents the Fourier component of $D_{ik}(\mathbf{r}t)$. For $D_{xx} = -D$ it is equal to

$$\frac{2\pi i}{\omega} e^{-i\omega|\tau|} \sum_\alpha |e^{(\alpha)}_x|^2 = \frac{2\pi i}{\omega} e^{-i\omega|\tau|} \,. \tag{6.130}$$

To find $D_{xx}(k^2)$ we have to Fourier expand this function in time. This gives

$$\frac{2\pi i}{\omega} e^{-i\omega|\tau|} = -\int_{-\infty}^{\infty} \frac{dk_0}{2\pi} \frac{4\pi}{k_0^2 - \mathbf{k}^2 + i0} e^{-ik_0\tau} \,. \tag{6.131}$$

As shown above, integration here assumes encirclement of poles $k_0 = \pm|\mathbf{k}| = \pm\omega$ from below and above, correspondingly, so that for $\tau > 0$ the integral is determined

[9] Consider $\delta D^l(k^2)$: an arbitrary change of $D^l(k^2)$. Then we get $\delta D_{\mu\nu} = k_\mu k_\nu \delta D^l \equiv k_\mu \chi_\nu$, where $\chi_\nu = k_\nu \delta D^l(k^2)$.

by the residue at the pole $k_0 = +\omega$, for $\tau < 0$ by the residue at the pole $k_0 = -\omega$. Thus, we finally find

$$D(k^2) = \frac{4\pi}{k^2 + i0}\,. \tag{6.132}$$

Now it is obvious that the corresponding function in coordinate representation satisfies the equation

$$\Box D(x - x') = -4\pi\delta(x - x')\,, \tag{6.133}$$

so that it is the Green's function of the wave equation.

In most cases it is convenient to choose $D^l = 0$, so that the photon propagator has the form

$$D_{\mu\nu} = g_{\mu\nu} D(k^2) = \frac{4\pi}{k^2 + i0} g_{\mu\nu}\,, \tag{6.134}$$

which coincides with the result obtained in Chapter 4 and corresponds to the so-called Feynman gauge.

Sometimes it is convenient to choose $D^l = -D(k^2)/k^2$, so that

$$D_{\mu\nu} = \frac{4\pi}{k^2}\left(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right), \tag{6.135}$$

corresponding to the so-called Landau gauge. Then $D_{\mu\nu}k^\nu = 0$, and this choice is similar to the Lorentz gauge, where $k^\mu A_\mu = 0$.

The choice of the gauge $\operatorname{div} \mathbf{A} = 0$ leads to a similar gauge for the propagator:

$$D_{ii}k^l = 0\,, \quad D_{0l}k^l = 0\,, \tag{6.136}$$

which, together with $D_{xx} = -D = -4\pi/k^2$, gives

$$D_{il} = -\frac{4\pi}{\omega^2 - \mathbf{k}^2}\left(\delta_{il} - \frac{k_i k_l}{\mathbf{k}^2}\right), \quad D_{00} = -\frac{4\pi}{\mathbf{k}^2}\,, \quad D_{0i} = 0\,, \tag{6.137}$$

so that D_{00} simply equals to the Fourier transform of the Coulomb potential.

For massive particles with spin $s = 1$ there is no gauge invariance, and the choice of propagator is unique. Substituting the appropriate operators ψ_μ into the definition

$$G_{\mu\nu}(x - x') = -i\,\langle 0|T\psi_\mu(x)\psi_\nu^+(x')|0\rangle\,, \tag{6.138}$$

we obtain an expression which differs from (6.129) only in the form of the sum over polarizations, which takes into account three independent polarizations of the massive vector field. Dropping the technical details [6], we only mention that in momentum representation the propagator of the vector field is equal to

$$G_{\mu\nu} = -\frac{1}{p^2 - m^2 + i0}\left(g_{\mu\nu} - \frac{p_\mu p_\nu}{m^2}\right). \tag{6.139}$$

6.8 The Wick theorem and general diagram rules

From simple examples of calculations of matrix elements of scattering matrix, considered above, we already can see the advantages of diagrammatic approach. Let us consider now the general case. Matrix element of S for a transition between arbitrary initial and final states coincides with vacuum average of operator, which is obtained by multiplication of S from the right by creation operators of all initial particles and from the left by annihilation operators of all final particles. Then, in n-th order of perturbation theory this matrix element is written as:

$$\langle f|S^{(n)}|i\rangle = \frac{1}{n!}(-ie)^n \int d^4x_1 \cdots d^4x_n \Big\langle 0|\ldots b_{2f} b_{1f} \cdots a_{1f} \cdots c_{1f} T \times (\bar{\psi}_1 \gamma A_1 \psi_1) \cdots (\bar{\psi}_n \gamma A_n \psi_n) c^+_{1i} \cdots a^+_{1i} \cdots b^+_{1i} \cdots |0\Big\rangle . \qquad (6.140)$$

The indices $1i, 2i, \ldots$ enumerate the initial particles (electrons, positrons, and photons separately), while $1f, 2f, \ldots$ enumerate the final ones. The indices $1, 2, \ldots$ of operators ψ and A correspond to $\psi_1 = \psi(x_1)$ etc. Operators ψ and A are represented by linear combinations of the creation and annihilation operators of the appropriate particles in different states. Thus, in matrix element vacuum averages of products of creation and annihilation operators as well as their linear combinations appear. All these operators are taken in interaction representation, so that they satisfy the equations of motion and the commutation relations for *free* particles. Calculation of such averages is done using the *Wick theorem*, which is proved below.

The Wick theorem

Let us define the *normal* product of operators

$$N(ABCD\cdots XYZ), \qquad (6.141)$$

so that all the creation operators are to the left of the annihilation operators, with their sign corresponding to the parity of the permutation of the Fermion operators, which transforms the product to the necessary form. Obviously, the vacuum average of the normal product of operators equals zero, except in the case where under the sign of the normal product we have simply some c-number. Let us call the "pairing" ("contraction") of two operators the following difference:

$$A^{\cdot}B^{\cdot} = T(AB) - N(AB). \qquad (6.142)$$

It is easy to see that this expression is a c-number, as its right-hand side is either zero or coincides (up to a sign) with the commutator (anticommutator) of operators A and B. The main statement of the Wick theorem is that the T-product of an arbitrary number of

operators can be expressed through all possible N-products with all possible pairings (contractions):

$$T(ABCD\cdots XYZ) = N(ABCD\cdots XYZ) + N(A^{\cdot}B^{\cdot}CD\cdots XYZ) + N(A^{\cdot}BC^{\cdot}D\cdots XYZ) + \cdots + N(A^{\cdot}B^{\cdot\cdot}C^{\cdot\cdot\cdot}D\cdots X^{\cdot}Y^{\cdot\cdot}Z^{\cdot\cdot\cdot}), \quad (6.143)$$

i. e., the chronological product of the operators is equal to the normal product plus the sum of the normal products with one pairing (the pair can be chosen in all possible ways), plus the sum of the normal products with two pairings, etc. Pairing within the normal product is the c-number, which is (up to a sign ± 1) determined by equation (6.142). The minus sign is chosen when the permutation needed to bring paired operators out of the normal product is odd with respect to the Fermion operators.

To prove the theorem, we first note that the simultaneous permutation of the operators in both sides of (6.143) does not change this equality. Then, with no loss of generality we may assume that the time ordering of operators corresponds to that in (6.143). To obtain the N-product from the T-product we have to take all the creation operators and interchange their positions with all the annihilation operators on the left one by one, using the definition (6.142). In this way we get the sum of the N-products of the type written in (6.143). However, this will contain only the contractions of those operators, for which the order in the T-product is different from that in the N-product. But pairings of operators for which both orders are equivalent equal to zero, and we can say that the right-hand side of (6.143) contain normal products with all possible pairings (contractions). This proves the Wick theorem.

The Wick theorem helps to calculate the averages of the products of the operators over the vacuum state $|0\rangle$. The average of the normal products is obviously zero, so that a nonzero contribution comes only from those terms in the right-hand side of (6.143), where all the operators are paired:

$$\langle 0|T(ABCD\cdots XYZ)|0\rangle = \langle 0|T(AB)|0\rangle\langle 0|T(CD)|0\rangle\cdots\langle 0|T(YZ)|0\rangle \pm \langle 0|T(AC)|0\rangle\langle 0|T(BD)|0\rangle\cdots\langle 0|T(YZ)|0\rangle \pm \cdots, \quad (6.144)$$

where we have taken into account that

$$\langle 0|A^{\cdot}B^{\cdot}|0\rangle = \langle 0|T(AB)|0\rangle. \quad (6.145)$$

Thus, the average is represented by the sum of all the possible products of the averages (over the ground state) of the T-products of the pairs of operators. The sign of each term corresponds to the parity of permutation of the Fermion operators. From (6.144) it follows, in particular, that among the operators $A, B, C, D, \ldots$ there should necessarily be an even number of operators of each field. Recalling the definition of the Green's function, we conclude that the vacuum average of the T-product of an arbitrary (even) number of field operators is expressed via the sum of the products of the free Green's functions.

Applying Wick's theorem to the matrix element (6.140), we can represent it as the sum of terms, each being the product of some pairwise averages. Among these we shall meet pairings of operators ψ, $\bar{\psi}$, and A with the "external" operators of the creation of the initial particles or the annihilation of the final ones. These pairings can be expressed through the wave functions of the initial and final particles as

$$\begin{aligned} &\langle 0|Ac_p^+|0\rangle = A_p\,, \qquad &&\langle 0|c_pA|0\rangle = A_p^*\,, \\ &\langle 0|\psi a_p^+|0\rangle = \psi_p\,, \qquad &&\langle 0|a_p\bar{\psi}|0\rangle = \psi_p^*\,, \\ &\langle 0|b_p\psi|0\rangle = \psi_{-p}\,, \qquad &&\langle 0|\bar{\psi}b_p^+|0\rangle = \bar{\psi}_{-p}\,, \end{aligned} \tag{6.146}$$

where A_p, ψ_p are photon and electron wave functions with momentum p. The polarization indices are dropped here for brevity. There will also be pairings of "internal" operators, standing under the sign of the T-product. These pairings are replaced by the appropriate propagators.

Each of the terms in the sum for the matrix element of the S-matrix appearing as a result of the application of Wick theorem can be represented by the appropriate Feynman diagram. In the diagram of the n-th order there are n vertices, each associated with the corresponding integration variable $x_1, x_2, \ldots$. Each vertex is connected with three lines – two full ones (electrons) and one wavy one (photon), corresponding to the electron (ψ and $\bar{\psi}$) and photon (A) operators, as functions of the same variable x. The operator ψ corresponds to the incoming line and $\bar{\psi}$ – to the outgoing line.

To illustrate this we show several examples of the correspondence between the terms of the matrix element of third order and the diagrams. Dropping the signs of the integral and T-ordering, as well as the factors $-ie\gamma$ and the arguments of the operators, we write these terms in symbolic form, as shown in Figure 6.10, where pairings (contractions) are shown, as often done, by lines connecting the appropriate field operators. Note that for the internal photon pairings the direction of photon lines is of no importance, because of the even nature of the photon propagator as a function of $x - x'$.

Among the terms obtained in this way there are some that are equivalent and which differ only by the permutation of the numbers of vertices, reflecting the correspondence between the vertices and the number of variables $x_1, x_2, \ldots$, i. e., by a simple redefinition of the integration variables. The number of such permutations is $n!$. It cancels the factor of $1/n!$ in (6.140), so that we, in fact, do not need to take into account diagrams with all the permutations of the vertices. For example, two diagrams of the second order, shown in Figure 6.11, are just equivalent. In Figure 6.10 and Figure 6.11 we show only internal pairings, corresponding to internal lines in the diagrams (virtual electrons and photons). The remaining free operators are paired with external operators, which establishes the correspondence between the external "legs" of diagrams and the initial or final particles. For example, pairing $\bar{\psi}$ with operators a_f or b_i^+ gives the line of the final electron or initial positron, while ψ pairing with a_i^+ or b_f pro-

duces the line of the initial electron or final positron. Free operator A pairing with c_i^+ or with c_f may correspond to both the initial or final photon. In this way we obtain several "topologically equivalent" (i. e., consisting of the same number of lines with identical configurations of the graphs) diagrams, differing only by the permutations of the initial and the final incoming and outgoing external legs. Any such permutation is equivalent to some permutation of external operators $a, b, \ldots$. If among the initial or final particles there are identical fermions, the relative sign of the diagrams, differing by the odd permutation of the corresponding free legs, should be opposite.

The nonoverlapping sequence of the full lines in the diagrams forms the electron line with an arrow along it, conserving the continuous direction. It may have two free external legs, or it can form a loop, as shown in Figure 6.12. The conservation of direction along the electronic line is a graphical expression of the charge conservation law: the "incoming" charge to every vertex is equal to the "outgoing' one. Placement of the bispinor indices along the continuous electron line corresponds to writing the matrices from left to right, moving against the arrows. Bispinor indices of different electron lines never intermix. Along the nonclosed line, the sequence of indices ends at the free external legs on the electron (or positron) wave functions. On the closed loop the sequence of indices also closes, so that the loop corresponds to the trace of the product of the matrices along it. It is easy to see that this trace should always be taken with a minus sign, In fact, the loop with k vertices corresponds to the set of k pairings:

$$(\bar{\psi}^{\cdot} A \psi^{\cdot\cdot})(\bar{\psi}^{\cdot\cdot} A \psi^{\cdot\cdot\cdot}) \ldots (\bar{\psi}^{\cdot\cdot\cdot} A \psi^{\cdot}) \tag{6.147}$$

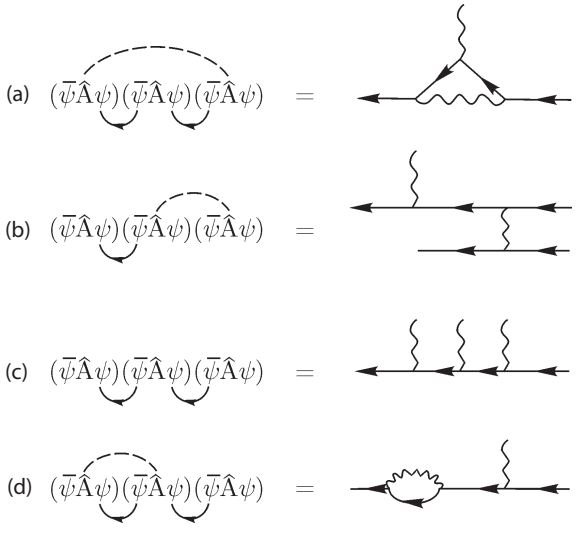


Figure 6.10

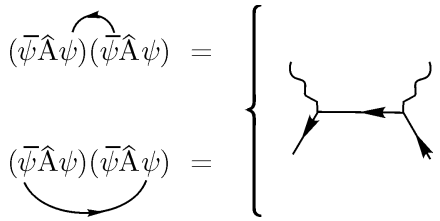


Figure 6.11

or the other similar, differing by permutation of vertices. In the $(k-1)$-th pairing the operators ψ and $\bar{\psi}$ are already in the correct order ($\bar{\psi}$ to the right of ψ) in which they should stand in an electron propagator. Those operators at the edges can be moved to become neighbors with the help of an even number of permutations with other ψ-operators, to get at the end the correct order $\bar{\psi}\psi$. As $\langle 0|T\bar{\psi}'\psi|0\rangle = -\langle 0|T\psi\bar{\psi}'|0\rangle$, the replacement of this pairing by the corresponding propagator is related to the change of the total sign of the whole expression.

Transformation to momentum representation leads to the general 4-momentum conservation law and also to the similar conservation law at every vertex. However, these laws may be *insufficient* for fixing momenta of all internal lines in a given diagram. In these cases we should perform the integration of all momenta of internal lines $d^4p/(2\pi)^4$ which remain indeterminate.

In a similar way we may analyze the case with an external electromagnetic field (cf. Chapter 4), i. e., the field, created by "passive" particles, whose states are not changed during the scattering process (these may be heavy "classical" charges). Let $A^{(e)}(x)$ be the 4-potential of an external field. It enters the Lagrangian together with the photon operator A as a sum $A + A^{(e)}$. Because of the classical nature of $A^{(e)}$ it is actually a c-number field containing no operators, and it cannot pair with other operators. Thus, in Feynman diagrams external fields may correspond only to external lines. Let us introduce Fourier expansion for $A^{(e)}$:

$$A^{(e)}(x) = \int \frac{d^4q}{(2\pi)^4} e^{-iqx} A^{(e)}(q)\,, \qquad A^{(e)}(q) = \int d^4x e^{iqx} A^{(e)}(x)\,. \tag{6.148}$$

In expressions for matrix elements in the momentum representation 4-vector q will be present along the 4-momenta of other external lines, corresponding to real particles. To each line of an external field we associate the factor $A^{(e)}(q)$, with the corresponding line considered as "incoming" in accordance with the sign in the exponent e^{-iqx} in Fourier expansion for $A^{(e)}(q)$ (the "outgoing" line should be associated with $A^{(e)*}(q)$). If the 4-momentum conservation law,

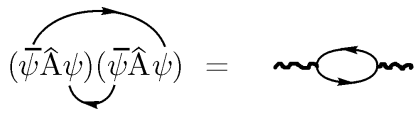


Figure 6.12

with given values of the 4-momenta of all real particles, does not fix the 4-momenta for the lines of external field, we have to integrate over "free" $d^4q/(2\pi)^4$, as well as over all other nonfixed 4-momenta in the diagram.

If external field does not depend on time,

$$A^{(e)}(q) = 2\pi\delta(q^0)A^{(e)}(\mathbf{q})\,, \tag{6.149}$$

where $A^{(e)}(\mathbf{q})$ is the three-dimensional Fourier component,

$$A^{(e)}(\mathbf{q}) = \int d^3\mathbf{r}A^{(e)}(\mathbf{r})e^{-i\mathbf{qr}}\,. \tag{6.150}$$

The xternal line is now associated with the factor $A^{(e)}(\mathbf{q})$ and the 4-momentum $q^\mu = (0, \mathbf{q})$. Energies of electron lines entering (along with the line of the external field) the vertex are the same due to the conservation law. Over the remaining nonfixed three-dimensional momenta $\mathbf{p}$ of the internal lines we perform the integration $d^3\mathbf{p}/(2\pi)^3$.

Let us now present the final summary of the diagrammatic rules for scattering amplitude (more precisely for iM_{fi}) of QED in momentum representation.

1. Contributions of the n-th order of perturbation theory are represented by diagrams with n vertices, each with one incoming and one outgoing electron (full) line and one photon (wavy) line. The scattering amplitude is described by all diagrams with free external legs (external lines), corresponding to the initial and final particles.

2. To each external incoming full line (leg) we associate the amplitude of initial electron $u(p)$ or final positron $u(-p)$. To each outgoing full line we associate the amplitude of final electron $\bar{u}(p)$ or initial positron $\bar{u}(-p)$.

3. To each vertex we associate the 4-vector $-ie\gamma^\mu$.

4. To each external incoming wavy line we associate the amplitude of initial photon $\sqrt{4\pi}e_\mu$, and to the outgoing wavy line the amplitude $\sqrt{4\pi}e^*_\mu$ of the final photon. The vector index μ coincides with index of the γ^μ matrix in the corresponding vertex, so that we have the scalar product.

5. To each internal full line we associate the factor $iG(p)$, and to the internal wavy line we associate the factor of $-iD_{\mu\nu}(p)$. Tensor indices $\mu\nu$ coincide with indices of matrices γ^μ, γ^ν in vertices, connected by a wavy line.

6. Along the continuous sequence of electron lines the arrows have the same direction, while the positions of bispinor indices correspond to writing matrices from left to right against the arrows. A closed loop is associated with the trace of the product of the corresponding bispinor matrices.

7. In each vertex, the 4-momenta of the lines entering or leaving it satisfy the conservation law, i. e., the sum of the momenta of incoming lines equals the sum of the

momenta of the outgoing lines. The momenta of the external lines (legs) are fixed (and obey the total conservation law for the scattering process under discussion), with the positron line associated with the momentum $-p$. Integration $d^4p/(2\pi)^4$ is performed over all momenta of the internal lines, remaining nonfixed, after taking into account the conservation laws in all vertices.

8. An incoming external line (leg), corresponding to the external field, is associated with factor $A^{(e)}(q)$, where the 4-vector q is related to the 4-momenta of other lines in the vertex by the conservation law. If the external field does not depend on time, this external leg corresponds to the factor of $A^{(e)}(\mathbf{q})$, and integration $d^3\mathbf{p}/(2\pi)^3$ is performed over nonfixed three-dimensional momenta of internal lines.

9. To each closed fermion loop we associate an extra factor of (-1). If among the initial or final particles there are several electrons or positrons, the relative sign of the diagrams, differing by odd permutations of the identical particles (i. e., the corresponding external legs), should be opposite.

Finally, let us recall that in the presence of identical fermions the total sign of the scattering amplitude is irrelevant.

Chapter 7

Exact propagators and vertices

7.1 Field operators in the Heisenberg representation and interaction representation

Above we have expressed the terms of a perturbation series via field operators in the interaction representation, with time-dependence determined by the Hamiltonian H_0 of free particles. Exact scattering amplitudes are more conveniently expressed via field operators in the Heisenberg representation, where the time-dependence is determined by the total Hamiltonian of interacting particles $H = H_0 + H_I$. According to the general rule for Heisenberg operators we have

$$\psi(x) \equiv \psi(\mathbf{r}t) = \exp(iHt)\psi(\mathbf{r})\exp(-iHt) \tag{7.1}$$

and similar expressions for $\bar{\psi}(x)$ and $A_\mu(x)$. Here $\psi(\mathbf{r})$ are time-independent (Schroedinger) operators. Heisenberg operators, taken at the same moments of time, satisfy the same commutation rules as Schroedinger operators and operators in the interaction representation. In fact, we have

$$\left\{\psi(\mathbf{r}t), \bar{\psi}(\mathbf{r}'t)\right\}_+ = \exp(iHt)\left\{\psi(\mathbf{r}), \bar{\psi}(\mathbf{r}')\right\}_+ \exp(-iHt) = \gamma^0\delta(\mathbf{r}-\mathbf{r}')\,. \tag{7.2}$$

In a similar way, $\psi(\mathbf{r}t)$ and $A_\mu(\mathbf{r}t)$ commute:

$$[\psi(\mathbf{r}t), A_\mu(\mathbf{r}'t)] = 0\,. \tag{7.3}$$

This is not so for operators taken at different moments of time!

The equation of motion for the Heisenberg ψ-operator is written as

$$-i\frac{\partial\psi}{\partial t} = H\psi(x) - \psi(x)H \equiv [H, \psi(x)]\,. \tag{7.4}$$

For the Hamiltonian itself the Schroedinger and Heisenberg representations coincide.

During the calculation of the right-hand side of (7.4), in the Hamiltonian we can drop the part depending only on the operator $A_\mu(x)$ (the Hamiltonian of a free electromagnetic field), as it commutes with ψ. Then

$$\begin{aligned} H &= \int d^3\mathbf{r}\psi^*(\mathbf{r}t)(\alpha\mathbf{p} + \beta m)\psi(\mathbf{r}t) + e\int d^3\mathbf{r}\bar{\psi}(\mathbf{r}t)\gamma^\mu A_\mu(\mathbf{r}t)\psi(\mathbf{r}t) \\ &= \int d^3\mathbf{r}\bar{\psi}(\mathbf{r}t)[\boldsymbol{\gamma}\cdot\mathbf{p} + m + e\gamma^\mu A_\mu(\mathbf{r}t)]\psi(\mathbf{r}t)\,. \end{aligned} \tag{7.5}$$

Calculating the commutator $[H, \psi(x)]$ using (7.2) and excluding the δ-function by $d^3\mathbf{r}$ integration, we obtain the equation of motion for the operator ψ in the explicit form

$$(\gamma^\mu p_\mu - e\gamma^\mu A_\mu - m)\psi(\mathbf{r}t) = 0\,, \tag{7.6}$$

which naturally coincides with Dirac's equation in an electromagnetic field.

Equations of motion for electromagnetic potential $A_\mu(\mathbf{r}t)$ are obvious in advance from the correspondence with the classical limit (large occupation numbers), when operator equations of motion should reduce to the usual Maxwell equations for potentials, so that in an arbitrary gauge we have

$$\partial_\mu \partial^\mu A^\nu(x) - \partial^\nu \partial_\mu A^\mu(x) = 4\pi e j^\nu(x)\,, \tag{7.7}$$

where $j^\nu(x) = \bar{\psi}(x)\gamma^\nu\psi(x)$ is the current operator, satisfying the continuity equation

$$\partial_\nu j^\nu = 0\,. \tag{7.8}$$

The system of equations (7.6) , (7.7) is the gauge invariant

$$A_\mu \to A_\mu(x) - \partial_\mu\chi(x)\,, \quad \psi(x) \to \psi(x)e^{ie\chi(x)}\,, \quad \bar{\psi}(x) \to e^{-ie\chi(x)}\bar{\psi}(x)\,, \tag{7.9}$$

where $\chi(x)$ is an arbitrary Hermitian operator, commuting (at the same moment of time) with ψ. Here we are dealing with operators in the Heisenberg representation. In interaction representation the gauge transformation of electromagnetic potential does not act on ψ operators at all!

Let us now establish the relation between operators in the Heisenberg and interaction representations. In accordance with the adiabatic hypothesis we assume that interaction $H_I(t)$ is slowly "switched on" from the time moment $t = -\infty$ to finite times. Then for $t \to -\infty$ both representations (Heisenberg and interaction) simply coincide. The corresponding wave functions (state vectors) Φ è Φ_{int} also coincide:

$$\Phi_{int}(t = -\infty) = \Phi\,. \tag{7.10}$$

On the other hand, the wave function in the Heisenberg representation does not depend on time at all (all time dependence is moved to the operators!), while in the interaction representation the time-dependence of the wave function has the form

$$\Phi_{int}(t) = S(t, -\infty)\Phi_{int}(-\infty)\,, \tag{7.11}$$

where[1]

$$S(t_2, t_1) = T\exp\left\{-i\int_{t_1}^{t_2} dt\; H_I(t)\right\} \tag{7.12}$$

with the obvious properties

$$S(t, t_1)S(t_1, t_0) = S(t, t_0)\,, \quad S^{-1}(t, t_1) = S(t_1, t)\,. \tag{7.13}$$

[1] Note, that similar operator in the previous Chapter was denoted as $U(t_2, t_1)$.

Comparing (7.11) and (7.10) we find

$$\Phi_{int}(t) = S(t, -\infty)\Phi\,, \tag{7.14}$$

which establishes the relation between the wave functions in both representations. The corresponding expression for the transformation of operators has the form

$$\psi(\mathbf{r}t) = S^{-1}(t, -\infty)\psi_{int}(\mathbf{r}t)S(t, -\infty) = S(-\infty, t)\psi_{int}(\mathbf{r}t)S(t, -\infty) \tag{7.15}$$

and a similar expression also holds for $\bar{\psi}$ and A_μ.

7.2 The exact propagator of photons

The exact photon propagator is defined as

$$\mathcal{D}_{\mu\nu}(x - x') = i\,\langle 0|TA_\mu(x)A_\nu(x')|0\rangle\,, \tag{7.16}$$

where $A_\mu(x)$ are the Heisenberg field operators, while previously we considered

$$D_{\mu\nu}(x - x') = i\,\langle 0|TA^{int}_\mu(x)A^{int}_\nu(x')|0\rangle\,, \tag{7.17}$$

built upon the operators in the interaction representation. Green's function (7.17) is usually referred to as the propagator of free photons.

Let us now express the exact propagator $\mathcal{D}_{\mu\nu}$ via the operators in the interaction representation. Consider $t > t'$; then from the relation between A_μ and A^{int}_μ of the type of (7.15), we obtain

$$\begin{aligned}\mathcal{D}_{\mu\nu}(x - x') &= i\,\langle 0|TA_\mu(x)A_\nu(x')|0\rangle \\ &= i\,\langle 0|S(-\infty, t)A^{int}_\mu(x)S(t, -\infty)S(-\infty, t')A^{int}_\nu(x')S(t', -\infty)|0\rangle\,.\end{aligned} \tag{7.18}$$

Using (7.13) we have

$$S(t, -\infty)S(-\infty, t') = S(t, t')\,, \qquad S(-\infty, t) = S(-\infty, +\infty)S(\infty, t)\,. \tag{7.19}$$

Then (7.18) is written as

$$\mathcal{D}_{\mu\nu}(x - x') = i\,\langle 0|S^{-1}[S(\infty, t)A^{int}_\mu(x)S(t, t')A^{int}_\nu(x')S(t', -\infty)]|0\rangle\,, \tag{7.20}$$

where, for brevity, we introduced

$$S = S(+\infty, -\infty) \tag{7.21}$$

and have taken into account that $S^{-1}(\infty, -\infty)S(\infty, t) = S(-\infty, t)$. As $S(t_2, t_1)$ contains only operators taken at the time moments between t_1 and t_2, placed in chronological order, it becomes obvious that all operator factors in square brackets in (7.20) are placed from left to right in the order of decreasing time. Placing the T-ordering

symbol before the bracket, in the following we can make any permutations of factors here, because the T-ordering will, in any case, place everything in the correct order. Then we can rewrite the bracket as:

$$T\left[A_\mu^{int}(x)A_\nu^{int}(x')S(\infty,t)S(t,t')S(t',-\infty)\right] = T\left[A_\mu^{int}(x)A_\nu^{int}(x')S\right]. \quad (7.22)$$

Thus, we obtain

$$\mathcal{D}_{\mu\nu}(x-x') = i\langle 0|S^{-1}TA_\mu^{int}(x)A_\nu^{int}(x')S|0\rangle\,. \quad (7.23)$$

Repeating all the previous arguments, we can show that this expression is also valid for the case of $t < t'$.

It can be shown that S^{-1} can be moved out of the averaging procedure in the form of a phase factor. In fact, a Heisenberg wave function of the vacuum Φ^0 (as any other Heisenberg function) coincides, according to (7.10), with the value of the vacuum wave function $\Phi_{int}^0(-\infty)$ in the interaction representation. On the other hand, we have

$$S\Phi_{int}^0(-\infty) \equiv S(+\infty,-\infty)\Phi_{int}^0(-\infty) = \Phi_{int}^0(+\infty)\,. \quad (7.24)$$

But a vacuum (ground state) in a stable system is strictly a stationary state in which there is no possibility of any spontaneous processes of creation or annihilation of particles. In other words, as time goes by, the vacuum remains a vacuum. This means that $\Phi_{int}^0(+\infty)$ can differ from $\Phi_{int}^0(-\infty)$ only by some phase factor $e^{i\alpha}$. Then

$$S\Phi_{int}^0(-\infty) = e^{i\alpha}\Phi_{int}^0(-\infty) = \langle 0|S|0\rangle\Phi_{int}^0(-\infty) \quad (7.25)$$

or, making a complex conjugation and taking into account the unitarity of S,

$$\Phi_{int}^{0*}(-\infty)S^{-1} = \langle 0|S|0\rangle^{-1}\Phi_{int}^{0*}(-\infty) \quad (7.26)$$

Now it is clear that (7.23) can be rewritten as

$$\mathcal{D}_{\mu\nu}(x-x') = i\frac{\langle 0|TA_\mu^{int}(x)A_\nu^{int}(x')S|0\rangle}{\langle 0|S|0\rangle}\,. \quad (7.27)$$

Substituting into the numerator and denominator perturbation expansion of the S-matrix defined by (6.56), and performing averaging with the help of the Wick theorem, we can obtain the expansion of $\mathcal{D}_{\mu\nu}$ in powers of the coupling constant e^2.

In the numerator of (7.27), the expressions being averaged differ from similar expressions for the matrix elements of the scattering matrix, as analyzed in the previous chapter, by the replacement of the "external" creation and annihilation operator of photons by the operators $A_\mu^{int}(x)$ and $A_\nu^{int}(x')$. As all factors here stand under the symbol of T-product, the pairings of these operators with "internal" operators $A_\mu^{int}(x_1)$, $A_\nu^{int}(x_2)$ will produce photon propagators $D_{\mu\nu}$. Thus, the results of averaging will be expressed by sets of diagrams with two external legs, which are constructed by the rules, given in the previous chapter, with the only difference that both the ex-

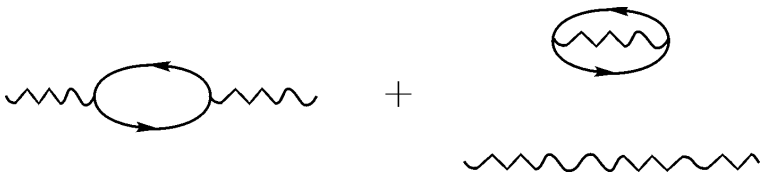


Figure 7.1

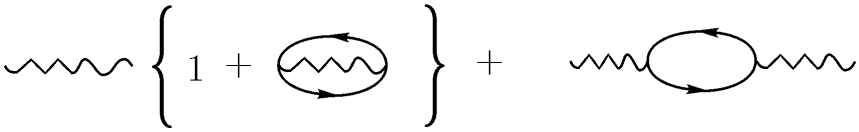


Figure 7.2

ternal and internal photon lines are now associated with the propagators $D_{\mu\nu}$, instead of the amplitudes of real photons. In a zero-order approximation, when $S = 1$, the numerator of (7.27) just coincides with $D_{\mu\nu}(x - x')$. The next nonzero terms are of the order of $\sim e^2$. These are represented by diagrams with two external legs and two vertices, as shown in Figure 7.1. The second of these diagrams consists of two disconnected parts: a wavy line (corresponding to $-iD_{\mu\nu}$) and a closed loop. This means that analytic expression for this diagram consists of two independent factors. Adding to the diagrams in Figure 7.1 the wavy line, corresponding with a zero order approximation, and moving it "outside the brackets", we obtain, up to the terms of the order of $\sim e^2$, that the numerator of (7.27) is expressed by the diagrams in Figure 7.2. The expression $\langle 0|S|0\rangle$ in the denominator of (7.27) represents the amplitude of a "vacuum–vacuum" transition. Its expansion into the perturbation series contains only diagrams with no external legs. In the zero-order approximation $\langle 0|S|0\rangle = 1$, while up to terms of the order of $\sim e^2$ this amplitude is expressed diagrammatically, as shown in Figure 7.3. Dividing (up to the same accuracy $\sim e^2$) the numerator of (7.27) by the denominator, we obtain the diagrams shown in Figure 7.4. so that the contribution of "vacuum"

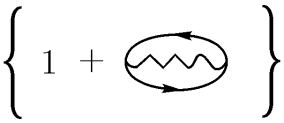


Figure 7.3

Figure 7.4

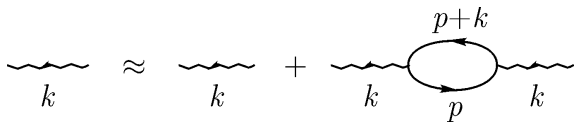


Figure 7.5

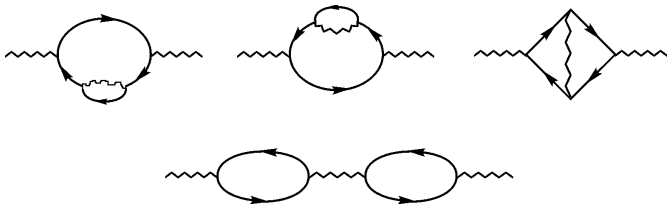


Figure 7.6

terms (under the figure brackets) is completely cancelled. Thus, the disconnected diagram in Figure 7.1(b) drops out of the answer. This result is, in fact, of a quite general nature. A more detailed analysis of the diagrams in the numerator and denominator of (7.27) shows, that the role of the denominator $\langle 0|S|0\rangle$ reduces to the general cancellation of all disconnected diagrams (in any order of perturbation theory), so that the exact propagator $\mathcal{D}_{\mu\nu}$ is expressed only by diagrams without disconnected parts, or by *connected* diagrams only!

Note that diagrams without external legs (closed loops) are of no physical importance, as these loops represent radiation corrections to the diagonal element of the S-matrix, describing vacuum–vacuum quantum transitions, and we already noted that the sum of all such loops (together with 1 from the zero-th order) produces only an irrelevant phase factor which does not influence physical results.

Transformation from coordinate to momentum representation is done in the usual way. For example, up to the terms of the order of $\sim e^2$, the propagator $-i\,\mathcal{D}_{\mu\nu}(k)$ is given by the diagrams shown in Figure 7.5, where the propagator itself is shown as a "fat" wavy line in the left-hand side. The analytic expression corresponding to these diagrams is

$$\mathcal{D}_{\mu\nu}(k) = D_{\mu\nu}(k) + ie^2 D_{\mu\lambda}(k)\left[\int \frac{d^4p}{(2\pi)^4} Sp\gamma^\lambda G(p+k)\gamma^\rho G(p)\right] D_{\rho\nu}(k)\,. \tag{7.28}$$

Terms of the higher orders are constructed in a similar way and are graphically represented by diagrams with two external photon lines and the necessary number of vertices, corresponding to the order of perturbation theory. For example, terms of the order of $\sim e^4$ are represented by diagrams with four vertices, as shown in Figure 7.6. The diagram shown in Figure 7.7 also contains four vertices, with an electron loop in its upper part. This loop corresponds to the pairing (contraction) $\bar{\psi}(x)\gamma\psi(x)$, i. e., simply to the vacuum average of the current $\langle 0|j(x)|0\rangle$. Even from the definition of the

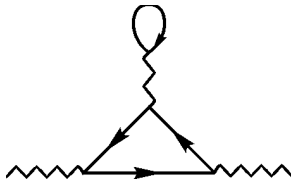


Figure 7.7

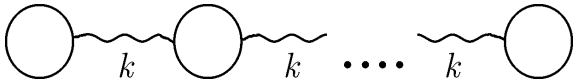


Figure 7.8

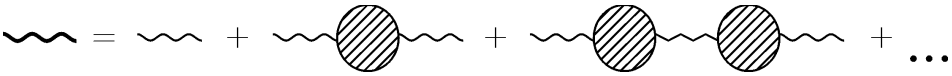


Figure 7.9

vacuum itself it is obvious, that this average should identically be zero, and this fact cannot be changed by any radiation (higher-order) corrections to this loop (though direct calculation, by the way, produces an infinite result here!). Thus, no diagrams with electron loops of this kind should be taken into account in any order of perturbation theory.

Part of the diagram ("block") between two photon lines (external or internal) is called a photon *self-energy part*[2]. In the general case, such a block can itself be divided into parts connected by one photon line, as shown in Figure 7.8, where circles denote blocks which cannot be further divided in this way. Such blocks are called *irreducible* (or single-particle irreducible). Let us denote the sum (of an infinite number!) of all irreducible photon self-energy parts as $i\mathcal{P}_{\mu\nu}/4\pi$ and call it the *polarization operator*. Classifying diagrams by the number of full irreducible self-energy parts (polarization operators), we can represent the exact photon propagator $\mathcal{D}_{\mu\nu}$ by the diagrammatic series shown in Figure 7.9, where each dashed circle represents $i\mathcal{P}_{\mu\nu}/4\pi$. The orresponding analytic expression is written as

$$\begin{aligned}\mathcal{D} &= D + D\frac{\mathcal{P}}{4\pi}D + D\frac{\mathcal{P}}{4\pi}D\frac{\mathcal{P}}{4\pi}D + \cdots \\ &= D\left\{1 + \frac{\mathcal{P}}{4\pi}\left[D + D\frac{\mathcal{P}}{4\pi}D + \cdots\right]\right\}. \end{aligned} \tag{7.29}$$

[2] Or, for brevity, just self-energy!

It is obvious that the series in square brackets again produces the complete series for $\mathcal{D}$. Thus, we obtain

$$\mathcal{D}_{\mu\nu}(k) = D_{\mu\nu}(k) + D_{\mu\lambda}(k)\frac{\mathcal{P}^{\lambda\rho}(k)}{4\pi}\mathcal{D}_{\rho\nu}(k)\,. \tag{7.30}$$

Multiplying this equality from the left side by the inverse tensor $(D^{-1})^{\tau\mu}$ and from the right side by $(\mathcal{D}^{-1})^{\nu\sigma}$, we get

$$\mathcal{D}^{-1}_{\mu\nu} = D^{-1}_{\mu\nu} - \frac{1}{4\pi}\mathcal{P}_{\mu\nu}\,. \tag{7.31}$$

Everything that was said in the previous chapter about tensor structure and gauge dependence of the free photon propagator $D_{\mu\nu}$ is also valid for the exact propagator $\mathcal{D}_{\mu\nu}$. Let us write its general form as

$$\mathcal{D}_{\mu\nu}(k) = \mathcal{D}(k^2)\left(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right) + \mathcal{D}^l(k^2)\frac{k_\mu k_\nu}{k^2}\,, \tag{7.32}$$

where $\mathcal{D}^l(k^2)$ is an arbitrary function determined by the choice of gauge. For the free propagator we write similarly

$$D_{\mu\nu}(k) = D(k^2)\left(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right) + D^l(k^2)\frac{k_\mu k_\nu}{k^2}\,, \tag{7.33}$$

which is formally different from the form used in the previous chapter, but is, in fact, equivalent to it, differing only by the definition of $D^l(k^2)$. The longitudinal part of the propagator (the second term in these expressions) is related to unphysical longitudinal part of the 4-potential and does not participate in interactions. Thus, interaction does not change it and we can always assume that

$$\mathcal{D}^l(k^2) = D^l(k^2)\,. \tag{7.34}$$

Let us now introduce inverse tensors which satisfy the following equalities:

$$\mathcal{D}^{-1}_{\mu\nu}\mathcal{D}^{\nu\lambda} = \delta^\lambda_\mu\,, \quad D^{-1}_{\mu\nu}D^{\nu\lambda} = \delta^\lambda_\mu\,. \tag{7.35}$$

For (7.32) and (7.33) the inverse tensors, taking (7.34) into account, has the form

$$\mathcal{D}^{-1}_{\mu\nu} = \frac{1}{\mathcal{D}}\left(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right) + \frac{1}{D^l}\frac{k_\mu k_\nu}{k^2}\,, \tag{7.36}$$

$$D^{-1}_{\mu\nu} = \frac{1}{D}\left(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right) + \frac{1}{D^l}\frac{k_\mu k_\nu}{k^2}\,. \tag{7.37}$$

Now it follows that the polarization operator is actually the transverse tensor:

$$\mathcal{P}_{\mu\nu} = \mathcal{P}(k^2)\left(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\,,\right) \tag{7.38}$$

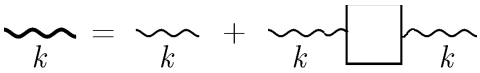


Figure 7.10

where $\mathcal{P}(k^2) = k^2 - 4\pi/\mathcal{D}(k^2)$, so that[3]

$$\mathcal{D}(k^2) = \frac{4\pi}{k^2[1 - \mathcal{P}(k^2)/k^2]}\,. \tag{7.39}$$

Thus, in contrast to the photon propagator, the polarization operator is gauge invariant!

Sometimes it is useful to introduce photon self-energy, defined as the sum of all (not only irreducible) the diagrams. Let us denote it as $i\,\Pi_{\mu\nu}/4\pi$; then we have

$$\mathcal{D}_{\mu\nu} = D_{\mu\nu} + D_{\mu\lambda}\frac{\Pi^{\lambda\rho}}{4\pi}D_{\rho\nu}\,, \tag{7.40}$$

which is shown by the diagrams in Figure 7.10. Determining now $\Pi_{\mu\lambda}$, we obtain

$$\frac{1}{4\pi}\Pi_{\mu\nu} = D^{-1}_{\mu\lambda}\mathcal{D}^{\lambda\rho}D^{-1}_{\rho\nu} - D^{-1}_{\mu\nu} \tag{7.41}$$

and, using (7.32), (7.33), (7.36), and (7.37), we get

$$\Pi_{\mu\nu} = \Pi(k^2)\left(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right), \quad \Pi = \frac{\mathcal{P}}{1 - \mathcal{P}/k^2}\,. \tag{7.42}$$

Now we see that $\Pi_{\mu\nu}$, as well as $\mathcal{P}_{\mu\nu}$, is the gauge invariant tensor.

7.3 The exact propagator of electrons

An exact electron propagator is defined as

$$\mathcal{G}(x - x') = -i\,\langle 0|T\,\psi(x)\bar{\psi}(x')|0\rangle\,, \tag{7.43}$$

which is different from the case of the free particle propagator

$$G(x - x') = -i\,\langle 0|T\,\psi^{int}(x)\bar{\psi}^{int}(x')|0\rangle \tag{7.44}$$

by replacement of the ψ-operators in interaction representation by operators in the Heisenberg representation. Similarly to the case of photon propagator, discussed above, equation (7.43) can be transformed to

$$\mathcal{G}(x - x') = -i\,\frac{\langle 0|T\,\psi^{int}(x)\bar{\psi}^{int}(x')S|0\rangle}{\langle 0|S|0\rangle}\,. \tag{7.45}$$

[3] It is useful to note that $\mathcal{P}(k^2) = \mathcal{P}^\mu_\mu(k^2)/3$.

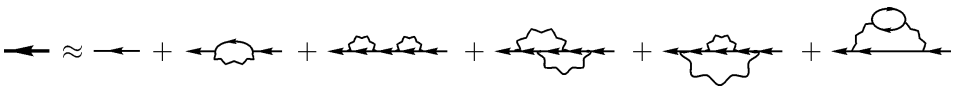


Figure 7.11

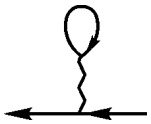


Figure 7.12

Expansion of this expression in powers of e^2 leads to a diagrammatic series for the $\mathcal{G}$-function. The role of the denominator in (7.45) again reduces to the cancellation of vacuum–vacuum transitions, so that the diagrammatic expansion for Green's function contains diagrams without an isolated vacuum loops. Up to the terms of the order $\sim e^4$ diagrams for $\mathcal{G}$ are shown in Figure 7.11, where the exact propagator itself is shown by a "fat" line. Diagrams like that in Figure 7.12, as noted before, should not be taken into account. In momentum representation, a "fat" line corresponds to $i\mathcal{G}(p)$, and the usual full and wavy lines represent propagators of free particles $iG(p)$ and $-iD(k)$.

Let us present a formal proof of the cancellation of vacuum diagrams. Consider the n-th order correction to the Green's function (propagator) of an electron, described by some disconnected diagram, representing two multiples. The first one includes all H_I, contracted with $\psi(x)$ and $\bar{\psi}(x')$, i. e., corresponds to a connected block with external legs. The second one describes the rest of the diagram. Thus, the analytic expression for this correction looks like

$$-i\frac{(-i)^n}{n!}\int dt_1\cdots\int dt_m\langle 0|T\left[\psi^{int}(x)\bar{\psi}^{int}(x')H_I(t_1)\cdots H_I(t_m)\right]|0\rangle_c \times\int dt_{m+1}\cdots\int dt_n\langle 0|T(H_I(t_{m+1})\cdots H_I(t_n))|0\rangle\,. \quad (7.46)$$

Here $\langle 0|\cdots|0\rangle_c$ and $\langle 0|\cdots|0\rangle$ corresponds to some definite set of contractions (pairings) described by the Wick theorem, and the symbol $\langle\cdots\rangle_c$ denotes that pairings in this expression produce a connected diagram.

It is easily seen, that some of the diagrams give identical contributions. In fact, if we change the pairings just by making different permutations of H_I between $\langle\cdots\rangle_c$ and $\langle\cdots\rangle$, this will simply correspond to the renaming integration variables and will not change the value of the correction to $\mathcal{G}$. The number of such diagrams is equal to the number of ways to break n operators in H_I into groups of m and $n-m$ operators, i. e., it will be equal to $\frac{n!}{m!(n-m)!}$.

The total contribution of these diagrams is given by

$$-i\frac{(-i)^m}{m!}\int dt_1\cdots\int dt_m\langle 0|T(\psi^{int}(x)\bar{\psi}^{int}(x')H_I(t_1)\cdots H_I(t_m))|0\rangle_c \times\frac{(-i)^{n-m}}{(n-m)!}\int dt_{m+1}\cdots\int dt_n\langle 0|T(H_I(t_{m+1})\cdots H_I(t_n))|0\rangle\,. \quad (7.47)$$

Let us sum the contributions of all diagrams of an arbitrary order, containing the definite connected part and the arbitrary disconnected parts. Obviously, we shall get

$$-i\frac{(-i)^m}{m!}\int dt_1\cdots\int dt_m\langle 0|T(\psi^{int}(x)\bar{\psi}^{int}(x')H_I(t_1)\cdots H_I(t_m))|0\rangle_c$$
$$\times\left\{1-i\int dt_{m+1}\langle 0|H_I(t_{m+1})|0\rangle-\frac{1}{2}\int dt_{m+1}\int dt_{m+2}\langle 0|T(H_I(t_{m+1})H_I(t_{m+2}))|0\rangle\right.$$
$$\left.+\cdots+\frac{(-i)^k}{k!}\int dt_{m+1}\cdots\int dt_{m+k}\langle 0|T(H_I(t_{m+1})\cdots H_I(t_{m+k}))|0\rangle+\cdots\right\}. \qquad (7.48)$$

Let us return to the initial expression (7.45). If we expand $\langle 0|S|0\rangle$ in the denominator in a series in powers of H_I, we shall get exactly the same expression, which stands in figure brackets in (7.48). Thus,

$$\langle 0|T\psi^{int}(x)\bar{\psi}^{int}(x')S|0\rangle=\langle 0|T\psi^{int}(x)\bar{\psi}^{int}(x')S|0\rangle_c\langle 0|S|0\rangle\,, \qquad (7.49)$$

so that, according to (7.45),

$$\mathcal{G}(x-x')=-i\langle 0|T\psi^{int}(x)\bar{\psi}^{int}(x')S|0\rangle_c\,, \qquad (7.50)$$

which completes our proof! This rule is valid for calculations of arbitrary expressions like (7.27) or (7.45) with an arbitrary number of field operators. In practice this means that we can just drop the factor of $\langle 0|S|0\rangle$ in the denominator, and do not take into account disconnected diagrams.

Further simplifications appear due to the fact that all types of pairings in

$$-i\frac{(-i)^m}{m!}\int dt_1\cdots\int dt_m\langle 0|T(\psi^{int}(x)\bar{\psi}^{int}(x')H_I(t_1)\cdots H_I(t_m))|0\rangle_c\,, \qquad (7.51)$$

differing only by permutations of H_I, give the same contributions. Because of this, we can drop the factor of $1/m!$ and take into account only those pairings which lead to *topologically* nonequivalent diagrams, i. e., those which can not be obtained from each other by permutation of operators H_I. Now the contribution of each diagram does not contain a factor with relevant dependence on the order of diagram m. Due to this fact, each diagram can be separated into elements which can be considered separately as corrections to one or another Green's function. There may be irrelevant dependences on m, such as factors like λ^m, where λ is some constant. Such factors do not prohibit separation of the diagram into different elements (blocks). On the contrary, the presence of a factor like $1/m$ does not allow such separation and separate summations within different parts of the diagram.

The block between two electron lines is called the *electron self-energy part*. As in the case of photons, it is called *irreducible* (or single particle irreducible), if it cannot be cut into two self-energies by cutting one electron line. We denote the sum of all irreducible self-energy parts as $-i\mathcal{M}(p)$, and call $\mathcal{M}(p)$ the *mass operator*. Up to terms of the order of $\sim e^4$ the mass operator is represented by diagrams shown in Figure 7.13. Performing summation, similar to that done during the derivation of equation (7.30), we obtain the *Dyson equation*,

$$\mathcal{G}(p)=G(p)+G(p)\mathcal{M}(p)\mathcal{G}(p) \qquad (7.52)$$

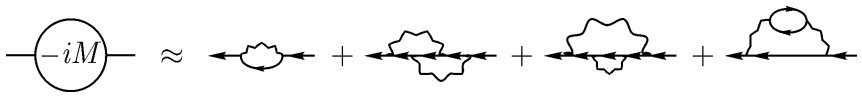


Figure 7.13

or, in terms of inverse matrices,

$$\mathcal{G}^{-1}(p) = G^{-1}(p) - \mathcal{M}(p) = \gamma^{\mu} p_{\mu} - m - \mathcal{M}(p)\,. \tag{7.53}$$

Equation (7.30) can also be called a Dyson equation for a photon propagator. Below we shall return many times to the discussion of these equations.

Heisenberg ψ-operators (in contrast to ψ-operators in interaction representation), as previously noted, *do change* under gauge transformations. In a similar way, the exact telectron propagator $\mathcal{G}$ is also *not* a gauge invariant quantity. It is clear that the change of $\mathcal{G}$ under gauge transformations should be expressible via the same arbitrary function D^l, which is added to the photon propagator. It is clear from the fact that, during the calculations of $\mathcal{G}$ via a diagrammatic perturbation series, any term in this series is expressed via the photon Green's functions D, and no other terms related to an electromagnetic field are present at all. We can make some assumptions about the properties of the operator χ in (7.9), with the only limitation that the result be expressed via D^l. Under (7.9) the propagators $\mathcal{D}$ and $\mathcal{G}$ transform into

$$\mathcal{D}_{\mu\nu} \to i\,\langle 0|T[A_{\mu}(x) - \partial_{\mu}\chi(x)][A_{\nu} - \partial'_{\nu}\chi(x')]|0\rangle\,, \tag{7.54}$$

$$\mathcal{G} \to -i\,\langle 0|T\psi(x)e^{ie\chi(x)}e^{-ie\chi(x')}\bar{\psi}(x')|0\rangle\,. \tag{7.55}$$

We shall assume that operators χ are averaged over the vacuum independently of others, which is natural, as gauge invariance of electrodynamics requires that the "field" χ does not take part in the interactions. Now put also $\langle 0|\chi(x)|0\rangle = 0$. Then, in equations (7.54) and (7.55), terms containing χ are separated and we obtain

$$\mathcal{D}_{\mu\nu} \to \mathcal{D}_{\mu\nu} + i\,\langle 0|T\,\partial_{\mu}\chi(x)\partial'_{\nu}\chi(x')|0\rangle\,, \tag{7.56}$$

$$\mathcal{G} \to \mathcal{G}\,\langle 0|Te^{ie\chi(x)}e^{-ie\chi(x')}|0\rangle\,. \tag{7.57}$$

Let us stress once again that here the χ are *operators*. Next we consider the case of infinitesimal gauge transformations and introduce $\delta\chi$ instead of χ. Transformation (7.56), independently of the smallness of $\delta\chi$, can be written as

$$\mathcal{D}_{\mu\nu} \to \mathcal{D}_{\mu\nu} + \delta\mathcal{D}_{\mu\nu}\,, \quad \delta\mathcal{D}_{\mu\nu} = \partial_{\mu}\partial'_{\nu}d^l(x - x')\,, \tag{7.58}$$

where

$$d^l(x - x') = i\,\langle 0|T\delta\chi(x)\delta\chi(x')|0\rangle\,. \tag{7.59}$$

Now we see that d^l determines the change of the longitudinal part of the photon propagator $\mathcal{D}^l$ under the gauge transformation.

In (7.57) we can expand the exponents in powers of $\delta\chi$ up to quadratic terms, so that

$$\langle 0|Te^{ie\delta\chi(x)}e^{-ie\delta\chi(x')}|0\rangle \approx 1 - \frac{1}{2}e^2\langle 0|\delta\chi^2(x) + \delta\chi^2(x') - 2T\delta\chi(x)\delta\chi(x')|0\rangle\,. \tag{7.60}$$

Taking into account the definition (7.59) we get

$$\mathcal{G} \to \mathcal{G} + \delta\mathcal{G}\,, \quad \delta\mathcal{G} = ie^2\mathcal{G}(x-x')[d^l(0) - d^l(x-x')]\,. \tag{7.61}$$

In momentum representation

$$\delta\mathcal{G}(p) = ie^2\int\frac{d^4q}{(2\pi)^4}[\mathcal{G}(p) - \mathcal{G}(p-q)]d^l(q)\,, \tag{7.62}$$

with

$$q^2d^l(q) = \delta\mathcal{D}^l(q)\,. \tag{7.63}$$

These expressions give the general gauge transformation rules for exact propagators in QED.

7.4 Vertex parts

Besides the self-energy parts, in more complicated diagrams we can introduce additional blocks with special physical meaning. Consider the function

$$K^\mu(x_1,x_2,x_3) = \langle 0|TA^\mu(x_1)\psi(x_2)\bar\psi(x_3)|0\rangle\,. \tag{7.64}$$

Due to the homogeneity of space-time, this function depends only on differences of its arguments. After transformation to interaction representation we have

$$K^\mu(x_1,x_2,x_3) = \frac{\langle 0|TA^\mu_{int}(x_1)\psi^{int}(x_2)\bar\psi^{int}(x_3)S|0\rangle}{\langle 0|S|0\rangle}\,. \tag{7.65}$$

In momentum representation we can write

$$K^\mu(p_2,p_1;k)(2\pi)^4\delta(p_1+k-p_2) = \int d^4x_1\int d^4x_2\int d^4x_3 e^{-ikx_1+ip_2x_2-ip_1x_3}K^\mu(x_1,x_2,x_3)\,. \tag{7.66}$$

In the diagram technique function K^μ is described by a "three leg" graph, shown in Figure 7.14, with one photon and two electron legs, with the 4-momenta satisfying the conservation law

$$p_1 + k = p_2\,. \tag{7.67}$$

The zero-th order term in the perturbation expansion of this function is obviously zero,

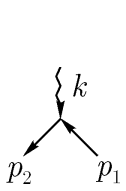


Figure 7.14

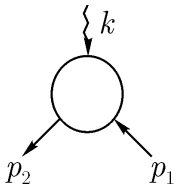


Figure 7.15

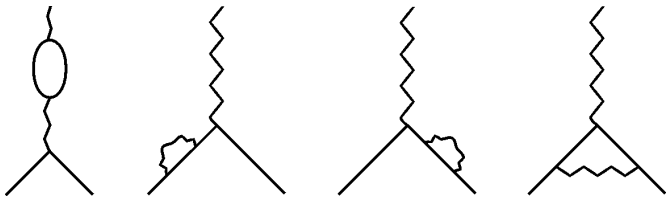


Figure 7.16

while the first-order term in coordinate representation is

$$K^{\mu}(x_1, x_2, x_3) = e \int d^4x G(x_2 - x)\gamma_{\nu} G(x - x_3) D^{\nu\mu}(x_1 - x) \tag{7.68}$$

or, in momentum representation

$$K^{\mu}(p_2, p_1; k) = eG(p_2)\gamma_{\nu} G(p_1) D^{\nu\mu}(k)\,, \tag{7.69}$$

which is shown by the diagram in Figure 7.15. In higher orders, diagrams become more complicated due to the addition of extra vertices. For example, in the third order the diagrams shown in Figure 7.16 appear. In the first three diagrams of Figure 7.16 we can separate the obvious self-energy parts of the photon and electrons. However, there are no such blocks in the fourth diagram. This is a general situation: corrections of self-energy type simply replace in (7.69) Green's functions G and D by $\mathcal{G}$ and $\mathcal{D}$. The sum of the remaining terms of expansion lead to the change of the factor γ^{μ} in (7.69). Denoting this quantity as Γ^{μ} we have, by definition,

$$K^{\mu}(p_2, p_1; k) = \{i\mathcal{G}(p_2)[-ie\Gamma_{\nu}(p_2, p_1; k)]i\mathcal{G}(p_1)\}\,[-i\mathcal{D}^{\nu\mu}(k)]\,. \tag{7.70}$$

The block connected with the other parts of diagram by a single photon and two electron lines is called the *vertex part*, if this block can not be separated into parts, which are connected to each other by single (electron or photon) lines. Block Γ^{μ}, introduced above, representing the sum of all possible vertex parts, including the simple vertex γ^{μ}, is called the vertex operator (or vertex function). Up terms of the fifth order it is expressed by the diagrams shown in Figure 7.17. All three momenta here cannot be simultaneously related to real particles: we have already seen that absorption (emission) of a photon by the free electron is impossible because of 4-momentum conservation.

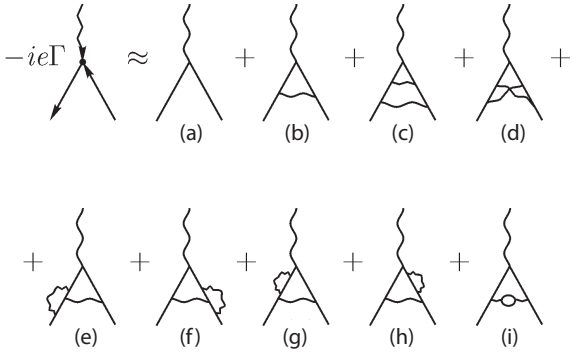


Figure 7.17

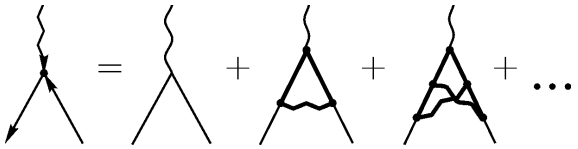


Figure 7.18

Thus, one of the legs in this graph can only be related a virtual particle (or external field).

We can now introduce the notions of compact and noncompact vertex parts. Those vertex parts which do not contain self-energy corrections to internal lines are called compact, and we cannot separate the parts representing corrections to the internal vertices. Among the graphs shown in Figure 7.17, only diagrams (b) and (d) are compact. Diagrams (g,h,i) contain self-energy corrections either to the electron or the photon lines. In diagram (c) the upper horizontal wavy line can be considered as a correction to the upper vertex, while the wavy lines at the sides of diagrams (e) and (f) represent corrections to vertices at these sides. Replacing internal lines in compact diagrams by "fat" lines representing exact Green's functions, we obtain the expansion of the vertex operator in the form, shown in Figure 7.18, which is usually called a "skeleton" diagram expansion. This expansion, in fact, produces an integral equation for Γ, but with an infinite number of terms in the right-hand side; there is no closed equation for the vertex parts, similar to that of the Dyson equation for Green's functions (propagators).

We can also introduce vertices with a larger number of external legs, e. g., the "four-leg" vertex shown in Figure 7.19. We can obtain such a vertex considering the function

$$K(x_1, x_2; x_3, x_4) = \langle 0|T\psi(x_1)\psi(x_2)\bar{\psi}(x_3)\bar{\psi}(x_4)|0\rangle\,, \tag{7.71}$$

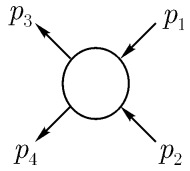


Figure 7.19

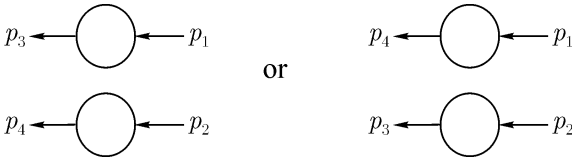


Figure 7.20

which is usually called a *two particle* Green's function. It also depends on differences of its arguments, and its Fourier transformation can be written as

$$\int d^4x_1 \int d^4x_2 \int d^4x_3 \int d^4x_4 K(x_1, x_2; x_3, x_4) e^{i(p_3x_1 + p_4x_2 - p_1x_3 - p_2x_4)} = (2\pi)^4 \delta(p_1 + p_2 - p_3 - p_4) K(p_3, p_4; p_1, p_2), \quad (7.72)$$

where

$$K(p_3, p_4; p_1, p_2) = (2\pi)^4 \delta(p_1 - p_3) \mathcal{G}(p_1) \mathcal{G}(p_2) - (2\pi)^4 \delta(p_2 - p_3) \mathcal{G}(p_1) \mathcal{G}(p_2) + \mathcal{G}(p_3) \mathcal{G}(p_4) [-i\Gamma(p_3, p_4; p_1, p_2)] \mathcal{G}(p_1) \mathcal{G}(p_2). \quad (7.73)$$

The first two terms here exclude from the definition of $\Gamma(p_3, p_4; p_1, p_2)$ diagrams like those shown in Figure 7.20. Similarly, in the third term in (7.73), the factors $\mathcal{G}$ exclude from the definition of Γ-vertex those graphs which represent corrections to external electron lines. Using the properties of the T-product of fermion operators, we may easily see that $\Gamma(p_3, p_4; p_1, p_2)$ has the following (anti)symmetry properties:

$$\Gamma(p_3, p_4; p_1, p_2) = -\Gamma(p_4, p_3; p_1, p_2) = -\Gamma(p_3, p_4; p_2, p_1). \quad (7.74)$$

This vertex describes, e. g., the process of scattering of two electrons; its amplitude can be found if we associate with the external legs the amplitudes of the initial and final particles (instead of propagators $\mathcal{G}$):

$$iM_{fi} = \bar{u}(p_3)\bar{u}(p_4)[-ie\Gamma(p_3, p_4; p_1, p_2)]u(p_1)u(p_2), \quad (7.75)$$

and here Γ describes all the possible interaction processes in all orders of perturbation theory.

Figure 7.21

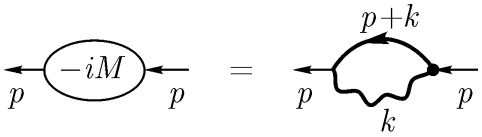


Figure 7.22

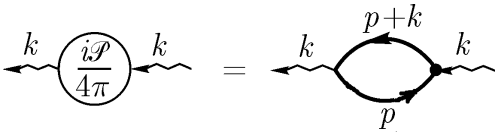


Figure 7.23

7.5 Dyson equations

Exact propagators and vertex parts are connected to each other, as we have already seen, by certain integral relations. Let us analyze these relations in more detail. Consider diagrams for irreducible self-energies of an electron. It is easy to see that among the infinite number of these diagrams, only one, shown in Figure 7.21, is compact in the sense discussed in the previous section, while the others can be considered as introducing corrections to one of its vertices. It is clear that all vertex corrections should be attributed only to one (any of two) vertices of this diagram, while the other remains "bare" (to avoid double counting). Correspondingly, the sum of all irreducible self-energy parts (i. e., the mass operator) can be expressed by only one skeleton diagram, shown in Figure 7.22. The appropriate analytic expression has the form

$$\mathcal{M}(p) = G^{-1}(p) - \mathcal{G}^{-1}(p) = -ie^2 \int \frac{d^4k}{(2\pi)^4} \gamma^\nu \mathcal{G}(p+k)\Gamma^\mu(p+k, p; k)\mathcal{D}_{\mu\nu}(k)\,. \tag{7.76}$$

Similar expression can be also written for the polarization operator. Among irreducible self-energies for a photon, again only one is compact, and the polarization operator is represented by the diagram shown in Figure 7.23. the orresponding analytic expression is

$$\frac{1}{4\pi}\mathcal{P}_{\mu\nu}(k) = D^{-1}_{\mu\nu}(k) - \mathcal{D}^{-1}_{\mu\nu}(k) = ie^2 Sp \int \frac{d^4p}{(2\pi)^4} \gamma_\mu \mathcal{G}(p+k)\Gamma_\nu(p+k, p; k)\mathcal{G}(p)\,. \tag{7.77}$$

Equations (7.76) and (7.77) give an explicit form of the Dyson equations (7.52) and (7.30), which are integral equations for exact propagators, expressing them via exact vertex parts. However, there are no similar "closed" integral equations for vertex parts, so that in practical cases we have to solve the Dyson equations using different types of approximations for the vertex part, e. g., based partial summation of Feynman diagrams.

7.6 Ward identity

There are certain exact relations between propagators and vertices which are simpler than Dyson type equations. Consider the electron propagator. Let us make a gauge transformation (7.9), assuming $\chi(x) = \delta\chi(x)$, where $\delta\chi(x)$ is an infinitesimal *non-operator* function of the coordinates x. Then, the electron propagator will change as

$$\delta\mathcal{G}(x, x') = ie\mathcal{G}(x - x')[\delta\chi(x) - \delta\chi(x')]\,. \tag{7.78}$$

Such a gauge transformation breaks the homogeneity of space-time, and $\delta\mathcal{G}$ now depends on x and x' separately, not only on $x - x'$. Now we have to make a Fourier transformation over x and x' separately, so that in momentum representation $\delta\mathcal{G}$ becomes the function of two 4-momenta:

$$\delta\mathcal{G}(p_2, p_1) = \int d^4x \int d^4x' \delta\mathcal{G}(x, x') e^{ip_2x - ip_1x'}\,. \tag{7.79}$$

Substituting here (7.78) and integrating over $d^4x d^4\xi$ or $d^4x' d^4\xi$, where $\xi = x - x'$, we obtain

$$\delta\mathcal{G}(p + q, p) = ie\delta\chi(q)[\mathcal{G}(p) - \mathcal{G}(p + q)]\,. \tag{7.80}$$

On the other hand, the same gauge transformation applied to the operator of electromagnetic vector-potential $A_\mu(x)$ produces

$$\delta A^{(e)}_\mu(x) = -\frac{\partial}{\partial x^\mu}\delta\chi\,, \tag{7.81}$$

which may be considered as an infinitesimal *external field*. In momentum representation,

$$\delta A^{(e)}_\mu(q) = iq_\mu\delta\chi(q)\,. \tag{7.82}$$

The value of $\delta\mathcal{G}$ can be also calculated as the change of propagator under the influence of this field. Up to terms of the first order over $\delta\chi$, this change can be expressed by the single skeleton diagram shown in Figure 7.24, where the "fat" wavy line denotes the effective external field

$$\delta A^{(e)}_\mu(q) + \delta A^{(e)}_\lambda(q)\frac{1}{4\pi}\mathcal{P}^{\lambda\nu}(q)\mathcal{D}_{\nu\mu}(q)\,, \tag{7.83}$$

which takes into account the self-energy corrections. However, the 4-vector $\delta A^{(e)}_{\lambda}(q)$ is longitudinal (with respect to q), while the tensor $\mathcal{P}^{\lambda\nu}$ is transversal (cf. (7.38, (7.42)). Thus, the second term here simply gives zero, so what remains is, in fact, the contribution of the diagram shown in Figure 7.24, where the line of the external field can be taken as "thin" and equal to $\delta A^{(e)}_{\mu}(q)$. In analytic form,

$$\delta\mathcal{G}(p+q,p) = e\mathcal{G}(p+q)\Gamma^{\mu}(p+q,p;q)\mathcal{G}(p)\delta A^{(e)}_{\mu}(q)\,. \tag{7.84}$$

Substituting here (7.82) and comparing with (7.80), we find

$$\mathcal{G}(p+q) - \mathcal{G}(p) = -\mathcal{G}(p+q)q_{\mu}\Gamma^{\mu}(p+q,p;q)\mathcal{G}(p) \tag{7.85}$$

or, in terms of inverse matrices,

$$\mathcal{G}^{-1}(p+q) - \mathcal{G}^{-1}(p) = q_{\mu}\Gamma^{\mu}(p+q,p;q)\,. \tag{7.86}$$

For $q \to 0$, comparing the coefficients before infinitesimal q_{μ} in both sides of this relation, we get

$$\frac{\partial}{\partial p_{\mu}}\mathcal{G}^{-1}(p) = \Gamma^{\mu}(p,p;0) \tag{7.87}$$

which is called the *Ward identity* in differential form. Relation (7.86) is also called a Ward identity, but for finite q. From (7.87) we can see that the derivative of $\mathcal{G}^{-1}(p)$ over the momentum coincides with a vertex operator with zero momentum transfer. The derivative of Green's function $\mathcal{G}(p)$ itself is equal to

$$-\frac{\partial}{\partial p_{\mu}} i\mathcal{G}(p) = i\mathcal{G}(p)[-i\,\Gamma^{\mu}(p,p;0)]i\mathcal{G}(p) \tag{7.88}$$

In zero-th approximation this identity is obvious, as from $G^{-1} = \gamma^{\mu}p_{\mu} - m$ we immediately obtain $\frac{\partial G^{-1}}{\partial p_{\mu}} = \gamma^{\mu}$. Now it is easy to obtain a diagrammatic derivation of Ward identity: from Dyson equation (7.53) it is obvious, that the differentiation of the inverse Green's function over the momentum is equivalent to all possible insertions of lines of a fictitious external field, with zero momentum transfer, into all diagrams for irreducible self-energy, which generates all the diagrams for the corresponding vertex part. The Ward identity is of great importance for checking the self-consistency of concrete approximations in different problems of quantum field theory.

A little more technical is a similar derivation of similar identities for an exact photon propagator (polarization operator). The details of this derivation can be found in [6].

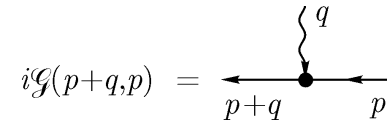


Figure 7.24

Chapter 8

Some applications of quantum electrodynamics

8.1 Electron scattering by static charge: higher order corrections

In this chapter we shall consider the calculations of some specific effects of quantum electrodynamics (QED), as well as some conceptual problems related to the foundations of QED. It should be noted that QED is actually an example of the quite successful theory of interacting elementary particles. It allows exceptionally precise calculations of different effects due to electromagnetic interactions, which are in an ideal agreement with current. The detailed analysis of the vast number of QED effects can be found in [2, 6], while here we shall limit ourselves to only few of the most typical cases. During our discussion we shall more or less skip the technical details, concentrating on the qualitative aspects of the theory.

Let us return to the previously discussed problem (cf. Chapter 4) of electron scattering by static charge of the nuclei (Rutherford scattering). In the first order of perturbation theory this scattering process is described by the diagram shown in Figure 8.1(a), where the static charge is denoted by a cross. According to the general rules of diagram technique, the corresponding scattering amplitude is written as

$$M_{fi} = -i \int d^4x \langle f | j_\mu(x) | i \rangle A^\mu(x) \,, \tag{8.1}$$

where the matrix element of transition current is

$$\langle f | j_\mu(x) | i \rangle = e \bar{u}_f \gamma_\mu u_i e^{-iqx} \,, \tag{8.2}$$

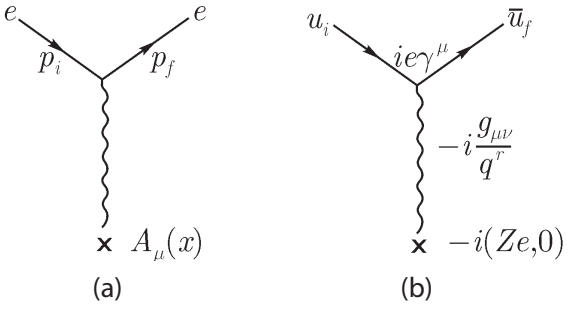


Figure 8.1

with $q = p_i - p_f$, and we introduced the spinors of an initial and final state of an electron. Vector potential $A_\mu(x)$ describes electromagnetic field of the static charge. Then we can write

$$M_{fi} = -ie\bar{u}_f \gamma_\mu u_i A^\mu(q)\,, \tag{8.3}$$

where

$$A^\mu(q) = \int d^4x\, e^{-iqx} A^\mu(x)\,. \tag{8.4}$$

For a static charge, the value of $A^\mu(x)$ is time independent, so that

$$A^\mu(q) = \int dt\, e^{-i(E_i - E_f)t} \int d^3\mathbf{r}\, e^{i\mathbf{qr}} A_\mu(\mathbf{r}) = 2\pi\delta(E_f - E_i) A^\mu(\mathbf{q})\,. \tag{8.5}$$

The static Maxwell equation is written as

$$\nabla^2 A^\mu(\mathbf{r}) = -4\pi j^\mu(\mathbf{r})\,. \tag{8.6}$$

Then we have

$$A^\mu(\mathbf{q}) = \frac{4\pi}{|\mathbf{q}|^2} j^\mu(\mathbf{q})\,. \tag{8.7}$$

Accordingly, from (8.3) and (8.5) we obtain

$$M_{fi} = -2\pi i\delta(E_f - E_i) e\bar{u}_f \gamma_\mu u_i \frac{4\pi}{|\mathbf{q}|^2} j^\mu(\mathbf{q})\,. \tag{8.8}$$

To shorten expressions to follow, we drop the δ-function for the conservation law and define the amplitude M as

$$-iM = ie\bar{u}_f \gamma_\mu u_i \frac{4\pi}{|\mathbf{q}|^2} j^\mu(\mathbf{q})\,. \tag{8.9}$$

During static charge scattering, an electron changes its momentum so that $\mathbf{p}_i \neq \mathbf{p}_f$, but energy is conserved, and $E_i = E_f$, or $q_0 = 0$. Thus,

$$q^2 = -|\mathbf{q}|^2 < 0 \tag{8.10}$$

is a space-like scattering vector, and (8.9) is rewritten as

$$-iM = (ie\bar{u}_f \gamma^\mu u_i) \left(\frac{-4\pi i g_{\mu\nu}}{q^2} \right) (-ij^\nu(\mathbf{q}))\,. \tag{8.11}$$

Here the first factor describes the vertex part, while the second represents the photon propagator. For static nuclei with charge Ze we have

$$j^0(\mathbf{r}) = \rho(\mathbf{r}) = Ze\delta(\mathbf{r})\,, \qquad \mathbf{j}(\mathbf{r}) = 0\,, \tag{8.12}$$

so that

$$-iM = (ie\bar{u}_f\gamma^0 u_i)\left(\frac{-4\pi i}{q^2}\right)(-iZe)\,, \tag{8.13}$$

which is expressed by the diagram in Figure 8.1(a) and coincides in fact with (4.75). These expressions describe Rutherford scattering and the corresponding crossection is given by (4.79):

$$\frac{d\sigma}{d\Omega} \sim |M|^2 \sim q^{-4} \sim \frac{1}{\sin^4\frac{\theta}{2}}\,, \tag{8.14}$$

where θ is the scattering angle determined by kinematics:

$$q^2 = (p_i - p_f)^2 \approx -2k^2(1-\cos\theta) = -4k^2\sin^2\frac{\theta}{2}\,, \tag{8.15}$$

where we have neglected electron mass (in comparison to that of nuclei) and introduced $k \equiv |\mathbf{p}_i| = |\mathbf{p}_f|$.

This is the result of the first order perturbation theory. Let us discuss higher order (radiation) corrections. Let us consider as an example the third order diagram, shown in Figure 8.2. Using the general rules of diagram technique we obtain the corresponding

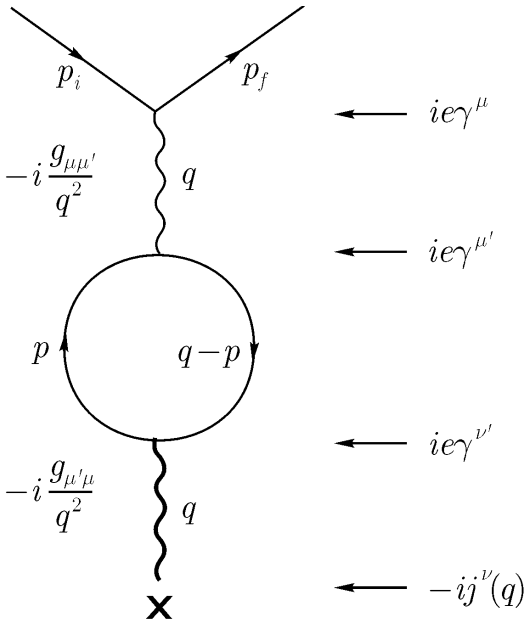


Figure 8.2

analytic expression

$$-iM = (-1)(ie\bar{u}_f\gamma^\mu u_i)\left(-i\frac{4\pi g_{\mu\mu'}}{q^2}\right)\int\frac{d^4p}{(2\pi)^4}Sp \tag{8.16}$$
$$\times\left\{(ie\gamma^{\mu'})\frac{i(\hat{p}+m)}{p^2-m^2}(ie\gamma^{\nu'})\frac{i(\hat{q}-\hat{p}+m)}{(q-p)^2-m^2}\right\}\left(-i\frac{4\pi g_{\nu'\nu}}{q^2}\right)(-ij^\nu(\mathbf{q}))\,.$$

In comparison to the first order result (8.11), here we observe the obvious modification of the photon propagator by the single-loop polarization "insertion", so that

$$-i\frac{4\pi g_{\mu\nu}}{q^2}\to -i\frac{4\pi g_{\mu\nu}}{q^2}+\left(-i\frac{4\pi g_{\mu\mu'}}{q^2}\right)I^{\mu'\nu'}\left(-i\frac{4\pi g_{\nu'\nu}}{q^2}\right)=$$
$$-i\frac{4\pi g_{\mu\nu}}{q^2}+\frac{(-4\pi i)}{q^2}I_{\mu\nu}(q^2)\frac{(-4\pi i)}{q^2}\,, \tag{8.17}$$

where

$$I_{\mu\nu}(q^2)=(-1)\int\frac{d^4p}{(2\pi)^4}Sp\left\{(ie\gamma^\mu)\frac{i(\hat{p}+m)}{p^2-m^2}(ie\gamma^\nu)\frac{i(\hat{q}-\hat{p}+m)}{(q-p)^2-m^2}\right\}. \tag{8.18}$$

Immediately we see that for $|p|\to\infty$ the integral in $I_{\mu\nu}$ contains the contribution of a term (from the polarization loop) like $\int dp\frac{p^3}{p^2}$, which is seems to be quadratically divergent at the upper limit. This is a a typical divergence that appears in higher orders of perturbation theory in practically every model of quantum field theory. The physical origin of this divergence is obviously related to the point-like (local) nature of field interactions in relativistic theory. In fact, divergence here is weaker (logarithmic), but the problem remains. Below we shall discuss its qualitative aspects.

Direct, but rather tedious, calculations show [2] that $I_{\mu\nu}$ can be written as

$$I_{\mu\nu}(q^2)=-ig_{\mu\nu}q^2I(q^2)+\cdots\,, \tag{8.19}$$

where

$$I(q^2)=\frac{e^2}{3\pi}\int_{m^2}^{\infty}\frac{dp^2}{p^2}-\frac{2e^2}{\pi}\int_0^1 dz\,z(1-z)\ln\left[1-\frac{q^2z(1-z)}{m^2}\right], \tag{8.20}$$

and the multiple dot in (8.19) replaces the terms proportional to $q_\mu q_\nu$, which give a zero contribution after tensor contraction of the photon propagator with external charges (currents). The first term in (8.19) gives precisely the logarithmic divergence of the polarization loop[1].

It is useful to explicitly write expressions for $I(q^2)$ in the limits of large and small $(-q)^2$. To make the integral sensible we introduce in the first term of (8.20) the upper limit *cutoff* Λ^2 (with dimensionality of momentum (mass) $\Lambda^2\gg m^2$ squared). Then,

[1] Logarithmic, not quadratic, divergence here is due to some "hidden" algebra of the integrand [2].

for $(-q^2) \ll m^2$ we have

$$\ln\left[1 - \frac{q^2 z(1-z)}{m^2}\right] \approx -\frac{q^2 z(1-z)}{m^2} \tag{8.21}$$

and accordingly,

$$I(q^2) \approx \frac{e^2}{3\pi} \ln\left(\frac{\Lambda^2}{m^2}\right) + \frac{e^2}{15\pi}\frac{q^2}{m^2}\,. \tag{8.22}$$

For $(-q^2) \gg m^2$ we have

$$\ln\left[1 - \frac{q^2 z(1-z)}{m^2}\right] \approx \ln\left(\frac{-q^2}{m^2}\right), \tag{8.23}$$

so that

$$I(q^2) \approx \frac{e^2}{3\pi} \ln\left(\frac{\Lambda^2}{m^2}\right) - \frac{e^2}{3\pi} \ln\left(\frac{-q^2}{m^2}\right) = \frac{e^2}{3\pi} \ln\left(\frac{\Lambda^2}{-q^2}\right). \tag{8.24}$$

Now we can write the scattering amplitude with a single-loop correction at $(-q^2) \ll m^2$ in the form[2]

$$-iM = (ie\bar{u}_f \gamma_0 u_i)\left(-\frac{4\pi i}{q^2}\right)\left[1 - \frac{e^2}{3\pi}\ln\left(\frac{\Lambda^2}{m^2}\right) - \frac{e^2}{15\pi}\frac{q^2}{m^2} + O(e^4)\right](-iZe)\,. \tag{8.25}$$

This expression can be rewritten *with the same accuracy* as

$$-iM = (ie_R\bar{u}_f \gamma_0 u_i)\left(-\frac{4\pi i}{q^2}\right)\left[1 - \frac{e_R^2}{15\pi}\frac{q^2}{m^2}\right](-iZe_R)\,, \tag{8.26}$$

where we introduced the *renormalized* charge

$$e_R = e\left(1 - \frac{e^2}{3\pi}\ln\frac{\Lambda^2}{m^2}\right)^{1/2}. \tag{8.27}$$

Let us assume that the value of e_R from (8.27) represents the "true" (experimentally measurable) electric charge. Then the scattering amplitude (8.26) becomes *finite*, and its divergence is "concealed" in e_R, which is taken from the experiment and is not calculable within our theory. Thus we have explicitly performed the *renormalization* of the divergent radiation correction. In the following we shall see that in QED all divergences which appearing in higher orders of perturbation theory can be similarly "hidden" in the *finite* number of parameters, which should be determined experimentally. This reflects the fundamental property of the *renormalizability* of this theory. Only renormalizable models of quantum field theory are physically sensible.

[2] Perturbation expansion here is in powers of dimensionless parameter $e^2 \to \frac{e^2}{\hbar c} \approx \frac{1}{137}$.

8.2 The Lamb shift and the anomalous magnetic moment

The first term in equation (8.26) is obviously due to the Coulomb potential

$$V_0(r) = -Ze_R^2 \int \frac{d^3\mathbf{q}}{(2\pi)^3} e^{i\mathbf{qr}} \frac{4\pi}{|\mathbf{q}|^2} = -\frac{Ze_R^2}{r}\,. \tag{8.28}$$

The second term in (8.26) corresponds to quantum corrections to the Coulomb potential, related to the possibility of creating virtual e^+e^--pairs. The factor of $|\mathbf{q}|^2$ there, after the transformation to coordinate representation, is replaced by $-\nabla^2$. Then, taking into account (8.28) and the Fourier expansion of δ-function

$$\delta(\mathbf{r}) = \int \frac{d^3\mathbf{q}}{(2\pi)^3} e^{i\mathbf{qr}}\,, \tag{8.29}$$

or using the well-known relation [33] $\nabla^2 \frac{1}{r} = -4\pi\delta(\mathbf{r})$, we can see that equation (8.26) in coordinate representation corresponds to the interaction of the form

$$V(r) = -Ze_R^2 \left(1 - \frac{e_R^2}{60\pi^2 m^2}\nabla^2\right)\frac{1}{r} = -\frac{Ze_R^2}{r} - \frac{Ze_R^4}{15\pi m^2}\delta(\mathbf{r})\,. \tag{8.30}$$

Thus, the creation of virtual e^+e^--pairs (vacuum polarization) leads to the modification of Coulomb interaction at small distances, corresponding to some additional attraction to nuclei. Obviously this expression is not rigorous, being obtained from the asymtotics of a single-loop contribution in the limit of $(-q)^2 \ll m^2$. However, it is sufficient for simple estimates.

Consider the case of $Z = 1$ (proton). It is clear that the second term in equation (8.30) can lead to the shift of energy levels E_{nl} of hydrogen. Considering this term as perturbation, we easily obtain this shift as

$$\Delta E_{nl} = -\frac{e_R^4}{15\pi m^2}|\psi_{nl}(0)|^2\delta_{l0} = -\frac{8e_R^6}{15\pi n^3}\ \mathrm{Ry}\ \delta_{l0}\,, \tag{8.31}$$

where $\psi_{nl}(0)$ is a hydrogen wave function, corresponding to the main quantum number n and orbital moment l, and $\mathrm{Ry} = me^4/2$ is the Rydberg constant ($\mathrm{Ry} \approx 13.6\,\mathrm{eV}$). Due to the point-like nature of additional interaction in (8.30), it acts upon only the wave functions, which are nonzero at the nuclei (proton), i. e., upon the s-states (with $l = 0$). A corresponding shift of the levels is observed experimentally and measured with high accuracy. In some first experiments Lamb measured the energy difference between the $2s_{1/2}$ and $2p_{1/2}$ levels, which are degenerate according to the Schroedinger–Dirac theory, which does not take radiation correction into account. The observed value of the shift is equal to $+1057\,\mathrm{MHz}$. Calculations with equation (8.31) give the shift of $-27\,\mathrm{MHz}$. However, we should note that the contribution of vacuum polarization is responsible only for the part of the shift between $2s_{1/2}$ and $2p_{1/2}$. The complete set of Feynman diagrams responsible for the Lamb shift in this order of perturbation theory ($\sim e^3$) is shown in Figure 8.3. All divergences appearing in

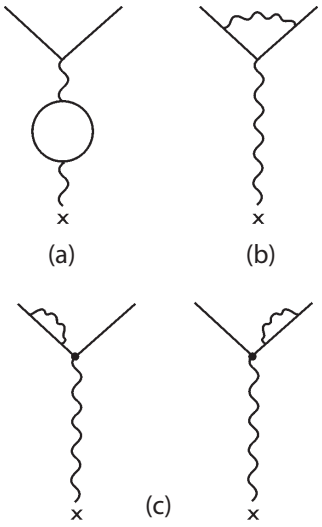


Figure 8.3

these diagrams can be "hidden" in electron charge, mass, and wave function renormalization. This allows the calculation of total Lamb shift, giving the result in an ideal correspondence with the experiment[3]. This was a triumph of the renormalization approach in QED. As the value of the Lamb shift is known with an accuracy of the order of $\sim 0.01\%$, one is easily convinced of the importance of the contributions of each of the diagrams in Figure 8.3, including the relatively small contribution of vacuum polarization expressed by the diagram in Figure 8.3(a). The main contribution is due to the renormalization of the electron mass (diagram in Figure 8.3(c)). Physically this effect is due to the fact that the value (formally infinite!) of radiation corrections to the mass of a free electron is different from that for an electron bound within an atom (which is also infinite!). The difference of these infinite corrections is *finite* [18, 60] and produces the main contribution to the shift of atomic levels.

Consider in more detail the effects connected with the diagram in Figure 8.3(b). In fact this diagram modifies the structure of electron transition current (vertex) $-e\bar{u}_f\gamma_\mu u_i$. Calculation of the finite part of this diagram in the limit of small $(-q^2)$ gives [2, 18, 33]

$$-e\bar{u}_f\gamma_\mu u_i \to -e\bar{u}_f\left\{\gamma_\mu\left[1+\frac{e^2}{3\pi}\frac{q^2}{m^2}\left(\ln\frac{m}{m_\gamma}-\frac{3}{8}\right)\right]-\left[\frac{e^2}{2\pi}\frac{1}{2m}i\sigma_{\mu\nu}q^\nu\right]\right\}u_i\,, \tag{8.32}$$

where $\sigma_{\mu\nu}=\frac{i}{2}(\gamma^\mu\gamma^\nu-\gamma^\nu\gamma^\mu)$. The expression in the first square brackets here gives the corresponding contribution to the Lamb shift, as this term is similar in form to (8.26). However, here we also meet the divergence at *small* momenta, which can be

[3] These calculations are very cumbersome, and we refer reader for details to [2, 6].

formally avoided in (8.32) by introducing a small fictitious photon mass m_γ. This divergence is related to the so-called *infrared catastrophe*. In fact, the contribution connected with the fictitious mass m_γ is exactly *cancelled* by similar terms, originating from diagrams in Figure 8.3(c). Infrared divergences in QED do not lead to major difficulties like in the case of *ultraviolet* divergences (appearing due to the divergence of Feynman integrals at the upper limit) discussed above. The infrared catastrophe is related to the ever-present possibility (for any QED process) of radiation of the large number of very "soft" photons with very small energy (frequency). Thus, the appearance of the infrared catastrophe is connected with the somehow inconsistent formulation of the problem: What is the probability for electron scattering by static nuclei with no photon being emitted? In reality we have to determine the scattering amplitude for an electron without a single photon emission, as well as the amplitudes with the emission of one, two, three etc. "soft" photons with energies less than m_γ. Each of these amplitudes diverges, but an artificial introduction of m_γ makes them finite. The sum of all these amplitudes does not diverge, and the fictitious parameter m_γ just cancels. This problem was analyzed in detail at the early stages of development of QED [2, 6, 18].

We are now interested in the second term in the square brackets in (8.32), which modifies γ_μ, i. e., the structure of the current. In fact, we can convince ourselves [2, 6, 18], that the contribution of the type $\sigma_{\mu\nu}q^\nu$ describes the magnetic moment of an electron $\boldsymbol{\mu} = -\frac{e}{2m}\boldsymbol{\sigma}$, which is usually written as $\boldsymbol{\mu} = -g\frac{e}{2m}\mathrm{s}$, with spin $\mathrm{s} = \frac{1}{2}\boldsymbol{\sigma}$, and g being the gyromagnetic ratio for an electron (in Dirac's theory $g = 2$). Accordingly, the second term in (8.32) describes the additional contribution to the magnetic moment of an electron, so that

$$\boldsymbol{\mu} = -\frac{e}{2m}\left(1 + \frac{e^2}{2\pi}\right)\boldsymbol{\sigma} \tag{8.33}$$

or

$$g = 2 + \frac{e^2}{\pi}\,. \tag{8.34}$$

Thus, in addition to Dirac's magnetic moment of an electron, there appears the so-called *anomalous* magnetic moment $e^2/2\pi$. A more precise expression for the anomalous contribution to the gyromagnetic ratio, obtained through very tedious calculations taking into account terms up to the order of $\sim e^6$, has the form

$$\frac{g-2}{2} = \frac{1}{2}\frac{e^2}{\pi} - 0.32848\left(\frac{e^2}{\pi}\right)^2 + (1.49 \pm 0.2)\left(\frac{e^2}{\pi}\right)^3 + \cdots = (1159655.4 \pm 3.3)\cdot 10^{-9}. \tag{8.35}$$

The uncertainty shown here is related to the difficulty of *calculating* the very large number of diagrams of the order of $\sim e^6$. The *experimental* value of the anomalous gyromagnetic ratio is

$$\left.\frac{g-2}{2}\right|_{exp} = (1159657.7 \pm 3.5)\cdot 10^{-9}\,. \tag{8.36}$$

This is the reason why QED is considered to be probably the most exact of the theories of interacting elementary particles. To the author's knowledge, up to now no discrepancies between QED predictions and experiments were ever found in purely electrodynamic phenomena.

The analysis of the radiation corrections using Feynman diagrams, being rigorous, is rather complicated and requires tedious calculations. To understand the physics of these effects, it is useful to refer to the qualitative approach proposed by Welton, which allows us to obtain their simple interpretation, based on the picture of the vacuum fluctuations of an electromagnetic field and the role of the electron–positron vacuum.

First of all, let us discuss the mean-square fluctuations of an electromagnetic field in the arbitrary point of a physical vacuum. Consider a field in some normalization volume V. Zero-point oscillation with frequency ω has the energy $\frac{\hbar\omega}{2}$. We can write the obvious relation

$$\frac{\hbar\omega}{2} = \frac{1}{8\pi}\overline{\int dV(\mathbf{E}_\omega^2 + \mathbf{H}_\omega^2)} = \frac{1}{4\pi}\overline{\int dV \mathbf{E}_\omega^2} = \frac{\mathbf{E}_{0\omega}^2}{8\pi}V\,, \tag{8.37}$$

where $\mathbf{E}_{0\omega}$ and $\mathbf{H}_{0\omega}$ are amplitudes of electric and magnetic fields in a vacuum, corresponding to zero-point oscillation with frequency ω, while the line denotes averaging over the oscillation period. From (8.37) we find the mean-square amplitude of a zero-point oscillation of the field, corresponding to frequency ω, as

$$\mathbf{E}_{0\omega}^2 = \frac{4\pi\hbar\omega}{V}\,. \tag{8.38}$$

Consider an electron bound within an atom. It is acted upon by a Coulomb field of nuclei and also by zero-point fluctuations of the electromagnetic field in a vacuum. Thus, the orbital motion of an electron is superposed with additional chaotic motion due to vacuum fluctuations of the electromagnetic field. Let $V(\mathbf{r})$ denote the potential energy of an electron at a point $\mathbf{r}$. We can write the electron coordinate as $\mathbf{r} = \mathbf{r}_0 + \mathbf{r}'$, where $\mathbf{r}_0$ denotes the usual coordinate, which is more or less regularly changing during its orbital motion, while $\mathbf{r}'$ is its small displacement under the influence of a random force from vacuum field fluctuations. Then we can write the change of the average potential energy of an electron under these random displacements as

$$\begin{aligned}\langle \Delta V(\mathbf{r})\rangle &= \langle V(\mathbf{r}_0 + \mathbf{r}') - V(\mathbf{r}_0)\rangle \approx \left\langle x_i' \frac{\partial V}{\partial x_i} + \frac{1}{2}(x_i' x_k')\frac{\partial^2 V}{\partial x_i \partial x_k}\right\rangle \\ &= \frac{1}{2}\nabla^2 V\langle (x_i')^2\rangle = \frac{1}{6}\nabla^2 V\langle (\mathbf{r}')^2\rangle\,.\end{aligned} \tag{8.39}$$

Here the angular brackets denote the average over all the possible values of the random variable $\mathbf{r}'$. During this averaging we take into account that $\langle x_i'\rangle = 0$, and $\langle x_i' x_k'\rangle = \frac{1}{3}\langle(\mathbf{r}')^2\rangle$ due to the spatial isotropy of these random displacements.

For the Coulomb field of the proton we have

$$\nabla^2 V(\mathbf{r}_0) = 4\pi e^2 \delta(\mathbf{r}_0)\,, \tag{8.40}$$

so that

$$\langle V(\mathbf{r})\rangle = V(\mathbf{r}_0) + \frac{2\pi e^2}{3}\delta(\mathbf{r}_0)\langle \mathbf{r}'^2\rangle\,. \tag{8.41}$$

To estimate the Lamb shift of an atomic level, we have to average (8.41) over the electron state of the atom, so that

$$\Delta E_{Lamb} = \frac{2\pi}{3} e^2 \int dV |\psi_n(\mathbf{r}_0)|^2 \delta(\mathbf{r}_0) \langle \mathbf{r}'^2 \rangle = \frac{2\pi}{3} e^2 |\psi_n(0)|^2 \langle \mathbf{r}'^2 \rangle \,, \tag{8.42}$$

where ψ_n is the wave function of the relevant atomic state.

To estimate $\langle \mathbf{r}'^2 \rangle$ we assume that electron displacement under the influence of field fluctuations is independent of its orbital motion. Let us write the classical equation of motion:

$$m \frac{d^2 \mathbf{r}'_\omega}{dt^2} = e\mathbf{E}_\omega = e\mathbf{E}_{0\omega} \sin(\mathbf{kr} - \omega t) \,, \tag{8.43}$$

which gives

$$\mathbf{r}'_\omega = -\frac{e\mathbf{E}_{0\omega}}{m\omega^2} \sin(\mathbf{kr} - \omega t) \,. \tag{8.44}$$

Accordingly,

$$\overline{\langle (\mathbf{r}'_\omega)^2 \rangle} = \frac{e^2}{2m^2\omega^4} \mathbf{E}^2_{0\omega} = \frac{2\pi e^2 \hbar}{m^2 \omega^3 V} \,, \tag{8.45}$$

where the line again denotes time averaging, and to get the last equality we used (8.38).

Zero-point oscillations with different frequencies are independent, so that their contribution to the mean square displacement of an electron can be written as a simple sum:

$$\langle \mathbf{r}'^2 \rangle = \frac{V}{\pi^2 c^3} \int d\omega \omega^2 \overline{\langle (\mathbf{r}'_\omega)^2 \rangle} = \frac{2e^2\hbar}{\pi c^3 m^2} \int_{\omega_{\min}}^{\omega_{\max}} \frac{d\omega}{\omega} \,. \tag{8.46}$$

In the absence of an electron–positron vacuum, the upper integration limit here can be arbitrarily large, and the integral diverges. In fact, for frequencies of the order of $mc^2/\hbar$ zero-point oscillations of the electromagnetic field begin to interact with the filled negative energy ("background") states of the electron–positron vacuum. We can imagine the interaction of current fluctuations due to random displacement of electrons with positive energy and similar currents due to random displacement of "background" electrons from the filled states. Due to the Pauli principle, all electrons tend to avoid each other, and these current fluctuations should be in the opposite phase, leading to their effective compensation. This leads to the effective cutoff in (8.46) for $\omega_{\max} \sim mc^2$. The cutoff at the lower limit in (8.46) is determined by some average frequency of electronic excitation in an atom, which is of the order of the Rydberg frequency: $\omega_{\min} = \omega_0 \sim \frac{\mathrm{Ry}}{\hbar} = \frac{me^4}{2\hbar^3}$. Then (8.46) reduces to

$$\langle (\mathbf{r}'^2) \rangle = \frac{2e^2\hbar}{\pi c^3 m^2} \ln \frac{mc^2}{\hbar\omega_0} = \frac{2}{\pi} \frac{e^2}{\hbar c} \left(\frac{\hbar}{mc} \right)^2 \ln \frac{mc^2}{\hbar\omega_0} \,. \tag{8.47}$$

Now we obtain for the value of Lamb shift (8.42):

$$\Delta E_{Lamb} = \frac{4}{3} \frac{e^2}{\hbar c} \left(\frac{\hbar}{mc} \right)^2 |\psi_n(0)|^2 \ln \frac{mc^2}{\hbar\omega_0} \,. \tag{8.48}$$

This shift is always positive, and the s-level of hydrogen is higher in energy than predicted by the standard Schroedinger–Dirac theory. For a hydrogen atom

$$|\psi_n(0)|^2 = \left(\frac{1}{na\pi^{1/3}} \right)^3 \,, \tag{8.49}$$

where $a = \frac{\hbar^2}{me^2}$ is the Bohr radius, and we obtain

$$\Delta E_{Lamb} = \frac{8}{3\pi}\left(\frac{e^2}{\hbar c}\right)^3 \frac{\mathrm{Ry}}{n^3} \ln \frac{mc^2}{\hbar\omega_0}. \tag{8.50}$$

More accurate and detailed calculations by Bethe produced more precise result: $\hbar\omega_0 \approx 18\,\mathrm{Ry}$. Then from (8.50) it follows the value of the Lamb shift for $2s$-state of hydrogen is $\Delta E_{Lamb}(2s) \approx 1040\,\mathrm{MHz}$, which is very close to the result of rigorous calculations based upon the general QED formalism and renormalization theory. Thus, the Lamb shift is another confirmation of the reality of the physical "vacuum" of quantum field theory.

8.3 Renormalization – how it works

Previous examples of the calculation of radiation corrections in QED demonstrated the major role of renormalization procedures, allowing us to get rid of the inevitable divergences of Feynman integrals in the higher orders of perturbation theory. The development of the theory of renormalization in QED leads to the development of practical methods for calculating the arbitrary physical effects due to electromagnetic interaction, as well as for analyzing some conceptual problems of the theory. The notion of *renormalizability* is crucial for modern quantum field theory. Models of interacting fields lacking this property are usually treated as nonphysical. Before moving to a rigorous treatment of the renormalization procedure, we shall discuss the qualitative aspects of this approach, using as an example the case of charge renormalization in a single-loop approximation.

Let us return to equation (8.27), which contains a $\ln \frac{\Lambda^2}{m^2}$ divergence. The value of the electric charge enters the theory via the diagram for an elementary vertex, shown in Figure 8.4. There is an infinite number of corrections to this vertex, with some examples shown in Figure 8.5, which actually change the value of the charge. The *physical* charge is determined by all the corrections of this type and the result of the summation of *all* diagrams for the vertex part is experimentally measured as the charge of an electron. Let us call the "initial" charge, associated with an elementary vertex of Figure 8.4, the "bare" charge e_0. Then for the "true" or "dressed" charge e we can write the perturbation expansion in powers of the "bare" charge, e. g., built upon the single-loop polarization correction, as represented by the diagrams in Figure 8.6, where dots replace the similar diagrams of higher orders. The relation between e^2 and e_0^2 can be

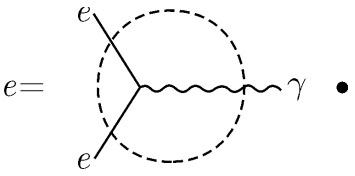


Figure 8.4

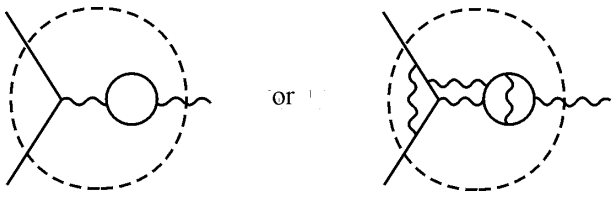


Figure 8.5

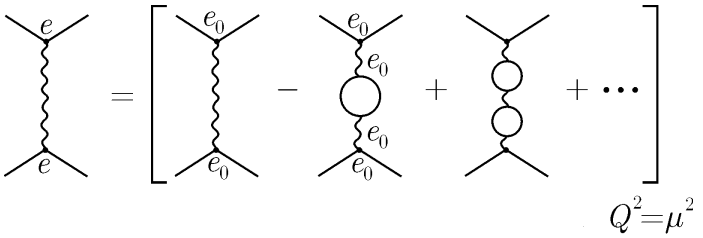


Figure 8.6

established, as shown in Figure 8.6, at some appropriate (from an experimental point of view) value of transferred (by photon line) momentum[4] $q^2 \equiv -Q^2 = -\mu^2$. In most traditional methods, to determine the charge value we use the low-energy limit of $Q^2 \ll m^2$. As a result, the expansion shown in Figure 8.6 can be schematically written as

$$e^2 = e_0^2[1 - I(Q^2 = \mu^2) + O(e_0^4)], \tag{8.51}$$

where the value of $I(Q^2)$ is determined by equations (8.17)–(8.20), i. e., by the single-loop approximation $\sim e_0^2$. Taking the square root of both sides of equation (8.51), we obtain

$$e = e_0\left[1 - \frac{1}{2}I(Q^2 = \mu^2) + O(e_0^4)\right], \tag{8.52}$$

which coincides with (8.27) after the square root expansion. Eexpansion (8.52) is shown in diagramatic form in Figure 8.7. Accordingly, taking into account all the

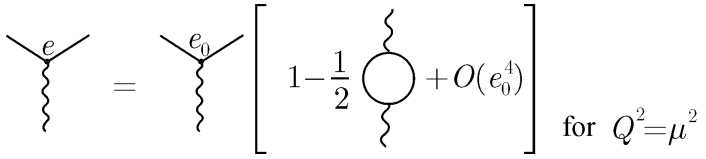


Figure 8.7

[4] The value of Q^2 is introduced here instead of $(-q^2)$, just for convenience, to deal with $Q^2 > 0$.

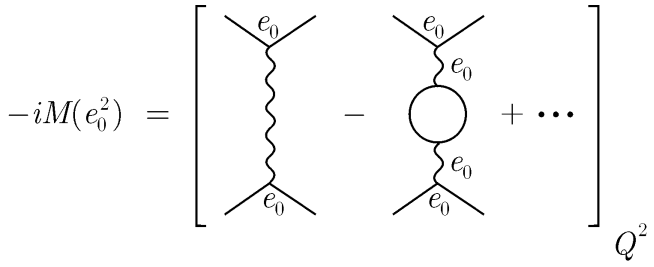


Figure 8.8

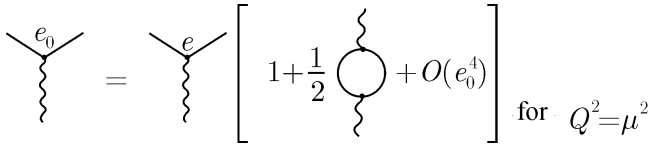


Figure 8.9

orders of perturbation theory, we have

$$e = e_0[1 + e_0^2 A_1(Q^2) + e_0^4 A_2(Q^2) + \ldots]_{Q^2=\mu^2}\,. \tag{8.53}$$

It is clear that $A_1(Q^2), A_2(Q^2), \ldots$ are infinite in the limit of $\Lambda^2 \to \infty$. Consider some physical scattering process, e. g., the one shown in the diagrams of Figure 8.8. In analytic form:

$$-i\,M(e_0^2) = e_0^2[F_1(Q^2) + e_0^2 F_2(Q^2) + O(e_0^4)]\,. \tag{8.54}$$

Here all the terms are also divergent. But now we are taking a crucial step. Let us renormalize the value of $-i\,M(e_0^2)$, expressing e_0 via e, by inversing (8.52), or, in other words, reconstructing the diagrams of Figure 8.7 with the same accuracy, as shown in Figure 8.9, and substituting this expansion into the vertices of the diagrams in Figure 8.8. Then we obtain the diagrammatic expansion shown in Figure 8.10. The first two diagrams of this expansion originate from the first diagram of Figure 8.8,

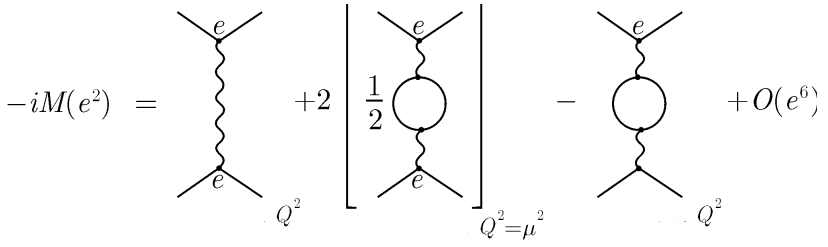


Figure 8.10

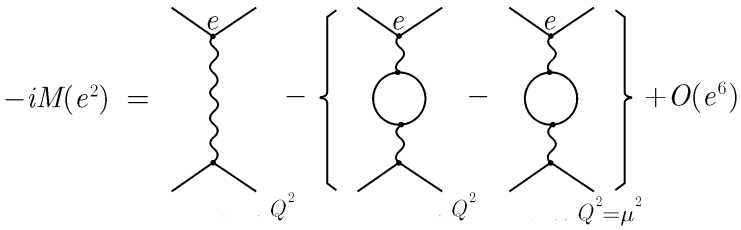


Figure 8.11

while the factor of 2 appears because we have to express e_0 via e in every vertex. In the remaining diagram of Figure 8.8 we can simply replace e_0 by e, as inaccuracy here is of the order of e^6. We can rewrite the expansion of Figure 8.10 as shown in Figure 8.11. In analytic form this expansion is written as

$$-iM(e^2) = e^2[F_1'(Q^2) + e^2 F_2'(Q^2) + O(e^4)]\,. \tag{8.55}$$

Now we have achieved everything we wanted: comparing (8.54) and (8.55) we can see that the new scattering amplitude is expressed only via the "experimental" charge e, defined according to (8.53) and measured at $Q^2 = \mu^2$. Actually, here we have not added or dropped anything, but just changed the parameters in (8.54), and in fact $M(e^2) = M(e_0^2)$. At the same time, the term $\sim e_0^4$ in (8.54) is infinite, while the term $\sim e^4$ in (8.55) is *finite*! It is clear from the fact that the "experimental" charge e is finite by definition, while two terms in the brackets in Figure 8.11 are of the opposite sign, so that after summation we obtain

$$\left[\frac{e^2}{3\pi}\ln\frac{\Lambda^2}{Q^2} - \frac{e^2}{3\pi}\ln\frac{\Lambda^2}{\mu^2}\right] = \frac{e^2}{3\pi}\ln\frac{\mu^2}{Q^2}\,, \tag{8.56}$$

which is *independent* of cutoff Λ^2. Different choices of parameter μ^2 (renormalization point) lead to different expansions (8.55). However, the observable value of $|M|^2$ should not depend on the choice of μ. This requirement can be written as the following differential equation:

$$\mu\frac{dM}{d\mu} = \left(\mu\frac{\partial}{\partial\mu} + \mu\frac{\partial e}{\partial\mu}\frac{\partial}{\partial e}\right)M = 0\,. \tag{8.57}$$

This means that the explicit dependence of M on μ which is contained in the coefficients $F_i'(Q^2, \mu^2)$ in expansion (8.55) is compensated for by the appropriate μ^2-dependence of $e^2(\mu^2)$. Equation (8.57) is a typical differential equation of the *renormalization group*, which is of great significance in quantum field theory. Below we shall once more return to a discussion of this (renormalization) invariance of the theory, which allows one to analyze conceptual foundations of quantum field theory and gives an effective formalism to perform calculations of specific effects.

8.4 "Running" the coupling constant

The expansion of Figure 8.6 can be redrawn as shown in Figure 8.12. If we limit ourselves only to loop diagrams, we obtain geometric a progression which is easily summed, as shown in Figure 8.13. We have seen above that divergences can be eliminated if we work with the physical (renormalized) charge e, which is determined by the expansion shown in Figure 8.13 at $Q^2 = \mu^2$. Actually we can use any value of μ^2. Different choices of $Q^2 = \mu_1^2, \mu_2^2, \ldots$ correspond to the perturbation expansion in powers of numerically different values of the physical charge $e(\mu_i^2)$. In fact, from Figure 8.13 we obtain

$$e^2(Q^2) = \frac{e_0^2}{1 + I(Q^2)}\,, \tag{8.58}$$

so that the experimentally observable charge depends on the value of transferred (during the scattering) momentum Q^2. The value of $e(Q^2)$ is called the "running" coupling constant. In the limits of large $Q^2 \equiv (-q^2)$ the value of $I(q^2)$ is given by (8.24), and we get

$$e^2(Q^2) = \frac{e_0^2}{1 - \frac{e_0^2}{3\pi} \ln\left(\frac{Q^2}{\Lambda^2}\right)}\,. \tag{8.59}$$

To exclude in equation (8.59) the explicit dependence of $e^2(Q^2)$ on the cutoff parameter Λ, we consider this expression at $Q^2 = \mu^2$ and express e_0 via $e^2(\mu^2)$. As a result,

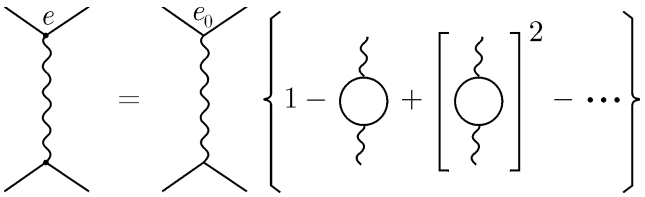


Figure 8.12

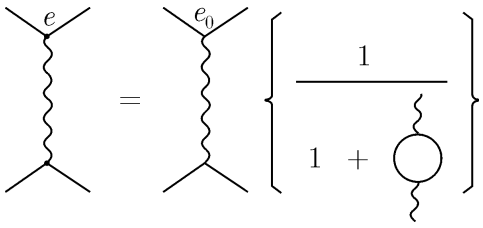


Figure 8.13

for large Q^2 we can rewrite (8.59) as

$$e^2(Q^2) = \frac{e^2(\mu^2)}{1 - \frac{e^2(\mu^2)}{3\pi} \ln\left(\frac{Q^2}{\mu^2}\right)}. \tag{8.60}$$

Here everything is finite! The "running" coupling constant $e(Q^2)$ describes the dependence of the effective charge on the transferred momentum Q^2, i. e., in fact on the distance between the charged particles. We shall see later that it is really the *observable* effect, and the corresponding dependence is precisely logarithmic. However, the result expressed by equation (8.60) raises a number of conceptual questions on the consistency of QED. The thing is that from (8.60) we can see that with the growth of Q^2 (reduction of distance) the value of effective charge *grows*, so that sooner or later perturbation theory becomes invalid at small distances, and for

$$Q^2 = \mu^2 \exp\left(\frac{3\pi}{e^2(\mu^2)}\right) \tag{8.61}$$

we obtain an obviously unphysical divergence ("ghost" pole). For Q^2 larger than this value the charge becomes imaginary! For historical reasons this behavior is called "Moscow zero" (or the "zero-charge" problem). In the following we shall return several times to the discussion of this situation and related problems.

Actually, somewhat prematurely, we note that in quantum chromodynamics (QCD) situation is just the opposite. There we also obtain the "running" coupling constant of gluons and quarks, which is expressed (similarly to (8.60)) as

$$g^2(Q^2) = \frac{g^2(\mu)}{1 + \frac{g^2(\mu^2)}{12\pi}(33 - 2n_f) \ln\left(\frac{Q^2}{\mu^2}\right)}, \tag{8.62}$$

where n_f is the number of *flavors* of quarks, while the constant factor of 33 is connected to the non-Abelian nature of gauge symmetry in QCD (in fact it is calculated as some constant, related to the properties of the matrices of generators of the color group $SU(3)$). Only for the world with $n_f > 16$, the sign in the denominator of equation (8.62) will be the same as in QED. In the real world we have $n_f = 6$. Thus, the effective charge in QCD does not grow, but *diminishes* with the growth of Q^2 and becomes small at small distances. This behavior is called "asymptotic freedom". For small enough Q^2 (at large distances between quarks) the effective coupling constant (in contrast to QED!) becomes large, which is directly related to the *confinement* of quarks ("infrared prison"). Let us denote the value of Q^2, corresponding to the pole ("ghost pole" again!) in (8.62), as Λ^2, so that

$$\Lambda^2 = \mu^2 \exp\left[-\frac{12\pi}{(33 - 2n_f)g^2(\mu^2)}\right]. \tag{8.63}$$

Then (8.62) can be rewritten as

$$g^2(Q^2) = \frac{12\pi}{(33 - 2n_f) \ln\left(\frac{Q^2}{\Lambda^2}\right)}. \tag{8.64}$$

For $Q^2 \gg \Lambda^2$ the effective coupling constant is small, and the interaction of quarks and gluons (at small distances of large momenta) can be described by perturbation theory, similarly to interactions of electrons and photons in QED (at large distances or small momenta). For $Q^2 \sim \Lambda^2$ such a description becomes impossible, while quarks and gluons form strongly coupled clusters: the hadrons. The experimental value of Λ is somewhere in the interval between 0.1 and 0.5 GeV. Then, for experiments at $Q^2 \sim (30\ \text{GeV})^2$ it follows from (8.64) that $g^2 \sim 0.1$, so that perturbation theory is valid, like in QED. In the limits of large Q^2 we can neglect all quark masses, but there is still a mass scale in the theory, given by μ^2, which appears in the process of renormalization.

8.5 Annihilation of e^+e^- into hadrons. Proof of the existence of quarks

As an interesting illustration of QED applications, let us show how purely electrodynamic experiments prove the existence of quarks [24]. This becomes possible via studies of the high-energy annihilation processes of electrons and positrons, with arbitrary hadrons in the final state. In fact, these reactions are going through the creation of quark–antiquark pairs, i. e., $e^+e^- \to q\bar{q}$, which afterwards form hadrons. We can show that the crossection for such processes can be obtained from an easily calculable QED cross section for electron–positron annihilation into muons: $e^+e^- \to \mu\bar{\mu}$.

To calculate the cross section of this process in QED it is sufficient to consider the second of the Feynman diagrams shown in Figure 6.6, where the final products of the reaction is the pair $\mu\bar{\mu}$[5]. The standard calculation, using the rules of QED diagram technique, gives the total cross section for such process as [24]

$$\sigma(e^+e^- \to \mu\bar{\mu}) = \frac{4\pi e^2}{3Q^2}\,, \tag{8.65}$$

where $Q^2 = 4E^2$ is the square of energy in the center of the mass reference frame (Mandelstam variable s). Then, the cross section for annihilation into the quark–antiquark pair is given by

$$\sigma(e^+e^- \to q\bar{q}) = 3e_q^2\sigma(e^+e^- \to \mu\bar{\mu})\,, \tag{8.66}$$

where e_q is the q-quark charge. An additional factor of 3 appears here due to three separate diagrams for different quark colors, which are to be summed, so that appropriate cross sections are also to be summed. To find the cross section for the creation of all possible hadrons it is necessary to sum over all quark flavors $q = u, d, s, \ldots$, so that

$$\sigma(e^+e^- \to \textit{hadrons}) = \sum_q \sigma(e^+e^- \to q\bar{q}) = 3\sum_q e_q^2\sigma(e^+e^- \to \mu\bar{\mu})\,. \tag{8.67}$$

[5] We remind, that muons are just like electrons, but with larger (about 200 times) rest mass.

Thus we obtain the very important prediction

$$R \equiv \frac{\sigma(e^+e^- \to hadrons)}{\sigma(e^+e^- \to \mu\bar{\mu})} = 3\sum_q e_q^2 \,. \tag{8.68}$$

As cross section $\sigma(e^+e^- \to \mu\bar{\mu})$ is well studied (and is in excellent agreement with equation (8.65)), the experimental measurements of the cross section for e^+e^--annihilation into hadrons give direct information on the number of quarks, their flavors, and their colors. We have

$$R = \begin{cases} 3\left[\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2\right] = 2 & \text{for } u,d,s\,. \\ 2 + 3\left(\frac{2}{3}\right)^2 = \frac{10}{3} & \text{for } u,d,s,c\,, \\ \frac{10}{3} + 3\left(\frac{1}{3}\right)^2 = \frac{11}{3} & \text{for } u,d,s,c,b \quad \text{etc.} \end{cases} \tag{8.69}$$

These predictions have been well confirmed by experiments! The value of $R = 2$ is observed for $Q < 2(m_c + m_u) \approx 3.7$ GeV, i. e., below the threshold for creation of c-quarks. Above the threshold for creation of five quark flavors, i. e., for $Q > 2m_b \approx 10$ GeV, the experimentally observed value of this ratio is $R = 11/3$. These experiments directly confirm the existence of three colors of quarks with the appropriate (fractional!) values of the electric charge.

Within QCD we can also take into account the contributions of diagrams with quarks (or antiquarks) emitting gluons [24]. In the first order over g^2, equation (8.68) is modified as follows:

$$R = 3\sum_q e_q^2 \left(1 + \frac{g^2(Q^2)}{\pi}\right), \tag{8.70}$$

so that the weak (logarithmic) dependence of R on Q^2 is also observed.

8.6 The physical conditions for renormalization

Let us now discuss more rigorously the basics of renormalizability in QED. It is clear that the general scheme of invariant perturbation theory and diagrammatic equations for exact propagators presented above was rather formal. We have operated with all the entities of the theory as with the usual finite mathematical expressions, though explicit calculations of $\mathcal{D}, G$, and Γ, using perturbation theory, inevitably produce diverging integrals. We shall explicitly show below that using certain recipes in QED allows us to perform the well-defined "subtraction" of all infinities and to obtain finite expressions for all the measurable physical characteristics. These recipes are based upon the obvious physical requirements of the photon mass being exactly zero, while the electron charge and mass are equal to their observable values. Our presentation will

necessarily be a bit schematic, but further details can be found in [6] and in especially detailed analysis in [2].

A physical photon has zero mass, so that its dispersion is given by $k^2 = 0$. This means that the exact photon propagator should always have a pole at $k^2 = 0$, so that

$$\mathcal{D}(k^2) = \frac{4\pi}{k^2} Z \quad \text{for } k^2 \to 0\,, \tag{8.71}$$

where Z is some constant. According to equation (7.39), the general form of the propagator is expressed via the polarization operator as

$$\mathcal{D}(k^2) = \frac{4\pi}{k^2(1 - \mathcal{P}(k^2)/k^2)}\,, \tag{8.72}$$

so that from (8.71) we get for polarization operator

$$\mathcal{P}(0) = 0\,. \tag{8.73}$$

Similarly, the constant Z in equation (8.71) can be defined as

$$\frac{1}{Z} = 1 - \left.\frac{\mathcal{P}(k^2)}{k^2}\right|_{k^2 \to 0}\,. \tag{8.74}$$

Further limitations on the behavior of $\mathcal{P}(k^2)$ can be obtained from the analysis of the physical definition of electric charge. Two classical (very heavy!) particles being at rest at some large distance from each other ($r \gg m^{-1}$, where m is electron mass), are interacting according to the Coulomb law: $V(r) = e^2/r$. On the other hand, this interaction is expressed by the diagram shown in Figure 8.14, where the "fat" wavy line denotes the exact propagator of the virtual photon and the upper and lower lines correspond to classical particles. Self-energy corrections for the photon are taken into account in its exact propagator. Any other self-energy corrections, acting upon the lines of heavy particles, lead to the corresponding diagrams being zero. In fact, an addition of some internal line into the diagram of Figure 8.14, e. g., joining 1 and 3 or 1 and 2 by a photon line, leads to the appearance in corresponding diagrams of heavy virtual particles (due to particle lines under the extra photon lines), with the propagators containing large mass M of the classical particle in denominators, giving zero contribution in the limit of $M \to \infty$. Then it is clear that the facto of $e^2\mathcal{D}(k^2)$ in

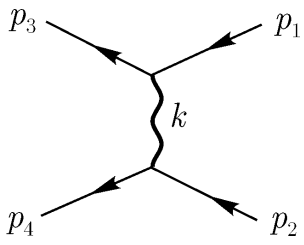


Figure 8.14

the diagram of Figure 8.14 is given (up to a sign) by the Fourier transform of the interaction potential of our particles. The static nature of interaction (particles at rest!) corresponds to the frequency of virtual photon $\omega = 0$, while the large distances correspond to the small wave vectors $\mathbf{k}$. As $\mathcal{D}$ depends only on $k^2 = \omega^2 - \mathbf{k}^2$, we arrive at the condition

$$e^2\mathcal{D} \to \frac{4\pi e^2}{k^2} \quad \text{for } k^2 \to 0\,, \tag{8.75}$$

so that in equation (8.71) we have to put $Z = 1$. Then from equation (8.74) it immediately follows that

$$\frac{\mathcal{P}(k^2)}{k^2} \to 0 \quad \text{for } k^2 \to 0\,. \tag{8.76}$$

Besides the previously derived condition (8.73) it follows now that

$$\mathcal{P}'(0) \equiv \left.\frac{d\mathcal{P}(k^2)}{dk^2}\right|_{k^2=0} = 0\,. \tag{8.77}$$

Note that the effective external line of a real photon should be associated with the factor of $\sqrt{4\pi}[1 + \frac{1}{4\pi}\mathcal{P}(k^2)\mathcal{D}(k^2)]e_\mu$. However, for the real photon we always have $k^2 = 0$, and due to (8.76) we conclude that in the lines of external photons we can safely drop all radiation corrections.

Thus, the natural physical requirements lead to definite values (zeroes!) for $\mathcal{P}(0)$ and $\mathcal{P}'(0)$. At the same time, direct calculation using the diagram rules of perturbation theory leads here to diverging integrals. We can get rid of these divergencies if we attribute the finite values dictated by physical requirements to these divergent expressions. This is the main idea of renormalization. Another way to formulate this operation e. g., for charge renormalization, is as follows. We can introduce the nonphysical "bare" charge e_0, as a parameter entering the initial expression for an operator of electromagnetic interaction, which is used in formal perturbation theory. After that, the renormalization condition is formulated as the requirement of

$$e_0^2\mathcal{D}(k^2) \to \frac{4\pi e^2}{k^2} \quad \text{for } k^2 \to 0\,, \tag{8.78}$$

where e is the true physical charge of a particle. The we find the relation

$$e^2 = Ze_0^2 \tag{8.79}$$

Now the unphysical e_0 is excluded from all expressions, determining the physical effects (while divergence is "hidden" in the renormalization factor Z). If we require $Z = 1$ from the very beginning, we actually perform renormalization "on the fly" [6], so that there is no need to introduce any fictitious entities into the intermediate calculations.

Let us now consider the physical conditions for renormalization of the electron propagator. It is obvious that the exact propagator $\mathcal{G}(p)$ should have a pole at $p^2 = m^2$,

where m is the mass of physical electron. Then we can write

$$\mathcal{G}(p) \approx Z_1 \frac{\gamma^\mu p_\mu + m}{p^2 - m^2 + i0} + g(p) \quad \text{for } p^2 \to m^2 , \tag{8.80}$$

where Z_1 is a scalar constant (renormalization factor) and $g(p)$ is finite for $p^2 \to m^2$. From (8.80) we immediately obtain the inverse propagator as

$$\mathcal{G}^{-1}(p) \approx \frac{1}{Z_1}(\gamma^\mu p_\mu - m) - (\gamma^\mu p_\mu - m) g(p) (\gamma^\mu p_\mu - m) \quad \text{for } p^2 \to m^2 . \tag{8.81}$$

The mass operator for $p^2 \to m^2$ now has the form

$$\mathcal{M}(p) = G^{-1}(p) - \mathcal{G}^{-1}(p) \approx \left(1 - \frac{1}{Z_1}\right)(\gamma^\mu p_\mu - m) + (\gamma^\mu p_\mu - m) g(p) (\gamma^\mu p_\mu - m) . \tag{8.82}$$

We associate the following factor with the effective external electron line (e. g., incoming) in the scattering diagram:

$$\mathcal{U}(p) = u(p) + \mathcal{G}(p)\mathcal{M}(p)u(p) , \tag{8.83}$$

where $u(p)$ is the usual electron bispinor, satisfying the Dirac equation $(\gamma^\mu p_\mu - m)u = 0$. Due to relativistic invariance ($\mathcal{U}$ is also the bispinor) the limiting value of $\mathcal{U}$(p) for $p^2 \to m^2$ can differ from $u(p)$ by a constant scalar factor (wave function renormalization)

$$\mathcal{U}(p) = Z' u(p) . \tag{8.84}$$

It is not difficult to show [6] the validity of a simple relation

$$Z' = \sqrt{Z_1} . \tag{8.85}$$

This is almost obvious, since Green's function (propagator) is quadratic in electron operators.

Now, after the establishment of the limiting behavior of the electron propagator, there is no need of any additional conditions for the vertex operators. Consider the diagram in Figure 8.15 and let us assume that it describes first-order electron scattering by an external field $A_\mu^{(e)}(k)$, taking into account all radiation corrections. In the limit of $k \to 0$ we have $p_2 \to p_1 \equiv p$ and the radiation corrections to the line of the

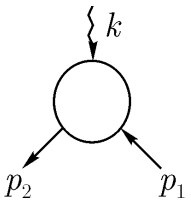


Figure 8.15

external field vanish (we already noted above that they vanish for arbitrary $k^2 = 0$). Then this diagram corresponds to the amplitude

$$M_{fi} = -e\bar{\mathcal{U}}(p)\Gamma^{\mu}(p, p; 0)\mathcal{U}(p)A_{\mu}^{(e)}(k \to 0)\,. \tag{8.86}$$

But for $k \to 0$ the potential $A_{\mu}^{(e)}(k)$ reduces to a constant independent of coordinates and time, which does not describe any physical field and cannot lead to any change of transition current. In other words, in this limit, transition current $\bar{\mathcal{U}}\Gamma^{\mu}\mathcal{U}$ should simply coincide with free current $\bar{u}\gamma^{\mu}u$:

$$\bar{\mathcal{U}}(p)\Gamma^{\mu}(p, p; 0)\mathcal{U}(p) = Z_1\bar{u}(p)\Gamma^{\mu}u(p) = \bar{u}(p)\gamma^{\mu}u(p)\,. \tag{8.87}$$

This relation is automatically satisfied due to Ward identity, independent of the value of Z_1. In fact, substituting $\mathcal{G}^{-1}(p)$ from (8.81) into (7.87), we obtain

$$\Gamma^{\mu}(p, p; 0) = \frac{1}{Z_1}\gamma^{\mu} - \gamma^{\mu}g(p)(\gamma^{\mu}p_{\mu} - m) - (\gamma^{\mu}p_{\mu} - m)g(p)\gamma^{\mu} \tag{8.88}$$

and (8.87) is satisfied due to $(\gamma^{\mu}p_{\mu} - m)u(p) = 0$ and $\bar{u}(p)(\gamma^{\mu}p_{\mu} - m) = 0$. This again simply gives us the definition of the physical electron charge. We see that the renormalization factor Z_1 just drops from the amplitude of the physical process. We can simply require

$$\bar{u}(p)\Gamma^{\mu}(p, p; 0)u(p) = \bar{u}(p)\gamma^{\mu}u(p) \qquad \text{for} \quad p^2 = m^2\,, \tag{8.89}$$

i. e., put $Z_1 = 1$. The convenience of such a definition is that now there is no necessity to introduce any corrections to external electron lines, and we simply have $\mathcal{U}(p) = u(p)$. It is also clear also that, for $Z_1 = 1$, for mass operator (8.82) we have

$$\mathcal{M}(p) = (\gamma^{\mu}p_{\mu} - m)g(p)(\gamma^{\mu}p_{\mu} - m)\,, \tag{8.90}$$

so that the second term in (8.83) obviously reduces to zero. Thus, there is no need to renormalize the external lines of all real particles, both photons and electrons.

8.7 The classification and elimination of divergences

The physical conditions of renormalization introduced above allow us, in principle. to obtain the finite and definite values for the amplitudes of any QED process in an arbitrary order of the perturbation theory.

Consider first the character of divergences appearing in different Feynman integrals. First of all we calculate the powers of the virtual 4-momenta, entering the integrand. Consider an arbitrary diagram of the n-th order (n is the number of vertices!), containing external lines of N_e electron and N_γ photon. The number N_e is always even. The total number of electron lines is equal to $2n$; of these N_e are external and I_e are

internal. During the calculation of the number of lines, the internal lines are counted twice, as each of them connects two vertices, so that

$$2n = N_e + 2I_e \,. \tag{8.91}$$

Correspondingly, the total number of internal electron lines in the diagram is

$$I_e = n - \frac{N_e}{2} \,. \tag{8.92}$$

Each vertex is connected with one photon line, and for N_γ vertices this line is external, while for the remaining $n - N_\gamma$ vertices this line is internal. As each internal photon line connects two vertices, the total number of these lines is equal to

$$\frac{n - N_\gamma}{2} \,. \tag{8.93}$$

Each internal photon line is associated with the propagator $D(k)$, which contains k to the power of -2. Each internal electron line is associated with the propagator $G(p)$, which behaves like p to power of -1 (for $p^2 \gg m^2$). Thus, the total power of the 4-momenta in the denominator of the integrand is

$$2\frac{n - N_\gamma}{2} + n - \frac{N_e}{2} = 2n - \frac{N_e}{2} - N_\gamma \,. \tag{8.94}$$

The number of integrations over d^4p and d^4k in the diagram is equal to the number of internal lines, but the conservation law of the 4-momentum in each vertex leads to an additional $n - 1$ constraint on integration momenta (one of these n conservation laws is connected with external momenta, and it corresponds to the general conservation law for the scattering process described by this diagram). Correspondingly, taking into account equations (8.92) and (8.93), we conclude that the total number of internal lines (both electron and photon) in the diagram is given by

$$n - \frac{N_e}{2} + \frac{n}{2} - \frac{N_\gamma}{2} = \frac{3}{2}n - \frac{N_e}{2} - \frac{N_\gamma}{2} \,, \tag{8.95}$$

which gives the number of integrations, not taking conservation laws into account. Then, subtracting $n - 1$ we obtain for the number of independent integration momenta

$$\frac{3}{2}n - \frac{N_e}{2} - \frac{N_\gamma}{2} - n + 1 = \frac{n}{2} + 1 - \frac{N_e}{2} - \frac{N_\gamma}{2} \,. \tag{8.96}$$

Multiplication by 4 gives the total number of integrations:

$$2(n - N_e - N_\gamma + 2) \,. \tag{8.97}$$

The difference between the number of integrations and the power of the momenta in the denominator of the integrand for our diagram is equal to the difference between (8.97) and (8.94):

$$r = 4 - \frac{3}{2}N_e - N_\gamma \,. \tag{8.98}$$

This number determines whether the appropriate Feynman integral is convergent or divergent[6]. Strictly speaking, the situation is more complicated, as the condition of $r < 0$ for the diagram as a whole is not sufficient to guarantee its convergence. We also have to require the negative values of r' for all internal blocks which can be contained within our diagram. The presence of internal blocks with $r' > 0$ leads to the divergence of the diagram as a whole, though all other integrals may be convergent. The condition $r < 0$ is sufficient to guarantee the convergence of the simplest diagrams.

For $r \geq 0$ the integral is always divergent. The power of divergence is not less than r if r is even, and not less than $r - 1$ if r is odd (the drop of divergence power by 1 in the last case is related to the integration of the product of odd number of 4-vectors over the whole 4-space giving zero!). The power of divergence can grow due to the presence of internal blocks with $r' > 0$.

Note that the divergence power of the diagram r, according to (8.98), does not depend on the diagram order n. This remarkable property, as we shall see later, makes the theory *renormalizable*. Briefly speaking, the important thing here is that from (8.98) it becomes immediately clear that only the finite number of *types* of divergence exists in such a theory, because with positivity of both N_e and N_γ we can obtain $r \geq 0$ for only a few pairs of the values of these integers, and thus only the finite number of the simplest *primitively diverging* diagrams. Correspondingly, we can introduce the finite number of parameters (to be determined from the experiments) to "hide" all divergences. In the case of n entering (with positive sign) into (8.98), the number of divergence types will grow with the growth of n, so that situation will become hopeless! For QED we can explicitly list all primitively diverging diagrams. From the very beginning we can exclude the cases of $N_e = N_\gamma = 0$ (vacuum loops) and $N_e = 0, N_\gamma = 1$ (the average value of vacuum current). All other cases are shown in Figure 8.16. For the first of these diagrams we have $r = 2$, and divergence is formally quadratic; in all other cases $r = 0$ or $r = 1$, and divergence is logarithmic.

The diagram of Figure 8.16(d) represents the first correction to the vertex. It should satisfy (8.89), which can be written as

$$\bar{u}(p)\Lambda^\mu(p, p; 0)u(p) = 0 \qquad \text{for} \quad p^2 = m^2\,, \tag{8.99}$$

where

$$\Lambda^\mu = \Gamma^\mu - \gamma^\mu\,. \tag{8.100}$$

Let us denote the Feynman integral, written according to diagram rules, as $\bar{\Lambda}^\mu(p_2, p_1; k)$. This integral is logarithmically divergent and does not satisfy (8.99). However, we can obtain the expression satisfying this condition by constructing the difference

$$\Lambda^\mu(p_2, p_2; k) = \bar{\Lambda}^\mu(p_2, p_1; k) - \bar{\Lambda}^\mu(p_1, p_1; 0)|_{p_1^2 = m^2}\,. \tag{8.101}$$

[6] Let us recall that in all cases we are dealing with divergences of integrals at the upper integration limit!

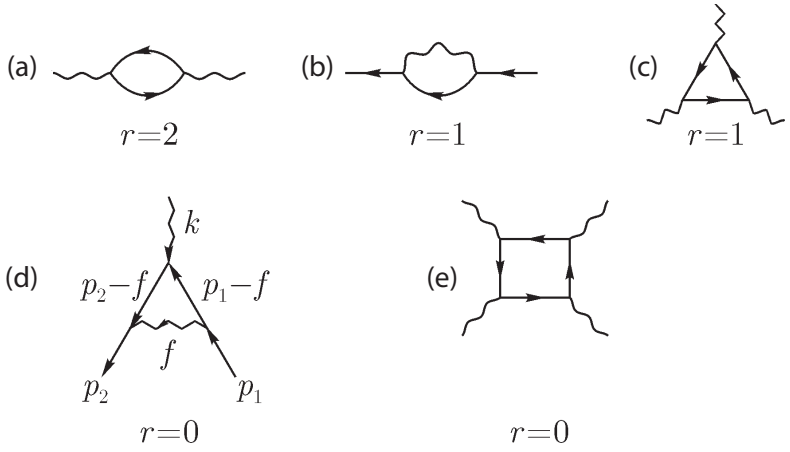


Figure 8.16

Divergence in the integral for $\bar{\Lambda}^{\mu}(p_2, p_1; k)$ can be separated if we consider the limit of a very large 4-momentum of the virtual photon f. Then we obtain

$$-4\pi i e^2 \int \frac{d^4 f}{(2\pi)^4} \gamma^{\nu} G(p_2 - f) \gamma^{\mu} G(p_1 - f) \gamma^{\lambda} D_{\lambda\nu}(f) \sim$$
$$-4\pi i e^2 \int \frac{d^4 f}{(2\pi)^4} \frac{\gamma^{\nu}(\gamma^{\kappa} f_{\kappa}) \gamma^{\mu} (\gamma_{\rho} f_{\rho}) \gamma_{\nu}}{f^2 f^2 f^2}, \quad (8.102)$$

which is independent of the values of the 4-momentum of the external lines. Then in the difference given by (8.101) divergences are cancelled, and we obtain the finite expression.

Such a procedure for cancelling divergence is called the subtraction scheme of renormalization. Let us stress that the possibility for cancelling divergence in $\bar{\Lambda}^{\mu}(p_2, p_1; k)$ by only one subtraction is guaranteed by (the weakest possible) logarithmic nature of the divergence.

After determining the first correction for Γ^{μ} (i. e., the first term of expansion for Λ^{μ}), the first correction for the electron propagator (diagram of Figure 8.16(b)) can be calculated using the Ward identity (7.87), which can be rewritten as

$$-\frac{\partial \mathcal{M}(p)}{\partial p_{\mu}} = \Lambda^{\mu}(p, p; 0), \quad (8.103)$$

introducing the mass operator $\mathcal{M}$ instead of $\mathcal{G}$ and Λ^{μ} instead of Γ^{μ}. This equation can be easily integrated with the boundary condition

$$\bar{u}(p) \mathcal{M}(p) u(p) = 0 \quad \text{for} \quad p^2 = m^2, \quad (8.104)$$

which follows from (8.90).

In principle, in a similar (though more tedious) way we can cancel divergences from the polarization operator of Figure 8.16(a) [2, 6], but here we have to make two subtractions:

$$\mathcal{P}(k^2) = \bar{\mathcal{P}}(k^2) - \bar{\mathcal{P}}(0) - k^2\bar{\mathcal{P}}'(0)\,, \tag{8.105}$$

where $\bar{\mathcal{P}}$ denotes the Feynman integral, corresponding to this diagram. It is obvious that (8.105) satisfies the physical requirements given by (8.73) and (8.77).

The next order of the perturbation theory for the vertex operator $\Lambda_\mu^{(2)}$ is determined by the diagrams shown in Figure 7.17(c–i). Of these, only the diagrams shown in Figure 7.17(d–f) are compact, which can be made finite with the help of one subtraction (8.101). Internal self-energy and vertex parts, contained within the noncompact diagrams, can be directly replaced by the already-known (renormalized) values of the first order, given by $\mathcal{P}^{(1)}$, $\mathcal{M}^{(1)}$ and $\Lambda_\mu^{(1)}$, so that integrals are again made finite by subtraction (8.101). Corrections $\mathcal{M}^{(2)}$ and $\mathcal{P}^{(2)}$ are then calculated using the Ward identity (8.103) and (8.105). The systematic application of such procedure gives, in principle, the rigorous way to obtain finite expressions for $\mathcal{P}$, $\mathcal{M}$, and Λ_μ in an arbitrary order of the perturbation theory [2, 6]. This makes possible the calculation of the amplitudes of the physical scattering processes, containing blocks like $\mathcal{P}$, $\mathcal{M}$, and Λ_μ. The physical conditions of renormalizability formulated above are sufficient to cancel divergences from all Feynman integrals. This is the manifestation of the quite nontrivial property of *renormalizability* of QED. Below we shall return several times to the discussion of renormalizability and its use in other models of quantum field theory.

8.8 The asymptotic behavior of a photon propagator at large momenta

Let us consider the conceptually very important problem of photon propagator asymptotic behavior at large momenta $|k^2| \gg m^2$. In the lowest order of perturbation theory the polarization operator is determined by the simple loop diagram shown in Figure 8.17. It is defined by the Feynman integral

$$\frac{i}{4\pi}\mathcal{P}^{\mu\nu}(k) = -e^2 \int \frac{d^4p}{(2\pi)^4} Sp\gamma^\mu G(p)\gamma^\nu G(p-k)\,. \tag{8.106}$$

However, this integral (over the whole 4-dimensional p-space) diverges (quadratically, according to a simple power counting of the previous section, but actually only

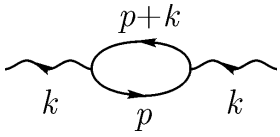


Figure 8.17

logarithmically due to a "hidden" algebra of the integrand). These divergence can be cancelled using the renormalization recipes of the previous sections. Direct calculations are rather tedious [2]. This analysis is much simplified in the asymptotic limit of $|k^2| \gg m^2$, which is of major interest to us. As we shall see below, after the renormalization 8.106), in this limit we have

$$\mathcal{P}(k^2) = \frac{e^2}{3\pi} k^2 \ln \frac{|k|^2}{m^2} \,. \tag{8.107}$$

In essence, this gives the first order correction inverse photon propagator $4\pi D^{-1} = k^2$ and it is valid until the following condition is satisfied:

$$\frac{e^2}{3\pi} \ln \frac{|k^2|}{m^2} \ll 1 \,, \tag{8.108}$$

which limits the validity region of our approximation at high values of $|k|^2$. In fact, equation (8.107) can be used even under the much weaker condition of

$$\frac{e^2}{3\pi} \ln \frac{|k^2|}{m^2} \lesssim 1 \,. \tag{8.109}$$

Now we shall give a proof of this statement and also obtain the result (8.107) itself [6]. First of all, let us note that, although for (8.109) there may be additional contributions to $\mathcal{P}(k^2)$ due to the higher orders of perturbation theory, in the n-th order it is sufficient to take into account only the terms of the order of $\sim (e^2)^n \ln^n(\frac{|k|^2}{m^2})$, containing the large logarithm, appearing in the limit of $|k^2| \gg m^2$. This logarithm should enter with the same power as e^2, because terms with lower powers of the logarithm are obviously smaller due to $e^2 \ll 1$. This is called the approximation of leading logarithms.

Consider now the Dyson equation for polarization operator (7.77)

$$\mathcal{P}(k^2) = \frac{4\pi i e^2}{3} Sp \int \frac{d^4 p}{(2\pi)^4} \gamma^\mu \mathcal{G}(p+k) \Gamma_\mu(p+k, p; k) \mathcal{G}(p) \,. \tag{8.110}$$

As we have shown above, $\mathcal{P}(k^2)$ is gauge invariant, so that calculating it using Feynman diagrams we can use any gauge for the propagators and vertices. Most convenient is the Landau gauge, when the photon propagator is written as ($D^l = 0$):

$$\mathcal{D}_{\mu\nu}(k) = \frac{4\pi}{k^2} \left(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \right) . \tag{8.111}$$

A detailed analysis of the correction diagrams for (8.106), which can be found in [6], shows that in this gauge perturbation theory series does not contain terms with the required powers of logarithms at all.

Then in (8.110) it is sufficient to use the zero-th order approximations $\mathcal{G} = G$ and $\Gamma^\mu = \gamma^\mu$. Then (8.110) reduces to the integral

$$\mathcal{P}(k^2) = \frac{4\pi i e^2}{3} Sp \int \frac{d^4 p}{(2\pi)^4} \gamma^\mu G(p+k) \gamma_\mu G(p) \,, \tag{8.112}$$

the same as in (8.106). Let us discuss the appearance of the logarithm in this integral. It is easily seen that it originates from the integration region

$$p^2 \gg |k^2| \quad \text{for } |k^2| \gg m^2 . \tag{8.113}$$

In fact, in this limit we can write[7]

$$G(p) \approx \frac{1}{\gamma^\mu p_\mu} = \frac{\gamma^\mu p_\mu}{p^2} , \tag{8.114}$$

$$\begin{aligned} G(p-k) \approx! & \frac{1}{\gamma^\mu p_\mu - \gamma^\mu k_\mu} = \frac{1}{\gamma^\mu p_\mu} + \frac{1}{\gamma^\mu p_\mu}\gamma^\nu k_\nu \frac{1}{\gamma^\alpha p_\alpha} \\ & + \frac{1}{\gamma^\mu p_\mu}\gamma^\nu k_\nu \frac{1}{\gamma^\alpha p_\alpha}\gamma^\rho k_\rho \frac{1}{\gamma^\beta p_\beta} + \cdots \\ = & \frac{\gamma^\mu p_\mu}{p^2} + \frac{(\gamma^\mu p_\mu)(\gamma^\nu k_\nu)(\gamma^\alpha p_\alpha)}{(p^2)^2} \\ & + \frac{(\gamma^\mu p_\mu)(\gamma^\nu k_\nu)(\gamma^\alpha p_\alpha)(\gamma^\rho k_\rho)(\gamma^\beta p_\beta)}{(p^2)^3} + \cdots . \end{aligned} \tag{8.115}$$

After substitution of these expressions into (8.112) the first term, independent of k, drops out due to renormalization in accordance with the condition $\mathcal{P}(0) = 0$ (the first subtraction in (8.105)). The second term also becomes zero after integration over the directions of p. The third integral is logarithmically divergent over p^2; it can be easily estimated making the integration from $p^2 \sim |k^2|$ (lower limit of the region (8.113)) up to some "cutoff parameter" Λ^2:

$$\int d^4 p \frac{p^4}{p^8} \sim \int dp\, p^3 \frac{p^4}{p^8} \sim \int_{|k^2|}^{\Lambda^2} dp^2\, p^2 \frac{p^4}{p^8} \sim \int_{|k^2|}^{\Lambda^2} dp^2 \frac{1}{p^2} \sim \ln \frac{\Lambda^2}{|k^2|} . \tag{8.116}$$

Finally we get

$$\mathcal{P}(k^2) = -\frac{e^2}{3\pi} k^2 \ln \frac{\Lambda^2}{|k^2|} . \tag{8.117}$$

This is not the end of our derivation; for the final cancellation of the divergence (at $\Lambda \to \infty$) we need to subtract from $\mathcal{P}(k^2)/k^2$ its value at $k^2 = 0$ (second subtraction in (8.105)). However, logarithmic accuracy of our calculations assumes $|k^2| \gg m^2$, so that it is sufficient to subtract the value of (8.117) at $|k^2| \sim m^2$, and Λ^2 in the argument of logarithm is simply replaced by m^2. Thus, we obtain the required result (8.107). In the Landau gauge there are no corrections to $\mathcal{G}$ and Γ with the "proper" powers of the logarithm, and equation (8.107) is actually valid under the condition (8.109).

The function $\mathcal{D}(k^2)$, corresponding to the polarization operator (8.107), has the form

$$\mathcal{D}(k^2) = \frac{4\pi}{k^2} \frac{1}{1 - \frac{e^2}{3\pi} \ln \frac{|k^2|}{m^2}} . \tag{8.118}$$

[7] The signs here are determined by the properties of the γ-matrices.

Because of (8.109) there is no need to expand this expression in the powers e^2. However, the validity of (8.118) is limited from the side of large $|k^2|$ due to the diminishing denominator. The derivation of (8.118) was based on logarithmic approximation and neglect of infinite sequences of the diagrams of higher orders, which do not contain leading logarithms. According to (8.118) an addition of each new "fat" photon line introduces an additional factor of $e^2\mathcal{D}$, and the small parameter of the perturbation theory instead of e^2 is given by

$$\frac{e^2}{1-\frac{e^2}{3\pi}\ln\frac{|k^2|}{m^2}} \ll 1\,, \tag{8.119}$$

which coincides with the "running" coupling constant (8.60) discussed above. As $|k^2|$ grows, this coupling becomes of the order of unity, so that the small expansion parameter actually disappears, and perturbation theory can not be further applied.

8.9 Relation between the "bare" and "true" charges

The situation with (8.118), (8.119) can be understood more clearly if, during the derivation of (8.118), we do not do renormalization "on the fly", but introduce the first "bare" charge e_0, which afterwards is fitted to obtain the correct observable value of charge e (or e_R in the notations used above). If the logarithmically divergent integral is cut off at the upper limit at some Λ^2 (above we have also used the notation M^2 as the cutoff parameter), the "bare" charge can be considered to be its function: $e_0 = e_0(\Lambda^2)$, and at the end we have to perform the limit $\Lambda \to \infty$. In this approach, the polarization operator takes the form (8.117)

$$\mathcal{P}(k^2) = -\frac{e_0^2}{3\pi}k^2\ln\frac{\Lambda^2}{|k^2|}\,. \tag{8.120}$$

Correspondingly,

$$\mathcal{D}(k^2) = \frac{4\pi}{k^2}\frac{1}{1+\frac{e_0^2}{3\pi}\ln\frac{\Lambda^2}{|k^2|}}\,. \tag{8.121}$$

Let us define the physical charge e according to

$$e_0^2\mathcal{D}(k^2) \to \frac{4\pi e^2}{k^2} \quad \text{where } k^2 \to m^2\,, \tag{8.122}$$

i. e., at distances of the order of m^{-1} (i. e., the Compton length of an electron $\hbar/mc$, which in quantum field theory defines its effective size). Then we obtain

$$e^2 = \frac{e_0^2}{1+\frac{e_0^2}{3\pi}\ln\frac{\Lambda^2}{m^2}}\,, \tag{8.123}$$

which, in fact, coincides with (8.59) if we choose the normalization point $\mu^2 = m^2$. Then

$$e_0^2 = \frac{e^2}{1 - \frac{e^2}{3\pi} \ln \frac{\Lambda^2}{m^2}} \,. \tag{8.124}$$

If we formally consider the limit of point-like charge $\Lambda \to \infty$ in (8.123), we shall obtain $e \to 0$, independently of the possible form of the function $e_0^2(\Lambda)$. This behavior is called the "zero charge" (or "Moscow zero"). It was first noted by Landau and Pomeranchuk and independently by Fradkin in the mid-1950s. In the opinion of Landau [49] this situation reflected the internal inconsistency of both the renormalization procedure and QED (and any other model of quantum field theory known at that time) itself.

Let us consider the arguments of Landau and Pomeranchuk, which lead them to this rather radical conclusion. Let the ratio $\frac{\Lambda^2}{|k^2|}$ be so large that

$$\frac{e_0^2}{3\pi} \ln \frac{\Lambda^2}{|k^2|} \gg 1 \,, \tag{8.125}$$

but, at the same time, we still have $e_0 \ll 1$. Then in equation (8.121) we can neglect unity in the denominator, so that

$$\mathcal{D}(k^2) = \frac{12\pi^2}{k^2 e_0^2 \ln \frac{\Lambda^2}{|k^2|}} \tag{8.126}$$

and, correspondingly, from (8.122) we have

$$e^2 = \frac{3\pi}{\nu \ln \frac{\Lambda^2}{m^2}} \,, \tag{8.127}$$

which is *independent* of the value of "bare" charge e_0. Note that here we divided (8.127) by an additional parameter ν, which denotes the number of the fundamental fermions which contribute to vacuum polarization (the corresponding contributions to polarization loops are additive!). Let us now introduce, instead of the standard 4-potential of electromagnetic field A^μ, a new 4-vector $\mathcal{A}^\mu = e_0 A^\mu$. Then the interaction Hamiltonian H_I will not contain the "bare" charge e_0, while the free electromagnetic field Hamiltonian H_0 (quadratic in A^μ) will contain e_0^2 in the denominator. The function $\tilde{\mathcal{D}}(k^2)$, defined with the help of $\mathcal{A}^\mu$ in the same way that $\mathcal{D}(k^2)$ is defined via A^μ, will be equal to

$$\tilde{\mathcal{D}}(k^2) = e_0^2 \mathcal{D}(k^2) = \frac{12\pi^2}{k^2 \ln \frac{\Lambda^2}{|k^2|}} \,. \tag{8.128}$$

This expression does not contain e_0, and this means that it corresponds to the neglect in the total Hamiltonian $H = H_0 + H_I$ (depending of e_0) of the term H_0. If this neglect of H_0 in comparison to H_I is possible (at large Λ) already for $e_0^2 \ll 1$, it is natural to assume, that it is even more justified at not so small e_0^2. Then equation (8.126), and also equation (8.127), become unrelated to the condition of $e_0^2 \ll 1$, so that the limit of $\Lambda \to \infty$ becomes feasible. Then $e^2 \to 0$ independent of the form of the function $e_0^2(\Lambda)$.

The cutoff parameter Λ, guaranteeing the validity of (8.127), is in any case very large. At corresponding (very small!) distances the effects of gravitation may exceed those of electromagnetism. This leads to the very attractive idea that the "crisis" of QED happens precisely

at those distances (energies) where gravitation coupling matches that of electromagnetism. Choosing the cutoff parameter of the order of the Planck length, we have

$$G_N \Lambda^2 \sim 1\,, \tag{8.129}$$

where G_N is the Newtonian gravitational constant. If we accept such a point of view, the value of the physical charge e will be automatically determined by the theory via equations (8.127) and (8.129), which will lead to the limitation of $\nu \approx 12$. In fact, if $\nu < 12$, the effects of gravitation will become important well before the effective charge becomes of the order of unity. In the opposite case of $\nu > 12$ the effects of gravitation will not "save" electrodynamics, becoming important "too late". Note that, according to the modern experimental data on elementary particles (see Chapter 1), there are precisely 12 fundamental fermions!

At the same time, we must stress the opinion of the majority of theorists, who believe that the limit of $\Lambda \to \infty$ in expressions like (8.123) and (8.124) cannot be performed without breaking the assumptions made during their derivation. From (8.124) it is seen, that as Λ grows (with fixed e^2) the value of e_0^2 also grows and for $e_0^2 \sim 1$ all these expressions just become invalid, as their derivation was based on the assumption of $e_0^2 \ll 1$, which is simply the criterion of the applicability of perturbation theory.

Note that for QED all these difficulties are of a rather "academic" importance, since they appear at fantastically high energies of no real interest: $\frac{e^2}{\pi} \ln(\frac{E^2}{m^2}) = 1$ is achieved for $E \sim 10^{93} m$, which is essentially due to the smallness of $e^2 = \frac{1}{137}$. Much earlier, as we shall see later, electromagnetic interactions become "intermixed" with the weak and strong interactions of the elementary particles, so that "pure" electrodynamics looses its meaning. At the end of this volume we shall return to the discussion of the problems of consistency of quantum field theory and its asymptotic behavior.

Fora better understanding of these problems we now present a simple qualitative discussion on coordinate space [5]. We can transform our asymptotic expressions of QED to "coordinate representation" by an obvious (from dimensionality arguments) replacement: $m \to r^{-1}$ and $\Lambda \to r_0^{-1}$, where r is the characteristic distance from the "center" of an electron (which can be taken to be of the order of its Compton wavelength), while r_0 is some fundamental length, characterizing the geometric size of the "bare" charge, which can be imagined to be a small sphere with radius r_0. Then (8.123) can be written as

$$e^2(r) = \frac{e^2(r_0)}{1 + \frac{2e^2(r_0)}{3\pi^2} \ln \frac{r}{r_0}}\,. \tag{8.130}$$

Let the value of the "bare" charge $e^2(r_0)$ be fixed. Our aim now is to go to the limit of a point-like "bare" charge, so we start to diminish r_0, with a fixed value of $e_0^2(r_0)$. Then, sooner or later we obtain $\frac{2e^2(r_0)}{3\pi^2} \ln \frac{r}{r_0} \gg 1$, and we can neglect unity in the denominator of (8.130). Correspondingly, we have

$$e^2(r) = \frac{1}{\frac{2}{3\pi^2} \ln \frac{r}{r_0}}\,. \tag{8.131}$$

But now, with a further diminishing of r_0 we obtain $e^2(r) \to 0$ for $r_0 \to 0$. This is "zero charge" behavior (or "Moscow zero"). On this ground, Landau and Pomeranchuk claimed [8]: "We are coming to the fundamental conclusion that the formal quantum electrodynamics apparently leads to the zero charge of an electron. The word "apparently" here is related to some lack of rigorousness in the above arguments". The physics here is that in this approximation vacuum polarization (due to creation of virtual electron–positron pairs) is so strong at small distances that on some distance the remnant charge is actually independent of the initial ("bare") charge. In the limit of the point-like "bare" charge nothing remains of it on any finite distance: we have the complete screening. Note that this result is quite transparent; the phenomenon of screening is well known in plasma and solid state physics [36], where it is described by quite similar calculations of the polarization operator in many body systems [1]. But in this situation, how we can understand the magnificent successes of QED?

Let us write (8.130) in the form solved with respect to $e^2(r_0)$ and put $r = \lambda_e = m^{-1}$ (the Compton wave length of an electron):

$$e^2(r_0) = \frac{e^2(\lambda_e)}{1 - \frac{2e^2(\lambda_e)}{3\pi^2} \ln \frac{\lambda_e}{r_0}} \,. \tag{8.132}$$

Here $e^2(\lambda_e)$ should be understood as the "physical" charge of an electron, i. e., that charge which is measured at large distances (of the order of λ_e) outside the effective region of vacuum polarization (screening). When we "enter" this region ($r_0 < \lambda_e$), the charge grows due to the diminishing screening inside the "cloud" of electron–positron pairs[9]. However, we cannot reach the limit of a very large charge due to the existence of the "Landau ghost pole", close to which equation (8.132) simply becomes invalid. From a practical point of view all this is not important at all, as we are speaking here about the region of $r_0 \sim \lambda_e \exp[-(137)(3\pi^2)/2]$. QED is a practical theory precisely because we are using nonexact solutions with point-like interaction and leave open the question of the correct behavior at small distances, where other interactions become quite important. And who knows; is there not some physical mechanism cutting off divergences at small distances (e. g., related to gravitation; cf. (8.129)!), which is still unknown to us, but which makes interactions in quantum field theory effectively nonpoint-like? Thus, the pragmatic (majority!) point-of-view is that we are dealing with the experimentally defined "physical" charge $e(\lambda_e)$, such that we can work with solutions in the form of a perturbation series, though modern theory becomes invalid at small distances. Thus, the problem of the asymptotic behavior of QED still remains unsolved (we shall return to this at the end of the book).

In asymptotically free field theories, e. g., in QCD, the situation is different. The sign in the denominators of expressions similar to (8.130) and (8.132) is opposite to that in QED, and the asymptotic behavior of the interaction "constant" (charge) is also opposite: $g^2(r_0) \to 0$ for $r_0 \to 0$. It is not that the charge at the finite distance is becoming zero for an arbitrary value of the "bare" charge, but the zero point-like charge corresponds to a finite charge at the finite distance: $g^2(r)$ grows with the growth of r. In QCD we are dealing with an effective "antiscreening" of the "bare" charge. However, we do not know which values of r_0 and $g(r_0)$ should be fixed and up to what values of r we can use an logarithmic expression like (8.130). We cannot use it infinitely, as the diminishing (with growth of r) denominator again makes

[8] Reports of the USSR Academy of Sciences **102**, 489 (1955)

[9] In fact, this corresponds to the dependence of the "fine structure constant" on momentum, transferred during the scattering process. We already noted that this effect is experimentally observable!

perturbation theory inapplicable. Qualitatively it is clear that this growth of the charge with r corresponds to the confinement force (acting upon quarks). These dependences are now measured experimentally, and we shall return later to the discussion of asymptotic freedom in QCD .

8.10 The renormalization group in QED

Let us show now how equations (8.130) and (8.124) can be derived using simple analysis, based on dimensional analysis and the notion of renormalizability which constitute the essence of the so-called renormalization group introduced in QED by Gell-Mann and Low. Consider again the square of the "bare" charge as a function of cutoff parameter $e_0^2(\Lambda)$ and introduce some function d which relates e_0^2 for two different values of its argument (cutoff):

$$e_0^2(\Lambda_2^2) = e_0^2(\Lambda_1^2)d. \tag{8.133}$$

For $\Lambda_1^2, \Lambda_2^2 \gg m^2$ the function d does not depend on m and, being dimensionless, can depend only on the dimensionless arguments $e_0^2(\Lambda_1^2)$ and Λ_2^2/Λ_1^2, so that we can write

$$e_0^2(\Lambda_2^2) = e_0^2(\Lambda_1^2)d\left(e_0^2(\Lambda_1^2), \frac{\Lambda_2^2}{\Lambda_1^2}\right). \tag{8.134}$$

This is the main relation of the renormalization group. Its physical meaning is quite clear: in renormalizable theory any change of the cutoff parameter can be compensated for by the appropriate change of the "bare" charge, with no change in the physical results (in this case, of the physical charge!). The functional equation (8.134) can be conveniently rewritten in a differential form. Consider the infinitesimally close values of cutoff parameters Λ_1^2 and Λ_2^2. Let us denote $\Lambda_1^2 = \xi$ and $\Lambda_2^2 = \xi + d\xi$. Then, from (8.134) we obtain

$$\begin{aligned} e_0^2(\xi + d\xi) &= e_0^2(\xi) + de_0^2(\xi) = e_0^2(\xi)d\left(e_0^2(\xi), 1 + \frac{d\xi}{\xi}\right) \\ &= e_0^2(\xi)\left[d(e_0^2(\xi), 1) + \left.\frac{\partial d(e_0^2(\xi), x)}{\partial x}\right|_{x=1}\frac{d\xi}{\xi}\right], \end{aligned} \tag{8.135}$$

which gives, on account of $d(e_0^2(\xi), 1) = 1$,

$$de_0^2(\xi) = e_0^2(\xi)\left.\frac{\partial d(e_0^2(\xi), x)}{\partial x}\right|_{x=1}\frac{d\xi}{\xi}, \tag{8.136}$$

giving the differential Gell-Mann–Low equation

$$\frac{de_0^2}{d\ln\xi} = \psi(e_0^2), \tag{8.137}$$

where we have introduced the Gell-Mann–Low function

$$\psi(e_0^2) = e_0^2\left[\frac{\partial d(e_0^2, x)}{\partial x}\right]_{x=1}. \tag{8.138}$$

Writing (8.137) as $\frac{de_0^2}{\psi(e_0^2)} = \frac{d\xi}{\xi}$ and integrating it from $\xi = \Lambda_1^2$ to $\xi = \Lambda_2^2$, we get

$$\ln \frac{\Lambda_2^2}{\Lambda_1^2} = \int_{e_0^2(\Lambda_1^2)}^{e_0^2(\Lambda_2^2)} \frac{de^2}{\psi(e^2)} . \tag{8.139}$$

If in the entire integration region the value of e_0^2 is small, we can use for $\psi(e^2)$ the expression obtained from the first order of perturbation theory. From the general expression $\mathcal{D}(k^2) = \frac{4\pi}{k^2}(1 - \frac{\mathcal{P}(k^2)}{k^2})^{-1}$ it is clear that corrections to the "bare" charge e_0^2 are given by $e_0^2 k^{-2}\mathcal{P}(k^2)$. Then, using for the polarization operator its lowest order expression (8.120), we find (cf. (8.123))

$$e^2 = e_0^2(\Lambda_1^2)\left[1 - \frac{e_0^2(\Lambda_1^2)}{3\pi} \ln \frac{\Lambda_1^2}{|k^2|}\right] = e_0^2(\Lambda_2^2)\left[1 - \frac{e_0^2(\Lambda_2^2)}{3\pi} \ln \frac{\Lambda_2^2}{|k^2|}\right] . \tag{8.140}$$

Then

$$d\left(e_0^2, \frac{\Lambda_2^2}{\Lambda_1^2}\right) = \frac{1 - \frac{e_0^2}{3\pi} \ln \frac{\Lambda_1^2}{|k^2|}}{1 - \frac{e_0^2}{3\pi} \ln \frac{\Lambda_2^2}{|k^2|}} \approx 1 + \frac{e_0^2}{3\pi} \ln \frac{\Lambda_2^2}{\Lambda_1^2} . \tag{8.141}$$

Correspondingly, using the definition (8.138), we obtain

$$\psi(e_0^2) = \frac{e_0^4}{3\pi} , \tag{8.142}$$

so that the Gell-Mann–Low function is quadratically growing with its argument. Now we can perform integration in (8.139) explicitly:

$$\frac{1}{3\pi} \ln \frac{\Lambda_2^2}{\Lambda_1^2} = \frac{1}{e_0^2(\Lambda_1^2)} - \frac{1}{e_0^2(\Lambda_2^2)} . \tag{8.143}$$

If we define the physical charge as $e^2 = \lim_{\Lambda_1^2 \to m^2} e_0^2(\Lambda_1^2)$, the expression (8.143) reduces to (8.123) and (8.124). Thus, the calculation of the Gell-Mann–Low function in the lowest order of perturbation theory and the subsequent integration of the differential renormalization group equation give the result obtained above by summation of the leading logarithms of diagrammatic expansion. In this sense, we can "overcome" the problems of the rigorous justification of this summation procedure. Sometimes it is said that the renormalization group provides an "improved" perturbation theory, where the role of the coupling constant is played by (8.119). However, all the main questions discussed above actually remain. The result (8.143) was obtained from an approximate expression for the Gell-Mann–Low function (8.142), which is valid only for $e_0^2 \ll 1$. It is not clear how it is changed by higher order corrections, and no reliable analysis of this problem is available. However, later we shall see that the qualitative analysis of the possible consequences of a differential equation like (8.137), based on certain assumptions on the form of the Gell-Mann–Low function for arbitrary values of its argument, is actually possible and may be quite useful for the general discussion on the asymptotic properties of quantum field theory.

8.11 The asymptotic nature of a perturbation series

The scheme of renormalization discussed above allows the total cancellation of divergences in separate diagrams, i. e., in separate terms of the perturbation expansion of the scattering matrix in powers of an electron charge, but not in the scattering matrix as a whole. The question arises of whether or not this renormalized perturbation series is convergent. There is an argument, due to Dyson, which proves that this series is actually divergent and belongs to the class of the so-called *asymptotic* expansions.

We have seen that the interaction between two electrons is determined by the function $e_R^2 \mathcal{D}(k^2)$, where e_R is the renormalized (physical) charge. Calculating, with the help of this function, some physical property $F(p; e_R^2)$, we obtain the infinite series in powers of e_R^2

$$F(p; e_R^2) = \sum_{n=0}^{\infty} e_R^{2n} f_n(p)\,, \tag{8.144}$$

where $f_n(p)$ are some functions of the 4-momenta of the particles. Assume that this series (with separate terms renormalized according to the procedure discussed above) is convergent for some value of e_R. Then $F(p; e_R^2) \equiv F(e_R^2)$ is an analytic function of e_R^2 for $e_R^2 \sim 0$, so that $F(-e_R^2)$ is also an analytic function, expressible as a power series. But $F(-e_R^2)$ represents our property F for the case of particle interaction given by $-e_R^2 \mathcal{D}(k^2)$, which corresponds to particle attraction rather than repulsion.

It can be easily seen that in this case the usual definition of the vacuum does not correspond to the state with the lowest possible energy! In fact, imagine the creation of N electron–positron pairs with all electrons being concentrated in one region of space and while all positrons in another region. If both regions are small and well separated, for large enough N the negative Coulomb energy of these attracting regions will become larger than their rest and kinetic energies. Let us call these states "pathological".

Assuming that charge interaction is determined by $-e_R^2 \mathcal{D}(k^2)$, consider some usual state characterized by the presence of several particles. In particular, this may be the usual vacuum state (state with no particles). This state is separated from the "pathological" state with the same energy by some energy barrier, and the height of this barrier is determined by the minimal energy needed to create N pairs, i. e., by the rest energy of these N particles.

Due to quantum mechanical tunneling there is a finite probability of transition from the usual to the "pathological" state. This means that every physical state is actually unstable towards the spontaneous creation of a large number of particles. The "pathological" state, to which our system tunnels, will not be stationary, because more and more particles will be created, so that the vacuum state in particular will be destroyed, and there will be no ground state for our system at all! Due to such "pathology", we cannot assume that QED interaction $-e_R^2 \mathcal{D}(k^2)$ leads to well-defined analytic functions. Actually, the function $F(-e_R^2)$ cannot be analytic and the perturbation series (8.144) can not be convergent for $e_R^2 \neq 0$.

Again we can pose the natural questions: What is the physical meaning of perturbation series (8.144), and why is QED, operating with such expansions, so successful in explaining experiments? The answer is that expansion (8.144) represents an asymptotic series. Such expansions, under certain conditions, can be used to describe the functions they represent with high (but always finite!) accuracy [71]. In contrast to a convergent series, the terms of the asymptotic series $e_R^{2n} f_n(p)$ first diminish with the growth of n, but then, starting from some number n_0, start to grow (and this growth is in general unlimited). The maximal accuracy for an asymptotic series to the approximate function F is determined by the value of f_{n_0}. The less this term is, the higher is this accuracy. In the case of QED, there are reasons to believe that in the series (8.144) the values of f_n will diminish up to n of the order of $n_0 \approx \hbar c/e_R^2 = 137$. This value of n_0 is so large that the accuracy of the QED series (8.144) in describing reality is very high. Apparently, the error here can be estimated as $\exp(-\hbar c/e_R^2)$, which is immensely small. For the practical tasks of QED, such accuracy is overwhelming!

Chapter 9

Path integrals and quantum mechanics

9.1 Quantum mechanics and path integrals

It is well known that quantum mechanics was initially formulated in two equivalent forms: matrix Heisenberg mechanics and wave mechanics, based on the Schroedinger equation. Later Feynman proposed [20] another quite elegant path integral formulation of quantum mechanics, which will be briefly described in this chapter. Of course, all these formulations of are equivalent and may be used to solve different practical problems, choosing those more convenient for the problem at hand. Conceptually, they stress different aspects of the same universal quantum theory and allow different ways of generalizing towards the appropriate quantum field theory. Feynman's formulation is especially convenient for this kind of generalization, as we shall see later.

Let $\psi(q_i, t_i)$ be a wave function of a quantum particle at the initial moment of time t_i, where q_i denotes the appropriate coordinate dependence. For simplicity we shall consider here only one-dimensional motion. We have seen in Chapter 4 that the value of the wave function at a later moment of time t can be written as

$$\psi(q_f, t_f) = \int dq_i K(q_f t_f; q_i t_i)\psi(q_i t_i)\,, \tag{9.1}$$

where $K(q_f t_f; q_i t_i)$ is the appropriate propagator (Green's function of the Schroedinger equation). According to standard interpretation, $\psi(q_f, t_f)$ represents the probability amplitude for finding the particle at spatial point q_f at time moment t_f. Correspondingly, propagator $K(q_f t_f; q_i t_i)$ represents the probability amplitude of particle transition from the initial point q_i at moment t_i to the final point q_f at moment t_f. The probability of this transition is given by

$$P(q_f t_f; q_i t_i) = |K(q_f t_f; q_i t_i)|^2\,. \tag{9.2}$$

Let us divide the time interval between moments t_i and t_f into two intervals separated by the time moment t. Repeated use of (9.1) gives

$$\psi(q_f, t_f) = \int dq_i \int dq K(q_f t_f; qt) K(qt; q_i t_i)\psi(q_i t_i)\,, \tag{9.3}$$

so that

$$K(q_f t_f; q_i t_i) = \int dq K(q_f t_f; qt) K(qt; q_i t_i)\,. \tag{9.4}$$

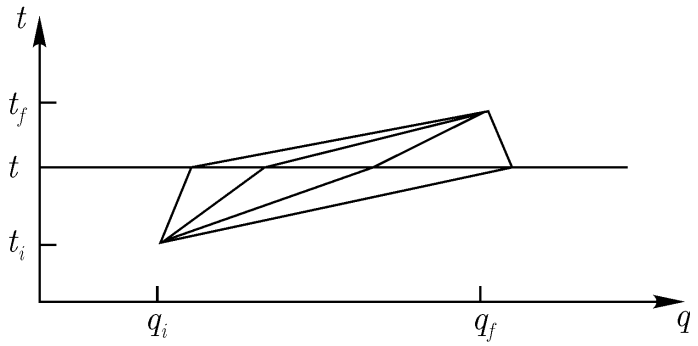


Figure 9.1

Thus, the $q_i t_i \rightarrow q_f t_f$ transition can be considered as the particle transition via all possible intermediate points (states), as shown in Figure 9.1. As an example, we may recall the notorious experiment on two slot electron diffraction. This is schematically shown in Figure 9.2, where slots are placed at points 2A and 2B. In this case the analogue of equation (9.4) can be written as

$$K(3;1) = K(3;2A)K(2A;1) + K(3;2B)K(2B;1)\,. \tag{9.5}$$

The intensity distribution at the screen, placed at point 3, is determined by

$$P(3;1) = |K(3;1)|^2\,, \tag{9.6}$$

with obvious interference contributions. It can be said that in this experiment the electron simultaneously moves along both paths (trajectories). Registering it somehow at one of the slots destroys the interference picture.

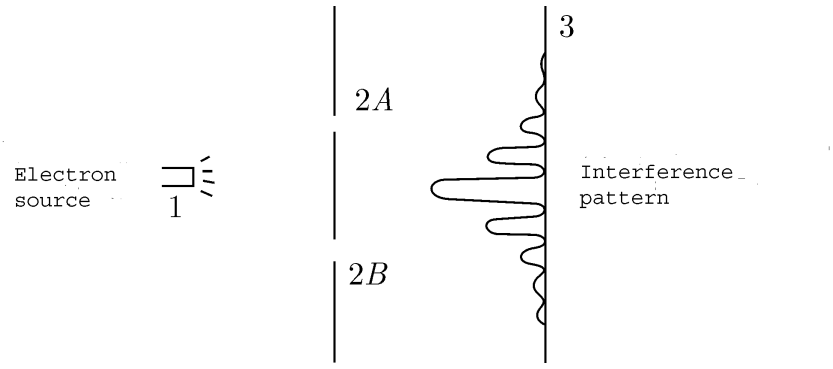


Figure 9.2

Let us introduce eigenvectors of the coordinate operator in Dirac's notations:

$$\hat{q}|q\rangle = q|q\rangle\,. \tag{9.7}$$

Then the wave function of our particle can be written as

$$\psi(qt) = \langle q|\psi_t\rangle_S\,, \tag{9.8}$$

where $|\psi_t\rangle_S$ is the state vector in the Schroedinger representation, related to the time-independent state vector in Heisenberg representation $|\psi\rangle_H$ by

$$|\psi_t\rangle_S = e^{-iHt/\hbar}|\psi\rangle_H\,. \tag{9.9}$$

Let us define the time-dependent state vector as

$$|qt\rangle = e^{iHt/\hbar}|q\rangle\,. \tag{9.10}$$

Then we can rewrite (9.8) as

$$\psi(qt) = \langle qt|\psi\rangle_H\,. \tag{9.11}$$

All these relations are well known from elementary quantum mechanics. Using the completeness of the set of state vectors (9.7), (9.10) we can write

$$\langle q_f t_f|\psi\rangle_H = \int dq_i \langle q_f t_f|q_i t_i\rangle\langle q_i t_i|\psi\rangle_H\,, \tag{9.12}$$

which reduces (with the account of (9.11)) to

$$\psi(q_f t_f) = \int dq_i \langle q_f t_f|q_i t_i\rangle\psi(q_i t_i)\,. \tag{9.13}$$

Comparing (9.13) with (9.1) we see that the propagator can be written as

$$K(q_f t_f; q_i t_i) = \langle q_f t_f|q_i t_i\rangle\,, \tag{9.14}$$

which (in a slightly different form) we already used in Chapter 4. Below we shall widely use (9.14) in our discussion.

Let us divide the time interval between moments t_i and t_f into $(n+1)$ equal segments of duration τ. Then the propagation of particle from $q_i t_i$ to $q_f t_f$ can be considered as shown in Figure 9.3, which by repeated use of (9.4), allows us to write the transition amplitude (propagator) as

$$\langle q_f t_f|q_i t_i\rangle = \int\cdots\int dq_1 dq_2\,\cdots dq_n \langle q_f t_f|q_n t_n\rangle\langle q_n t_n|q_{n-1} t_{n-1}\rangle\cdots\langle q_1 t_1|q_i t_i\rangle\,, \tag{9.15}$$

where the multiple integral is taken over *all* possible trajectories, connecting initial point q_i with final q_f. In the limit of $n\to\infty$ or $\tau\to 0$ equation (9.15) determines the propagator as a Feynman's *path integral* (continual or functional integral). Already at

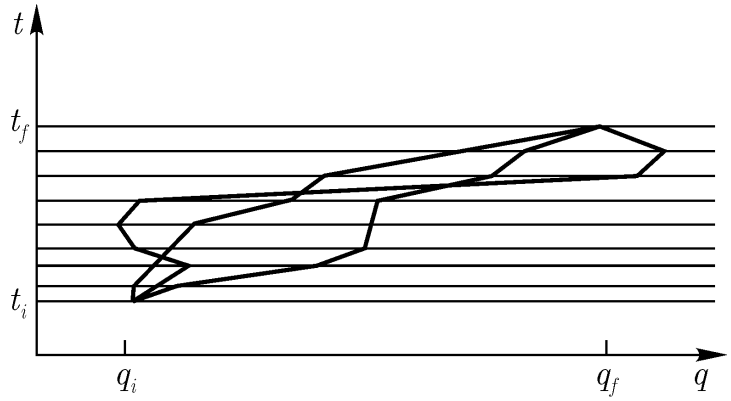


Figure 9.3

this level we can see the major difference between classical and quantum mechanics. The classical particle propagates from some initial point to the final point, moving along the *single* trajectory determined by the least action principle, while in quantum mechanics the particle motion involves the whole continuum of all possible trajectories, connecting these points!

The propagator at a small trajectory segment is easily calculated. From equation (9.10) we get

$$\begin{aligned}\langle q_{j+1}t_{j+1}|q_jt_j\rangle &= \langle q_{j+1}|e^{-iH\tau/\hbar}|q_j\rangle = \left\langle q_{j+1}\left|1-\frac{i}{\hbar}H\tau+O(\tau^2)\right|q_j\right\rangle \\ &= \delta(q_{j+1}-q_j)-\frac{i\tau}{\hbar}\langle q_{j+1}|H|q_j\rangle \\ &= \int\frac{dp}{2\pi\hbar}\exp\left[\frac{i}{\hbar}p(q_{j+1}-q_j)\right]-\frac{i\tau}{\hbar}\langle q_{j+1}|H|q_j\rangle\,, \qquad (9.16)\end{aligned}$$

where we have used an obvious representation of the δ-function via a Fourier integral. In the general case the Hamiltonian H is some function of q and p. Consider the most common case of a particle moving in a potential field, when

$$H=\frac{p^2}{2m}+V(q)\,. \qquad (9.17)$$

Then the kinetic energy term can be rewritten as

$$\left\langle q_{j+1}\left|\frac{p^2}{2m}\right|q_j\right\rangle = \int dp'\int dp\langle q_{j+1}|p'\rangle\left\langle p'\left|\frac{p^2}{2m}\right|p\right\rangle\langle p|q_j\rangle\,, \qquad (9.18)$$

so that using

$$\langle q_{j+1}|p'\rangle=\frac{1}{\sqrt{2\pi\hbar}}\exp\left(\frac{ip'q_{j+1}}{\hbar}\right),\quad \langle p|q_j\rangle=\frac{1}{\sqrt{2\pi\hbar}}\exp\left(-\frac{ipq_j}{\hbar}\right),$$

we obtain

$$\left\langle q_{j+1}\left|\frac{p^2}{2m}\right|q_j\right\rangle = \iint \frac{dpdp'}{2\pi\hbar}\exp\left[\frac{i}{\hbar}(p'q_{j+1}-pq_j)\right]\frac{p^2}{2m}\delta(p-p') = \int \frac{dp}{2\pi\hbar}\exp\left[\frac{i}{\hbar}p(q_{j+1}-q_j)\right]\frac{p^2}{2m}. \tag{9.19}$$

Note that in the left-hand side of this expression p is represented by operator, while in the right-hand side it is just a c-number! In a similar way we can obtain

$$\langle q_{j+1}|V(q)|q_j\rangle = V\left(\frac{q_{j+1}+q_j}{2}\right)\langle q_{j+1}|q_j\rangle = V\left(\frac{q_{j+1}+q_j}{2}\right)\delta(q_{j+1}-q_j) = \int \frac{dp}{2\pi\hbar}\exp\left[\frac{i}{\hbar}p(q_{j+1}-q_j)\right]V(\bar{q}_j), \tag{9.20}$$

where $\bar{q}_j = \frac{1}{2}(q_{j+1}+q_j)$. Now, from (9.19) and (9.20), we get

$$\langle q_{j+1}|H|q_j\rangle = \int \frac{dp}{2\pi\hbar}\exp\left[\frac{i}{\hbar}p(q_{j+1}-q_j)\right]H(p,\bar{q}), \tag{9.21}$$

so that (9.16) is rewritten as

$$\langle q_{j+1}t_{j+1}|q_jt_j\rangle = \int \frac{dp_j}{2\pi\hbar}\exp\left\{\frac{i}{\hbar}\left[p_j(q_{j+1}-q_j)-\tau H(p_j,\bar{q}_j)\right]\right\}, \tag{9.22}$$

where p_j is momentum at the moment between t_j and t_{j+1} (between q_j and q_{j+1}). The corresponding segments of the trajectory in momentum space are shown in Figure 9.4. Equation (9.22) defines the propagator on a small segment of one of the paths (trajectories). The complete propagator is obtained by substitution of (9.22) into (9.15),

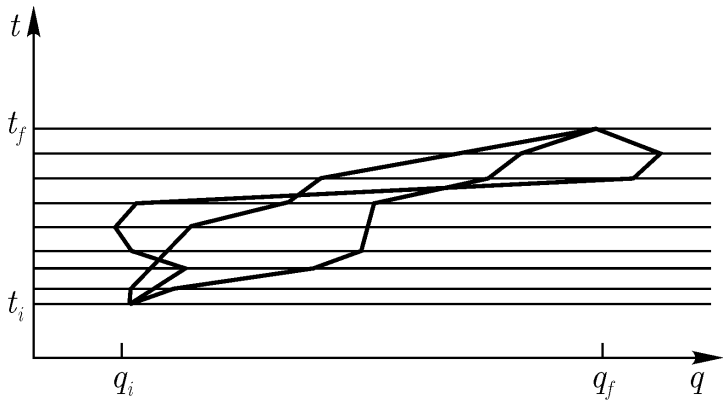


Figure 9.4

so that

$$\langle q_f t_f | q_i t_i \rangle = \lim_{n\to\infty} \int \prod_{j=1}^{n} dq_j \prod_{i=0}^{n} \frac{dp_i}{2\pi\hbar} \exp\left\{ \frac{i}{\hbar} \sum_{l=0}^{n} [p_l(q_{l+1} - q_l) - \tau H(p_l, \bar{q}_l)] \right\}, \tag{9.23}$$

where $q_0 = q_i$ and $q_{n+1} = q_f$. In fact we are dealing here with a multiple integral of an infinite order. Usually equation (9.23) is written in the symbolic form

$$\langle q_f t_f | q_i t_i \rangle = \int \frac{\mathcal{D}q(t)\mathcal{D}p(t)}{2\pi\hbar} \exp\left\{ \frac{i}{\hbar} \int_{t_i}^{t_f} dt\, [p\dot{q} - H(p,q)] \right\}, \tag{9.24}$$

where $q(t_i) = q_i$ and $q(t_f) = q_f$. This form defines the measure for integration over all trajectories $(q(t), p(t))$ in the *phase* space of a particle, and it has no other meaning except the compact notation for (9.23). The situation here is quite similar to the definition of the usual integral via the limiting behavior of Riemann sums. This notation (9.24) introduces the notion of a *functional* (continual) integral over all trajectories (paths) in the phase space. Variables $p(t)$ and $q(t)$ entering (9.24) are the usual c-number functions.

The definition of the propagator via the functional integral over all trajectories in phase space (9.24) is absolutely general and is valid for the arbitrary Hamiltonian $H(p,q)$. In the case of the Hamiltonian given by (9.17) we can make further simplifications and transform the propagator to a functional integral over all the paths in coordinate space only. In this case we can write

$$\langle q_f t_f | q_i t_i \rangle = \lim_{n\to\infty} \int \prod_{j=1}^{n} dq_j \prod_{i=0}^{n} \frac{dp_i}{2\pi\hbar} \exp\left\{ \frac{i}{\hbar} \sum_{l=0}^{n} \left[p_l(q_{l+1} - q_l) - \frac{p_l^2}{2m}\tau - V(\bar{q}_l)\tau \right] \right\}. \tag{9.25}$$

The integrals over p_j here are easily calculated using the standard expressions presented below. Then we get

$$\langle q_f t_f | q_i t_i \rangle = \lim_{n\to\infty} \left(\frac{m}{2\pi i\hbar\tau} \right)^{\frac{n+1}{2}} \int \prod_{j=1}^{n} dq_j \exp\left\{ \frac{i\tau}{\hbar} \sum_{l=0}^{n} \left[\frac{m}{2} \left(\frac{q_{l+1} - q_l}{\tau} \right)^2 - V(\bar{q}_l) \right] \right\}, \tag{9.26}$$

so that in a continuous limit we can write

$$\begin{aligned} \langle q_f t_f | q_i t_i \rangle &= \mathcal{N} \int \mathcal{D}q(t) \exp\left\{ \frac{i}{\hbar} \int_{t_i}^{t_f} dt \left[\frac{m}{2}\dot{q}^2 - V(q) \right] \right\} \\ &= \mathcal{N} \int \mathcal{D}q(t) \exp\left\{ \frac{i}{\hbar} \int_{t_i}^{t_f} dt L(q,\dot{q}) \right\} = \mathcal{N} \int \mathcal{D}q(t) \exp\left\{ \frac{i}{\hbar} S \right\}, \end{aligned} \tag{9.27}$$

where $L = T - V$ is the classical Lagrange function of our particle, while $S = \int_{t_i}^{t_f} dt L(q,\dot{q})$ is the classical action, calculated for an arbitrary trajectory $q(t)$, con-

necting the initial point $q(t_i)$ with final $q(t_f)$. The functional integral (9.27) is taken over all the possible trajectories, connecting the initial and final points. The normalization factor $\mathcal{N}$ introduced here is formally divergent in the limit of $n \to \infty$, but this is irrelevant; as we shall see later it always cancels from physical transition amplitudes.

The remarkable result (9.27) allows, in particular, a qualitative understanding of the physical origin of the classical principle of least action. We can see that in the classical limit of $\hbar \to 0$ the Feynman integral (9.27) contains the continuum of rapidly oscillating factors of $\exp(iS/\hbar)$, which "on the average" cancel each other. The only "surviving" one is the contribution of the most slowly changing factor with $S_{\min}$, which corresponds to the *single* trajectory described by the least action principle and the Newtonian equations of motion of classical mechanics.

Remarks on some useful integrals

Below we present some common integrals, which are useful for practical calculations with functional integrals. First of all, we have the well-known Gauss–Poisson integral:

$$\int_{-\infty}^{\infty} dx\, e^{-ax^2} = \sqrt{\frac{\pi}{a}}, \qquad a > 0\,. \tag{9.28}$$

This result follows immediately if we write

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, e^{-a(x^2+y^2)} = \frac{\pi}{a}\,, \tag{9.29}$$

which, after the transformation to polar coordinates in the (x, y) plane, reduces to

$$\int_0^{2\pi} d\theta \int_0^{\infty} dr\, re^{-ar^2} = \pi \int_0^{\infty} d(r^2)\, e^{-ar^2} = \frac{\pi}{a}\,. \tag{9.30}$$

The last equality is obvious and proves (9.28).

Consider now the integral of an exponent, depending on the quadratic form:

$$\int_{-\infty}^{\infty} dx\, e^{-ax^2+bx+c} \equiv \int_{-\infty}^{\infty} dx\, e^{q(x)}\,, \tag{9.31}$$

where we assume $a > 0$. Then we have $q'(x) = -2ax + b$, $q''(x) = -2a$, $q'''(x) = 0 \ldots$, and we easily find $\bar{x}$, the value of x corresponding to the minimum of $q(x)$:

$$\bar{x} = \frac{b}{2a}\,, \quad q(\bar{x}) = \frac{b^2}{4a} + c\,. \tag{9.32}$$

Now it is convenient to rewrite $q(x)$ as

$$q(x) = q(\bar{x}) - a(x - \bar{x})^2\,. \tag{9.33}$$

Then,

$$\int_{-\infty}^{\infty} dx\, e^{q(x)} = e^{q(\bar{x})} \int_{-\infty}^{\infty} dx e^{-a(x-\bar{x})^2} = e^{q(\bar{x})} \sqrt{\frac{\pi}{a}}\,, \tag{9.34}$$

so that finally we have

$$\int_{-\infty}^{\infty} dx e^{-ax^2+bx+c} \equiv \int_{-\infty}^{\infty} e^{q(x)} = \exp\left(\frac{b^2}{4a} + c\right) \sqrt{\frac{\pi}{a}}\,. \tag{9.35}$$

This expression was used to derive (9.26), (9.25).

Let us quote also the generalization of (9.35) for the case of n integration variables [56]:

$$\int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_n \exp\{i\lambda[(x_1-a)^2+(x_2-x_1)^2+\cdots+(b-x_n)^2]\} = \left[\frac{i^n\pi^n}{(n+1)\lambda^n}\right]^{1/2} \exp\left[\frac{i\lambda}{n+1}(b-a)^2\right], \tag{9.36}$$

which will be useful in the following.

Equation (9.27), in fact, contains the whole of the quantum mechanics of a particle and is widely used to solve practical problems [20]. Let us show how the common Schroedinger equation is derived from this representation. First of all, we write the basic relation (9.1), which connects the wave function at moment t_2 with its value at the previous moment t_1:

$$\psi(x_2,t_2) = \int_{-\infty}^{\infty} dx_1 K(x_2t_2;x_1t_1)\psi(x_1t_1)\,. \tag{9.37}$$

Let moments t_2 and t_1 be very close, so that $t_2 = t_1 + \varepsilon$, where $\varepsilon \to 0$. Then the propagator is determined by the contribution of a single small segment of the trajectory, so that using (9.26) we can write (9.37) as

$$\psi(x,t+\varepsilon) = A\int_{-\infty}^{\infty} \exp\left[\frac{i}{\hbar}\frac{m(x-y)^2}{2\varepsilon}\right] \exp\left[-\frac{i}{\hbar}\varepsilon V\left(\frac{x+y}{2},t\right)\right]\psi(y,t)dy\,, \tag{9.38}$$

where $A = \left(\frac{m}{2\pi i\hbar\varepsilon}\right)^{1/2}$. Due to the first exponent, a significant contribution to the integral originates only from the values y close to x. Making the variable transformation $y = x + \eta$ we rewrite (9.38) as

$$\psi(x,t+\varepsilon) = A\int_{-\infty}^{\infty} \exp\left(\frac{im\eta^2}{2\hbar\varepsilon}\right) \exp\left[-\frac{i\varepsilon}{\hbar}V\left(x+\frac{\eta}{2},t\right)\right]\psi(x+\eta,t)d\eta\,. \tag{9.39}$$

The main contribution here comes from the small values of η, and expanding both sides of (9.39) we have

$$\psi(x,t) + \varepsilon\frac{\partial\psi}{\partial t} = A\int_{-\infty}^{\infty} \exp\left(\frac{im\eta^2}{2\hbar\varepsilon}\right)\left[1-\frac{i\varepsilon}{\hbar}V(x,t)\right]\left[\psi(x,t)+\eta\frac{\partial\psi}{\partial x}+\frac{1}{2}\eta^2\frac{\partial^2\psi}{\partial x^2}\right]d\eta\,. \tag{9.40}$$

Now we can take into account that

$$A\int_{-\infty}^{\infty} e^{im\eta^2/2\hbar\varepsilon}d\eta = 1\,,$$
$$A\int_{-\infty}^{\infty} e^{im\eta^2/2\hbar\varepsilon}\eta d\eta = 0\,,$$
$$A\int_{-\infty}^{\infty} e^{im\eta^2/2\hbar\varepsilon}\eta^2 d\eta = \frac{i\hbar\varepsilon}{m}\,.$$

Then (9.40) reduces to

$$\psi(x,t) + \varepsilon\frac{\partial\psi}{\partial t} = \psi - \frac{i\varepsilon}{\hbar}V\psi - \frac{\hbar\varepsilon}{2im}\frac{\partial^2\psi}{\partial x^2}\,. \tag{9.41}$$

This equation is valid (for $\varepsilon \to 0$) if ψ satisfies the one-dimensional Schroedinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V(x,t)\psi\,. \tag{9.42}$$

This completes our derivation.

9.2 Perturbation theory

Let us consider potential $V(x)$ as a small perturbation. More strictly speaking we require the smallness (in comparison with $\hbar$) of the time integral of $V(x,t)$. Then we can write an expansion:

$$\exp\left\{-\frac{i}{\hbar}\int_{t_i}^{t_f} dt V(x,t)\right\} \approx 1 - \frac{i}{\hbar}\int_{t_i}^{t_f} dt V(x,t) - \frac{1}{2!\hbar^2}\left[\int_{t_i}^{t_f} dt V(x,t)\right]^2 + \cdots\,. \tag{9.43}$$

Using this type of expansion in equation (9.27), we can obtain the the perturbation expansion for the propagator $K(x_f t_f; x_i t_i)$:

$$K = K_0 + K_1 + K_2 + \cdots\,. \tag{9.44}$$

The zero-th order term here represents the free particle propagator:

$$K_0 = \mathcal{N}\int \mathcal{D}x \exp\left(\frac{i}{\hbar}\int dt\,\frac{1}{2}m\dot{x}^2\right)\,. \tag{9.45}$$

To make an explicit calculation, we return to the definition of the path integral (9.23) and write (9.45) as a limit of the multiple integral (cf. (9.26)):

$$K_0 = \lim_{n\to\infty}\left(\frac{m}{2\pi i\hbar\tau}\right)^{\frac{n+1}{2}}\int_{-\infty}^{\infty}\prod_{j=1}^{n} dx_j \exp\left[\frac{im}{2\hbar\tau}\sum_{l=0}^{n}(x_{l+1}-x_l)^2\right]. \tag{9.46}$$

Denoting the multiple integral here as I, we can calculate it using (9.36) and obtain

$$I = \frac{1}{(n+1)^{1/2}}\left(\frac{i2\pi\hbar\tau}{m}\right)^{n/2}\exp\left\{\frac{im}{2\hbar(n+1)\tau}(x_f - x_i)^2\right\}. \tag{9.47}$$

Taking $(n+1)\tau = t_f - t_i$, from (9.46) we get the explicit form of the free particle propagator

$$K_0(x_f t_f; x_i t_i) = \theta(t_f - t_i)\left(\frac{m}{2\pi i\hbar(t_f - t_i)}\right)^{1/2}\exp\left\{\frac{im(x_f - x_i)^2}{2\hbar(t_f - t_i)}\right\}\,, \tag{9.48}$$

where we have added a factor of $\theta(t_f - t_i)$, which guarantees causality. The generalization of this expression for a particle moving in three-dimensional space is quite obvious: the corresponding propagator reduces to the product of free propagators (9.48) along three axes x, y, z.

In Chapter 4 we have seen the that particle propagator satisfies Schroedinger equation with the δ-source:

$$\left[i\hbar \frac{\partial}{\partial t_f} - H(x_f) \right] K(x_f t_f; x_i t_i) = i\hbar \delta(t_f - t_i)\delta(x_f - x_i)\,. \tag{9.49}$$

For one-dimensional free particle motion $H(x_f) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_f^2}$. Correspondingly, the free particle propagator satisfies the equation

$$\left[i\hbar \frac{\partial}{\partial t_f} + \frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_f^2} \right] K_0(x_f t_f; x_i t_i) = i\hbar \delta(t_f - t_i)\delta(x_f - x_i)\,. \tag{9.50}$$

It can also be checked by direct substitution of (9.48) into this equation.

In equations (9.48) and (9.50) we make the replacements $t \to -i\hbar t$ and $\frac{\hbar}{2m} \to D$; equation (9.50) transforms into

$$\left[\frac{\partial}{\partial t_f} - D\frac{\partial^2}{\partial x_f^2} \right] K_0(x_f t_f; x_i t_i) = \delta(t_f - t_i)\delta(x_f - x_i)\,, \tag{9.51}$$

and $K_0(x_f t_f; x_i t_i)$ now represents the Green's function of the *diffusion equation* [70] with the diffusion coefficient D. All the imaginary terms of (9.48) disappear, and this expression describes the diffusion of particles from the point-like source. In fact, path integrals first appeared in the theory of diffusion processes, where these are called Wiener integrals. The disappearance of oscillations from (9.48) (which are replaced by rapidly the diminishing exponents of diffusion theory) is quite convenient for numerical calculations, particularly for calculations of path integrals by Monte-Carlo algorithms. Such formal transformation to the imaginary time is widely used in studies of different problems of quantum mechanics and quantum field theory.

There is one more aspect of transformation to imaginary time, which is even more fundamental for physics. Equilibrium statistical mechanics is based of the use of Gibbs canonical distribution, with the density matrix of the following form [36]:

$$\rho = \frac{1}{Z} e^{-\beta H}\,, \tag{9.52}$$

where H is the system Hamiltonian, Z is the partition function, and $\beta = \frac{1}{T}$ is the inverse temperature. Then it is easy to get

$$\frac{\partial \rho}{\partial \beta} = -H\rho\,. \tag{9.53}$$

But this equation (also called the Bloch equation) can be obtained from the usual Schroedinger equation:

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi \tag{9.54}$$

after the formal replacement $\psi \to \rho$, $t \to -i\hbar\beta$. Thus we may say that all of statistical mechanics is the same theory as quantum mechanics in "imaginary time". The calculation of

the equilibrium density matrix of the system of interacting particles can be performed solving equation (9.53) with the help of Green's function formalism (propagators) in imaginary (so-called "Matsubara") time [1]. These propagators can be represented by Feynman path integrals (Wiener integrals), which allows the development of an alternative general approach to problems of statistical physics [19].

Now let us calculate K_1, the first order correction over the potential $V(x)$. From (9.26) and (9.43) we have

$$K_1 = -\frac{i\tau}{\hbar}\lim_{n\to\infty}\left(\frac{m}{2\pi i\hbar\tau}\right)^{\frac{n+1}{2}}\sum_{i=1}^{n}\int dx_1\cdots dx_n V(x_i,t_i)\exp\left\{\frac{im}{2\hbar\tau}\sum_{j=0}^{n}(x_{j+1}-x_j)^2\right\}, \tag{9.55}$$

where we have replaced integration over t by summation over t_i. As V depends here on x_i, we break the sum in the exponent in two: one performed from $j=0$ to $j=i-1$ and the other from $j=i$ to $j=n$. Let us also separate the integral over x_i. As a result, equation (9.55) is rewritten as

$$\begin{aligned} K_1 = -\lim_{n\to\infty}\frac{i\tau}{\hbar}\sum_{i=1}^{n}\int & \\ \times dx_i\left\{\left(\frac{m}{2\pi i\hbar\tau}\right)^{\frac{n-i+1}{2}}\int dx_{i+1}\cdots dx_n\exp\left[\frac{im}{2\hbar\tau}\sum_{j=i}^{n}(x_{j+1}-x_j)^2\right]\right\} & \\ \times V(x_i,t_i)\left\{\left(\frac{m}{2\pi i\hbar\tau}\right)^{\frac{i}{2}}\int dx_1\cdots dx_{i-1}\exp\left[\frac{im}{2\hbar\tau}\sum_{j=0}^{i-1}(x_{j+1}-x_j)^2\right]\right\}. & \end{aligned} \tag{9.56}$$

The terms in the figure brackets are equal to $K_0(x_f t_f; xt)$ and $K_0(xt; x_i t_i)$, so that after the replacement $\tau\sum_i \int dx_i$ by $\int dx\int dt$ equation (9.56) reduces to

$$K_1 = -\frac{i}{\hbar}\int_{t_i}^{t_f} dt\int_{-\infty}^{\infty} dx\, K_0(x_f t_f; xt)V(x,t)K_0(xt; x_i t_i)\,. \tag{9.57}$$

Taking into account that $K_0(x_f t_f; xt) = 0$ for $t > t_f$, while $K_0(xt; x_i t_i) = 0$ for $t < t_i$, we can write equation (9.57) as

$$K_1 = -\frac{i}{\hbar}\int_{-\infty}^{\infty} dt\int_{-\infty}^{\infty} dx\, K_0(x_f t_f; xt)V(x,t)K_0(xt; x_i t_i)\,, \tag{9.58}$$

which is the final expression for the first order correction to the propagator (Green's function) of our particle.

Quite similarly, but by more tedious calculations, we can find the second-order correction:

$$K_2(x_f t_f; x_i t_i) = \left(-\frac{i}{\hbar}\right)^2 \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \times K_0(x_f t_f; x_2 t_2) V(x_2 t_2) K_0(x_2 t_2; x_1 t_1) V(x_1 t_1) K_0(x_1 t_1; x_i t_i)\,. \tag{9.59}$$

Now the structure of the higher orders becomes clear, and we obtain the perturbation series for the propagator:

$$\begin{aligned} K(x_f t_f; x_i t_i) = K_0(x_f t_f; x_i t_i) &- \frac{i}{\hbar} \int dt_1 dx_1 K_0(x_f t_f; x_1 t_1) V(x_1, t_1) K_0(x_1 t_1; x_i t_i) \\ &- \frac{1}{\hbar^2} \int dt_1 dt_2 dx_1 dx_2 K_0(x_f t_f; x_2 t_2) V(x_2 t_2) K_0(x_2 t_2; x_1 t_1) \\ &\times V(x_1 t_1) K_0(x_1 t_1; x_i t_i) + \cdots\,. \end{aligned} \tag{9.60}$$

which coincides with the similar expansion introduced in Chapter 4. Note that in equation (9.59) there is no factor of $1/2!$, which is present in expansion (9.43). This is due to the fact that two interactions V at different moments of time are equivalent, and we can write

$$\begin{aligned} &\frac{1}{2!} \int_{-\infty}^{\infty} dt' \int_{-\infty}^{\infty} dt'' V(t') V(t'') \\ &\quad = \int_{-\infty}^{\infty} dt' \int_{-\infty}^{\infty} dt'' [\theta(t' - t'') V(t') V(t'') + \theta(t'' - t') V(t') V(t'') \\ &\quad = \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 V(t_1) V(t_2) \theta(t_1 - t_2)\,. \end{aligned} \tag{9.61}$$

For the same reason, the correction of the arbitrary order K_n does not contain the factor of $1/n!$. It is clear that expansion (9.60) corresponds to the simple diagram technique: each term of the series can be expressed by a diagram, if we associate the straight line with the propagator and the wavy lines with the potential, acting at appropriate points of space at appropriate moments of time (over which we perform integration).

Substitution of expansion (9.60) into (9.1) gives

$$\begin{aligned} \psi(x_f t_f) &= \int dx_i\, K(x_f t_f; x_i t_i) \psi(x_i t_i) \\ &= \int dx_i\, K_0(x_f t_f; x_i t_i) \psi(x_i t_i) \\ &\quad - \frac{i}{\hbar} \int dt \int dx \int dx_i\, K_0(x_f t_f; x t) V(x, t) K_0(x t, x_i t_i) \psi(x_i t_i) + \cdots\,. \end{aligned} \tag{9.62}$$

The contribution of higher order terms, which are not written here, obviously reduces to the replacement of the last propagator K_0 by the complete propagator K. Corre-

spondingly we obtain the exact integral equation for the wave function:

$$\psi(x_f t_f) = \int dx_i\, K_0(x_f t_f; x_i t_i)\psi(x_i t_i) - \frac{i}{\hbar}\int dt \int dx K_0(x_f t_f; xt)V(x,t)\psi(xt)\,, \tag{9.63}$$

which is just equivalent to the Schroedinger equation for the problem under discussion. Assuming that for $t_i \to -\infty$ the wave function is the solution of the free particle Schroedinger equation (plane wave!) and denoting it by $\varphi(xt)$, we may rewrite (9.63) as

$$\psi(x_f t_f) = \varphi(x_f t_f) - \frac{i}{\hbar}\int dt \int dx K_0(x_f t_f; xt)V(x,t)\psi(xt)\,, \tag{9.64}$$

because the plane wave remains the plane wave during free particle motion.

For practical tasks it is more convenient to use the momentum representation. Let $K(\mathbf{p}_1 t_1; \mathbf{p}_0 t_0)$ be the probability amplitude for a particle with momentum $\mathbf{p}_0$ at moment t_0 to be registered at a later moment t_1 with momentum $\mathbf{p}_1$. This amplitude is given by

$$K(\mathbf{p}_1 t_1; \mathbf{p}_0 t_0) = \int d\mathbf{x}_0 \int d\mathbf{x}_1 \exp\left(-\frac{i}{\hbar}\mathbf{p}_1\mathbf{x}_1\right) K(\mathbf{x}_1 t_1; \mathbf{x}_0 t_0) \exp\left(\frac{i}{\hbar}\mathbf{p}_0\mathbf{x}_0\right), \tag{9.65}$$

where the free propagator $K(\mathbf{x}_1 t_1; \mathbf{x}_0 t_0)$ for a particle moving in three-dimensional space (in accordance with the remark made after equation (9.48)) has the form

$$K_0(\mathbf{x}_1 t_1; \mathbf{x}_0 t_0) = \theta(t_1 - t_0)\left(\frac{m}{2\pi i\hbar(t_1 - t_0)}\right)^{3/2} \exp\left\{\frac{im(\mathbf{x}_1 - \mathbf{x}_0)^2}{2\hbar(t_1 - t_0)}\right\}. \tag{9.66}$$

Then we have

$$K(\mathbf{p}_1 t_1; \mathbf{p}_0 t_0) = \theta(t_1 - t_0)\left[\frac{m}{2\pi i\hbar(t_1 - t_0)}\right]^{3/2} \times \int d\mathbf{x}_0 \int d\mathbf{x}_1 \exp\left[\frac{i}{\hbar}(\mathbf{p}_0\mathbf{x}_0 - \mathbf{p}_1\mathbf{x}_1)\right] \exp\left[\frac{im(\mathbf{x}_0 - \mathbf{x}_1)^2}{2\hbar(t_1 - t_0)}\right]. \tag{9.67}$$

Let us introduce the new integration variables

$$\mathbf{x} = \mathbf{x}_0 - \mathbf{x}_1\,, \quad \mathbf{X} = \mathbf{x}_0 + \mathbf{x}_1\,, \quad \mathbf{p} = \mathbf{p}_0 - \mathbf{p}_1\,, \quad \mathbf{P} = \mathbf{p}_0 + \mathbf{p}_1\,, \tag{9.68}$$

so that $2(\mathbf{p}_0\mathbf{x}_0 - \mathbf{p}_1\mathbf{x}_1) = \mathbf{P}\mathbf{x} + \mathbf{p}\mathbf{X}$. The Jacobian of this variables transformation is equal to $(1/2)^3 = 1/8$. Correspondingly, equation (9.67) is rewritten as

$$K(\mathbf{p}_1 t_1; \mathbf{p}_0 t_0) = \theta(t_1 - t_0)\left(\frac{\alpha}{i\pi}\right)^{3/2} \frac{1}{8}\int d\mathbf{X} \exp\left(\frac{i}{2\hbar}\mathbf{p}\mathbf{X}\right) \int d\mathbf{x} \exp\left(\frac{i}{2\hbar}\mathbf{P}\mathbf{x}\right) e^{i\alpha\mathbf{x}^2}, \tag{9.69}$$

where $\alpha = \frac{m}{2\hbar(t_1-t_0)}$. The first integral here is equal to $8(2\pi\hbar)^3\delta(\mathbf{p}) = 8(2\pi\hbar)^3\delta(\mathbf{p}_0 - \mathbf{p}_1)$, so that

$$K(\mathbf{p}_1 t_1; \mathbf{p}_0 t_0) = (2\pi\hbar)^3\theta(t_1 - t_0)\delta(\mathbf{p}_0 - \mathbf{p}_1)\left(\frac{\alpha}{i\pi}\right)^{3/2}\int d\mathbf{x}\exp\left(\frac{i}{2\hbar}\mathbf{P}\mathbf{x} + i\alpha\mathbf{x}^2\right), \quad (9.70)$$

and using (9.35) we obtain

$$K(\mathbf{p}_1 t_1; \mathbf{p}_0 t_0) = (2\pi\hbar)^3\theta(t_1 - t_0)\delta(\mathbf{p}_0 - \mathbf{p}_1)\exp\left[-\frac{i\mathbf{P}^2(t_1 - t_0)}{8m\hbar}\right], \quad (9.71)$$

where the δ-function expresses momentum conservation. Taking into account $\mathbf{P}^2 = 4\mathbf{p}_0^2$, we get finally

$$K(\mathbf{p}_1 t_1; \mathbf{p}_0 t_0) = (2\pi\hbar)^3\theta(t_1 - t_0)\delta(\mathbf{p}_0 - \mathbf{p}_1)\exp\left[-\frac{i\mathbf{p}_0^2(t_1 - t_0)}{2m\hbar}\right]. \quad (9.72)$$

At last we can calculate the Fourier transform of the propagator over time:

$$\begin{aligned} K(\mathbf{p}_1 E_1; \mathbf{p}_0 E_0) &= \int dt_0 \int dt_1 \exp\left(\frac{i}{\hbar}E_1 t_1\right) K(\mathbf{p}_1 t_1; \mathbf{p}_0 t_0)\exp\left(-\frac{i}{\hbar}E_0 t_0\right) \\ &= (2\pi\hbar)^3\delta(\mathbf{p}_0 - \mathbf{p}_1)\int dt_0 \int dt_1 \theta(\tau) \\ &\quad \times \exp\left(-\frac{i\mathbf{p}_1^2}{2m\hbar}\tau\right)\exp\left[\frac{i}{\hbar}(E_1 t_1 - E_0 t_0)\right], \end{aligned} \quad (9.73)$$

where we have introduced $\tau = t_1 - t_0$. Considering τ and t_0 as independent variables, we obtain

$$\begin{aligned} K(\mathbf{p}_1 E_1; \mathbf{p}_0 E_0) &= (2\pi\hbar)^3\delta(\mathbf{p}_0 - \mathbf{p}_1)\int_{-\infty}^{\infty} dt_0 \exp\left[\frac{i}{\hbar}(E_1 - E_0)t_0\right] \\ &\quad \times \int_{-\infty}^{\infty} d\tau\,\theta(\tau)\exp\left[\frac{i}{\hbar}(E_1 - \frac{\mathbf{p}_1^2}{2m})\tau\right]. \end{aligned} \quad (9.74)$$

The first integral here yields $(2\pi\hbar)\delta(E_1 - E_0)$, while the second one, because of the presence of $\theta(\tau)$, should be understood as[1]

$$\lim_{\delta\to+0}\int_0^{\infty} d\tau e^{i(E_1 - \mathbf{p}_1^2/2m + i\delta)\tau/\hbar} = \frac{i\hbar}{E_1 - \frac{\mathbf{p}_1^2}{2m} + i\delta}. \quad (9.75)$$

[1] Fourier transform of $\theta(t)$ is defined by

$$\theta(t) = \lim_{\delta\to+0}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}e^{-i\omega t}\frac{i}{\omega + i\delta},$$

which can be easily checked making the integration along the real axis and closing integration contour in upper or lower half-planes of complex ω, depending on the sign of t.

Thus finally we have

$$K(\mathbf{p}_1 E_1; \mathbf{p}_0 E_0) = (2\pi\hbar)^4 \delta(\mathbf{p}_0 - \mathbf{p}_1)\delta(E_0 - E_1)\frac{i\hbar}{E_1 - \frac{\mathbf{p}_1^2}{2m} + i\delta}, \quad \delta \to +0, \tag{9.76}$$

which is the Fourier transform of the *retarded* Green's function of the free particle, where δ-functions express the momentum and energy conservation laws. Note that the pole here is in fact determined by the kinetic energy of a particle, which reflects the general property of Green's functions [1]: their poles determine the energy spectrum of the corresponding particles (quasiparticles).

If we introduce the Fourier transform of potential, writing $V(\mathbf{x}, t)$ as

$$V(\mathbf{x}, t) = \int \frac{d\omega}{2\pi} \int \frac{d^3\mathbf{q}}{(2\pi)^3} e^{i(\mathbf{q}\mathbf{x} - \omega t)} V(\mathbf{q}\omega), \tag{9.77}$$

perturbation series (9.60) generates the standard diagram technique in momentum representation for the Green's function of a particle in an external field [1].

9.3 Functional derivatives

The Green's function (particle propagator) written in the form of Feynman path integral

$$\begin{aligned} \langle q_f t_f | q_i t_i \rangle &= \mathcal{N} \int \mathcal{D}q(t) \exp\left\{ \frac{i}{\hbar} \int_{t_i}^{t_f} dt \left[\frac{m}{2}\dot{q}^2 - V(q) \right] \right\} \\ &= \mathcal{N} \int \mathcal{D}q(t) \exp\left\{ \frac{i}{\hbar} \int_{t_i}^{t_f} dt L(q, \dot{q}) \right\} \end{aligned} \tag{9.78}$$

introduces the notion of the *functional integral*: integration is performed here over all *functions* (trajectories) $q(t)$, connecting the initial and final points. Thus, the calculation of (9.78) relates the whole set of functions $q(t)$ with some concrete (complex) *number*: the amplitude of quantum mechanical transition in the left-hand side. Thus, equation (9.78) is the concrete realization of the mathematical notion of *functional*: the mapping of the set of functions into the set of numbers:

- *Functional: function* $\Rightarrow$ *number*

In contrast, the usual function defines the mapping of one set of numbers into another set of numbers:

- *Function: number* $\Rightarrow$ *number*

In particular, the functional is not simply the function of another function (this is again just a function!).

Usually, the functional F of function $f(x)$ is denoted as $F[f(x)]$. A typical example of a functional is the definite integral: $F[f(x)] = \int_a^b dx f(x)$.

Let us now define the *functional derivative*. In analogy with the the usual differentiation, the functional (or variational) derivative of some functional $F[f(x)]$ over the function $f(y)$ is defined as

$$\frac{\delta F[f(x)]}{\delta f(y)} = \lim_{\varepsilon\to 0} \frac{F[f(x)+\varepsilon\delta(x-y)] - F[f(x)]}{\varepsilon}\,. \tag{9.79}$$

For example, for $F[f(x)]$ given by definite integral:

$$\frac{\delta F[f(x)]}{\delta f(y)} = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[\int dx[f(x)+\varepsilon\delta(x-y)] - \int dx f(x)\right] = \int dx\delta(x-y) = 1\,. \tag{9.80}$$

As another example, we consider the functional

$$F_x[f] = \int dy f(y)G(x,y)\,, \tag{9.81}$$

where the variable x in the left-hand side is considered as a parameter. Then we have

$$\begin{aligned}\frac{\delta F_x[f]}{\delta f(z)} &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[\int dy\{G(x,y)[f(y)+\varepsilon\delta(y-z)]\} - \int dy G(x,y)f(y)\right]\\ &= \int dy G(x,y)\delta(y-z) = G(x,z)\,.\end{aligned} \tag{9.82}$$

These expressions are sufficient for understanding all the expressions related to functional differentiation which will be used below.

9.4 Some properties of functional integrals

The amplitude of quantum particle transition from initial point $q_i t_i$ to final $q_f t_f$ is given by

$$\begin{aligned}\langle q_f t_f | q_i t_i\rangle &= \mathcal{N}\int_{q(t_i)=q_i}^{q(t_f)=q_f} \mathcal{D}q(t)\exp\left\{\frac{i}{\hbar}\int_{t_i}^{t_f} dt\left[\frac{m}{2}\dot{q}^2 - V(q)\right]\right\}\\ &= \mathcal{N}\int_{q(t_i)=q_i}^{q(t_f)=q_f} \mathcal{D}q(t)\exp\left\{\frac{i}{\hbar}\int_{t_i}^{t_f} dt L(q,\dot{q})\right\}.\end{aligned} \tag{9.83}$$

Let us derive some formal relations, which will be quite useful below during the generalization to quantum field theory. We can add to the Lagrange function of our particle an extra "source" term:

$$L \to L + \hbar J(t)q(t) \tag{9.84}$$

where $J(t)$ is some arbitrary function of time. Let us assume that $J(t)$ is nonzero at some time interval between moments t and t' ($t < t'$), which is shown in Figure 9.5. Consider also the moment T, previous to t, and another moment T', which is later

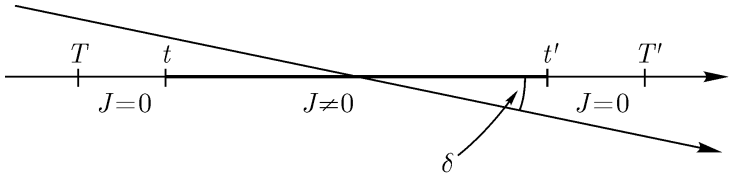


Figure 9.5

than t'. Then the transition amplitude of the system, interacting with source, between arbitrary states (points) in these moments of time is given by

$$\langle Q'T'|QT\rangle^J = \mathcal{N}\int \mathcal{D}q(t)\exp\left\{\frac{i}{\hbar}\int_T^{T'} dt\left[L(q,\dot{q})+\hbar Jq\right]\right\}. \tag{9.85}$$

On the other hand, using (9.4) we can write

$$\langle Q'T'|QT\rangle^J = \int dq'\int dq\langle Q'T'|q't'\rangle\langle q't'|qt\rangle^J\langle qt|QT\rangle\,, \tag{9.86}$$

where (due to our assumption on the form of $J(t)$) only the "intermediate" propagator is source dependent. Using (9.10) we have

$$\begin{aligned}\langle Q'T'|q't'\rangle &= \left\langle Q'\left|\exp\left(-\frac{i}{\hbar}HT'\right)\exp\left(\frac{i}{\hbar}Ht'\right)\right|q'\right\rangle = \\ &= \sum_m \varphi_m(Q')\varphi_m^*(q')\exp\left[\frac{i}{\hbar}E_m(t'-T')\right],\end{aligned} \tag{9.87}$$

where $\{\varphi_m(q)\}$ is the complete set of the eigenfunctions of the Hamiltonian (energy operator). In a similar way we obtain

$$\langle qt|QT\rangle = \sum_n \varphi_n(q)\varphi_n^*(Q)\exp\left[-\frac{i}{\hbar}E_n(t-T)\right]. \tag{9.88}$$

Substituting (9.87) and (9.88) into (9.86) and making the replacement $T' \to T'e^{-i\delta}$ and $T \to Te^{-i\delta}$ ("rotating" the time axis by an arbitrary angle $\delta < \pi/2$ in a complex plane of "time", as shown in Figure 9.5), we perform the limits of $T' \to \infty$ and $T \to -\infty$. In this case, because of the "damping" factor δ in the transition amplitude (9.86), all the contributions of the states with $E_n > 0, E_m > 0$ just vanish, while the term with $E_0 = 0$ survives, which gives the contribution of the *ground state* level of

our particle in potential $V(q)$[2]. Then we get

$$\lim_{T'\to\infty e^{-i\delta}}\lim_{T\to-\infty e^{-i\delta}}\langle Q'T'|QT\rangle^J = \varphi_0^*(Q)\varphi_0(Q')\exp\left[-\frac{i}{\hbar}E_0(T'-T)\right] \times\int dq'\int dq\varphi_0^*(q't')\langle q't'|qt\rangle^J\varphi_0(qt) \tag{9.89}$$

or

$$\int dq'\int dq\varphi_0^*(q't')\langle q't'|qt\rangle^J\varphi_0(qt) = \lim_{T'\to\infty e^{-i\delta}}\lim_{T\to-\infty e^{-i\delta}}\frac{\langle Q'T'|QT\rangle^J}{\varphi_0^*(Q)\varphi_0(Q')\exp\left[-\frac{i}{\hbar}E_0(T'-T)\right]}. \tag{9.90}$$

The left-hand side of this expression represents the transition amplitude (in the presence of the source), averaged over the ground state ("vacuum") of the system. Now we can make $t'\to\infty$ and $t\to-\infty$ and introduce denoting our averaged transition amplitude (9.90) as $\langle 0,\infty|0,-\infty\rangle^J$, which corresponds to a "vacuum–vacuum" transition during the infinite interval of time. The denominator in the right-hand side of (9.90) is a simple number, and we can write

$$\langle 0,\infty|0,-\infty\rangle^J \sim \lim_{T'\to\infty e^{-i\delta}}\lim_{T\to-\infty e^{-i\delta}}\langle Q'T'|QT\rangle^J \equiv Z[J], \tag{9.91}$$

where we have introduced the following functional of the source:

$$Z[J] = \lim_{T'\to\infty e^{-i\delta}}\lim_{T\to-\infty e^{-i\delta}}\mathcal{N}\int\mathcal{D}Q(t)\exp\left\{\frac{i}{\hbar}\int_T^{T'}dt[L(Q,\dot{Q})+\hbar JQ]\right\}. \tag{9.92}$$

Note that instead of time axis "rotation" in a complex plane to extract a ground state contribution, we could just add a small negative imaginary part to the Hamiltonian of our system (9.17), which can conveniently be written as $-\frac{1}{2}i\varepsilon q^2(\varepsilon\to+0)$. In this case the entire energy level will acquire small imaginary parts, which in the limit of $T'\to\infty, T\to-\infty$ will lead to the same effect of the exponential damping of the contributions of levels with $E_n > 0$[3]. In a Lagrange function this L is equivalent to

[2] Here it is important that the eigenlevels of energy can be ordered: $E_0 < E_1 < E_2 < \cdots < E_n < \cdots$, so that this procedure separates the contribution of the lowest energy level, which may be set as energy zero (or left explicitly as the most slowly vanishing term). At the end we may safely go to the limit of $\delta\to+0$, to get rid of "complex" time.

[3] It is obvious that the explicit coordinate dependence $\sim\frac{1}{2}q^2$ introduced here is just irrelevant for these argumentation. However, later we shall see its convenience.

an addition of the term $+\frac{1}{2}i\varepsilon q^2$. Then we can write

$$Z[J] = \mathcal{N}\int \mathcal{D}q(t)\exp\left\{\frac{i}{\hbar}\int_{-\infty}^{\infty} dt\left[L(q,\dot{q}) + \hbar Jq + \frac{1}{2}i\varepsilon q^2\right]\right\}, \quad \varepsilon \to +0\,. \tag{9.93}$$

We shall see that the thusly defined functional $Z[J]$ possesses a number of useful and interesting properties.

Instead of transition amplitude $\langle q_f t_f | q_i t_i\rangle$ we may consider the matrix element of coordinate operator $\langle q_f t_f | \hat{q}(t_{n1}) | q_i t_i\rangle$, where $t_f > t_{n1} > t_i$. Using the well-known general rules we can write

$$\begin{aligned}\langle q_f t_f | \hat{q}(t_{n1}) | q_i t_i\rangle = \int dq_1 \cdots dq_n \langle q_f t_f | q_n t_n\rangle \langle q_n t_n | q_{n-1} t_{n-1}\rangle \\ \cdots \langle q_{n1} t_{n1} | \hat{q}(t_{n1}) | q_{n1-1} t_{n1-1}\rangle \cdots . \langle q_1 t_1 | q_i t_i\rangle.\end{aligned} \tag{9.94}$$

Obviously,

$$\langle q_{n1} t_{n1} | \hat{q}(t_{n1}) | q_{n1-1} t_{n1-1}\rangle = q(t_{n1}) \langle q_{n1} t_{n1} | q_{n1-1} t_{n1-1}\rangle\,, \tag{9.95}$$

where in the right-hand side $q(t_{n1})$ is now not an operator, but a c-number (eigenvalue). Then we can repeat all the arguments used during the transformation from (9.15) to (9.24) and write (9.94) in the form of aFeynman path integral:

$$\langle q_f t_f | \hat{q}(t_1) | q_i t_i\rangle = \int \frac{\mathcal{D}q\,\mathcal{D}p}{2\pi\hbar} q(t_1) \exp\left\{\frac{i}{\hbar}\int_{t_i}^{t_f} dt\,[p\dot{q} - H(p,q)]\right\}. \tag{9.96}$$

Now let us calculate the matrix element $\langle q_f t_f | \hat{q}(t_{n1})\hat{q}(t_{n2}) | q_i t_i\rangle$. For $t_{n1} > t_{n2}$ we can write

$$\begin{aligned}\langle q_f t_f | \hat{q}(t_{n1})\hat{q}(t_{n2}) | q_i t_i\rangle = \int dq_1 \cdots dq_n \langle q_f t_f | q_n t_n\rangle \langle q_n t_n | q_{n-1} t_{n-1}\rangle \\ \cdots \langle q_{n1} t_{n1} | \hat{q}(t_{n1}) | q_{n1-1} t_{n1-1}\rangle \\ \cdots \langle q_{n2} t_{n2} | \hat{q}(t_{n2}) | q_{n2-1} t_{n2-1}\rangle \cdots \langle q_1 t_1 | q_i t_i\rangle\,,\end{aligned} \tag{9.97}$$

producing (as limiting behavior) the path integral of the form

$$\langle q_f t_f | \hat{q}(t_1)\hat{q}(t_2) | q_i t_i\rangle = \int \frac{\mathcal{D}q\,\mathcal{D}p}{2\pi\hbar} q(t_1) q(t_2) \exp\left\{\frac{i}{\hbar}\int_{t_i}^{t_f} dt\,[p\dot{q} - H(p,q)]\right\}. \tag{9.98}$$

Here we assumed $t_1 > t_2$. If we consider the case of $t_2 > t_1$, the matrix elements of the coordinate in time moments t_1 and t_2 in the right-hand side of (9.94) will interchange positions, so that this expression, as well as the path integral in the right-hand side of (9.98), reduces to $\langle q_f t_f | \hat{q}(t_2)\hat{q}(t_1) | q_i t_i\rangle$. Thus, in the general case the path integral in the right-hand side of (9.98) defines the matrix element of the *chronological* product of operators $\langle q_f t_f | T[\hat{q}(t_1)\hat{q}(t_2)] | q_i t_i\rangle$, where the operation of the T-ordering of two

operators is defined as

$$T[A(t_1)B(t_2)] = \begin{cases} A(t_1)B(t_2) & \text{for } t_1 > t_2\,, \\ B(t_2)A(t_1) & \text{for } t_2 > t_1\,. \end{cases} \tag{9.99}$$

Thus, in general case we can write

$$\langle q_f t_f | T[\hat{q}(t_1)\hat{q}(t_2)\cdots\hat{q}(t_n)] | q_i t_i \rangle = \int \frac{\mathcal{D}q\,\mathcal{D}p}{2\pi\hbar} q(t_1)q(t_2)\cdots q(t_n) \exp\left\{\frac{i}{\hbar}\int_{t_i}^{t_f} dt\,[p\dot{q} - H(p,q)]\right\}, \tag{9.100}$$

which gives the general expression for the average of the chronological product of the operators via the functional (path) integral. For the case, when the Hamiltonian can be written in the form given by equation (9.17), we can make additional simplifications and write

$$\langle q_f t_f | T[\hat{q}(t_1)\hat{q}(t_2)\cdots q(t_n)] | q_i t_i \rangle = \mathcal{N}\int \mathcal{D}q\, q(t_1)q(t_2)\cdots q(t_n) \exp\left\{\frac{i}{\hbar}\int_{t_i}^{t_f} dt L\right\}. \tag{9.101}$$

Using the definition of the functional $Z[J]$ (9.93) we can easily see that its functional (variational) derivative over the source J is written as

$$\frac{\delta Z[J]}{\delta J(t_1)} = i\,\mathcal{N}\int \mathcal{D}q(t)\, q(t_1) \exp\left\{\frac{i}{\hbar}\int_{-\infty}^{\infty} dt\left[L(q,\dot{q}) + \hbar J q + \frac{1}{2} i\varepsilon q^2\right]\right\}. \tag{9.102}$$

In the general case,

$$\frac{\delta^n Z[J]}{\delta J(t_1)\ldots\delta J(t_n)} = i^n\mathcal{N}\int \mathcal{D}\, q(t)q(t_1) \cdots q(t_n) \exp\left\{\frac{i}{\hbar}\int_{-\infty}^{\infty} dt\left[L(q,\dot{q}) + \hbar J q + \frac{1}{2} i\varepsilon q^2\right]\right\}. \tag{9.103}$$

Now putting $J = 0$ here, we get

$$\left.\frac{\delta^n Z[J]}{\delta J(t_1)\ldots\delta J(t_n)}\right|_{J=0} = i^n\mathcal{N}\int \mathcal{D}\, q(t)q(t_1) \cdots q(t_n) \exp\left\{\frac{i}{\hbar}\int_{-\infty}^{\infty} dt\left[L(q,\dot{q}) + \frac{1}{2} i\varepsilon q^2\right]\right\}. \tag{9.104}$$

Remembering that the term $\frac{i}{2}\varepsilon q^2$ allows us to extract the ground state contribution from the quantum averages, and using (9.101), we come to the following expression

for the "vacuum" average of chronological product of operators:

$$\left.\frac{\delta^n Z[J]}{\delta J(t_1)\ldots\delta J(t_n)}\right|_{J=0} \sim i^n \langle 0,\infty|T[\hat{q}(t_1)\cdots\hat{q}(t_n)]|0,-\infty\rangle\,. \tag{9.105}$$

Thus, the multiple functional differentiation of $Z[J]$ over the source J "generates" the averages of T-ordered products of the quantum operators, while the source itself can be put to zero at the end of calculations. Thus, the functional $Z[J]$ can be called the *generating functional* for these averages. As a byproduct, we obtain the representation of such averages in the form of functional (path) integrals.

We have seen above that vacuum averages of T-ordered products of field operators determine the whole set of Green's functions of quantum field theory. The transition from quantum mechanics to quantum field theory reduces to the generalization to the system with an infinite number of degrees of freedom, when the operators of coordinates are replaced by field operators at each point of space-time. Now it becomes clear that the path integral formulation of quantum mechanics can be used for the direct construction of quantum field theory, based on formalism of functional integrals over field variables. This will be our task in the next chapters.

Chapter 10

Functional integrals: scalars and spinors

10.1 Generating the functional for scalar fields

Now we we will begin discussing the modern functional formulation of quantum field theory. Let us first consider t the simplest case of the free scalar field $\varphi(x)$, interacting with an arbitrary source $J(x)$. Directly generalizing the analysis given in the previous chapter, we can introduce the generating functional:

$$Z[J] = \int \mathcal{D}\varphi(x^\mu) \exp\left\{ i \int d^4x [\mathcal{L}(\varphi) + J(x)\varphi(x) + \frac{i}{2}\varepsilon\varphi^2(x)] \right\}$$
$$\sim \langle 0, \infty | 0, -\infty \rangle^J , \tag{10.1}$$

which is proportional to the vacuum–vacuum transition amplitude. Here $\mathcal{L}(\varphi)$ is a Klein–Gordon Lagrangian, and we replaced integration over the trajectories of a particle to integration over all possible field configurations[1] in space-time: $\mathcal{D}q(t) \to \mathcal{D}\varphi(x^\mu)$. The meaning of such integration is rather simple. We can represent space-time as a set of small four-dimensional cubes (cells) of volume δ^4 and assume our field a constant within any of these cells (the average value of the field inside the cube): $\varphi \approx \varphi(x_i, y_j, z_k, t_l)$. Field derivatives can be expressed via finite differences as

$$\left.\frac{\partial\varphi}{\partial x_i}\right|_{i,j,k,l} \approx \frac{1}{\delta}[\varphi(x_i + \delta, y_j, z_k, t_l) - \varphi(x_i, y_j, z_k, t_l)], \tag{10.2}$$

etc. Replacing the set of indices (i, j, k, l) by a single index n which enumerates the cells (cubes), we can write

$$\mathcal{L}(\varphi_n, \partial_\mu\varphi_n) = \mathcal{L}_n . \tag{10.3}$$

In any of the indices (i, j, k, l) takes N values; the new index n takes N^4 values, and we can write the action as

$$S = \int d^4x \mathcal{L} = \sum_{n=1}^{N^4} \delta^4 \mathcal{L}_n . \tag{10.4}$$

---

[1] In classical field theory we are dealing with only one configuration of the field in space-time, those satisfying the Lagrange equations (principle of the least action). In quantum field theory all kinds of field configurations are "at work", each one entering the theory with the "weight" $\exp\{iS\}$, where S is the classical action.

Then generating functional $Z[J]$ takes the form

$$Z[J] = \lim_{N\to\infty} \int \prod_{n=1}^{N^4} d\varphi_n \exp\left\{ i \sum_{n=1}^{N^4} \delta^4 \left(\mathcal{L}_n + \varphi_n J_n + \frac{i}{2}\varepsilon\varphi_n^2 \right) \right\}, \tag{10.5}$$

which defines the meaning of the formal expression (10.1) and introduces the notion of the functional integral over field configurations (instead of the particle trajectories in quantum mechanics).

Let us calculate $Z[J]$ for the free field, when

$$\mathcal{L} \to \mathcal{L}_0 = \frac{1}{2}(\partial_\mu\varphi\partial^\mu\varphi - m^2\varphi^2) \tag{10.6}$$

is a Klein–Gordon Lagrangian. Then

$$Z_0[J] = \int \mathcal{D}\varphi \exp\left\{ i \int d^4x \left[\frac{1}{2}\left(\partial_\mu\varphi\partial^\mu\varphi - (m^2 - i\varepsilon)\varphi^2\right) + \varphi J \right] \right\}. \tag{10.7}$$

We can move further even without explicit calculation of the functional integral. Let us use the obvious identity $\partial_\mu(\varphi\partial^\mu\varphi) = \partial_\mu\varphi\partial^\mu\varphi + \varphi\partial_\mu\partial^\mu\varphi$ and write

$$\int d^4x \partial_\mu\varphi\partial^\mu\varphi = \int d^4x \partial_\mu(\varphi\partial^\mu\varphi) - \int d^4x \varphi\Box\varphi\,. \tag{10.8}$$

Then the first term in the right-hand side is transformed according to the Gauss theorem into a surface integral, which can be made zero if we move this surface to infinity (where we assume that $\varphi \to 0$). Then,

$$\int d^4x \partial_\mu\varphi\partial^\mu\varphi = -\int d^4x \varphi\Box\varphi \tag{10.9}$$

and generating the functional is rewritten as

$$Z_0[J] = \int \mathcal{D}\varphi \exp\left\{ -i \int d^4x \left[\frac{1}{2}\varphi(\Box + m^2 - i\varepsilon)\varphi - \varphi J \right] \right\}. \tag{10.10}$$

Let us stress that the field φ in this expression is arbitrary (integration variable!) and does not satisfy the Klein–Gordon equation at all! Now we can change the integration variable as

$$\varphi(x) \to \varphi_0(x) + \varphi(x) \tag{10.11}$$

and use the relation (which is derived similarly to (10.9)):

$$\int d^4x \varphi_0[\Box + m^2 - i\varepsilon]\varphi = \int d^4x \varphi(\Box + m^2 - i\varepsilon)\varphi_0\,. \tag{10.12}$$

Then we have

$$\int d^4x\left[\frac{1}{2}\varphi(\Box+m^2-i\varepsilon)\varphi-\varphi J\right]\to\int d^4x\left[\frac{1}{2}\varphi(\Box+m^2-i\varepsilon)\varphi+\right.$$
$$\left.+\varphi(\Box+m^2-i\varepsilon)\varphi_0+\frac{1}{2}\varphi_0(\Box+m^2-i\varepsilon)\varphi_0-\varphi J-\varphi_0 J\right]. \quad (10.13)$$

And now we can require that $\varphi_0(x)$ satisfy the Klein–Gordon equation with the source in the right-hand side:

$$(\Box+m^2-i\varepsilon)\varphi_0(x)=J(x)\,. \quad (10.14)$$

Then the integral of interest to us reduces to

$$\int d^4x\left[\frac{1}{2}\varphi(\Box+m^2-i\varepsilon)\varphi-\frac{1}{2}\varphi_0 J\right]. \quad (10.15)$$

The solution of equation (10.14) has the form

$$\varphi_0(x)=-\int d^4y\Delta_F(x-y)J(y)\,, \quad (10.16)$$

where $\Delta_F(x-y)$ is Feynman's propagator of a scalar field, satisfying the equation (already written in Chapter 4)

$$(\Box+m^2-i\varepsilon)\Delta_F(x)=-\delta(x)\,. \quad (10.17)$$

Substituting (10.16) into (10.15) we see that the expression in the exponent in (10.10) is equal to

$$-i\left\{\frac{1}{2}\int d^4x\varphi(\Box+m^2-i\varepsilon)\varphi+\frac{1}{2}\int d^4xd^4yJ(x)\Delta_F(x-y)J(y)\right\}. \quad (10.18)$$

Thus we obtain[2]

$$Z_0[J]=\exp\left\{-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right\}$$
$$\times\int\mathcal{D}\varphi\exp\left\{-\frac{i}{2}\int dx\varphi(\Box+m^2-i\varepsilon)\varphi\right\}. \quad (10.19)$$

However, the integral over $\mathcal{D}\varphi$ is simply some number (it is taken over all the possible configurations of the field φ)! Denoting this number as $\mathcal{N}$, we obtain finally

$$Z_0[J]=\mathcal{N}\exp\left\{-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right\}. \quad (10.20)$$

The value of $\mathcal{N}$ is of no special importance: this is just a normalization factor.

[2] Further we write for brevity dx instead of d^4x, etc.

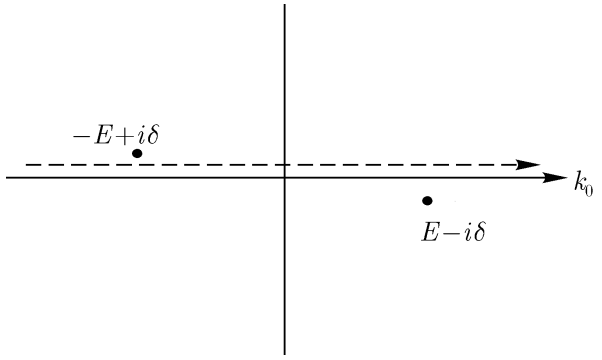


Figure 10.1

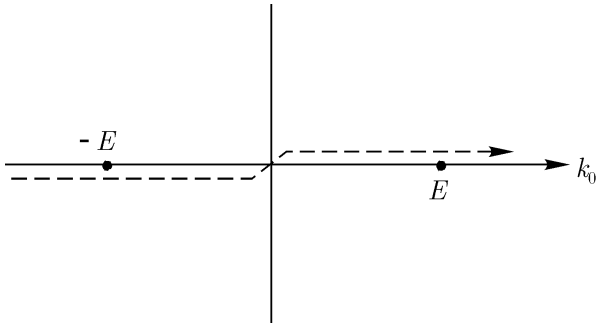


Figure 10.2

The Fourier expansion for $\Delta_F(x)$ has the form

$$\Delta_F(x) = \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ikx}}{k^2 - m^2 + i\varepsilon}\,. \tag{10.21}$$

The presence of $i\varepsilon \to i0^+$ in the denominator here dictates the choice of the integration contour in the integral over k_0 in accordance with Feynman rule to deal with the poles at $k_0 = \pm\sqrt{\mathbf{k}^2 + m^2}$. The poles are situated at the points (in the complex k_0 plane) determined by equation: $k_0^2 = \mathbf{k}^2 + m^2 - i\varepsilon$, i. e., at

$$k_0 = \pm\sqrt{\mathbf{k}^2 + m^2} \mp i\delta = \pm E \mp i\delta\,, \tag{10.22}$$

as shown in Figure 10.1. In the limit of $\delta \to 0(\varepsilon \to 0)$ these poles move to the real axis, and the integration contour goes as shown in Figure 10.2.

We have seen above that this approach corresponds to the "rotation" of the time axis by a small angle δ in the complex plane for time. This guarantees us the correct boundary conditions for the vacuum–vacuum transition amplitude. The same aim can be achieved by making finite angle rotation by $-\pi/2$, so that $t \to -it(\to -i\infty)$.

Introducing the notation

$$x_4 = it = ix_0\,, \tag{10.23}$$

we see that this limit corresponds to $x_4 \to \infty$. Such space-time (with imaginary time) is *Euclidean*, with the invariant interval (distance between two adjacent points) given by

$$ds^2 = -(dx^0)^2 - (dx)^2 - (dy)^2 - (dz)^2 = -\sum_{\mu=1}^{4}(dx^\mu)^2\,. \tag{10.24}$$

In momentum space we can similarly introduce

$$k_4 = -ik_0\,, \tag{10.25}$$

so that

$$k^2 = -(k_1^2 + k_2^2 + k_3^2 + k_4^2) = -k_E^2\,, \quad d^4k_E = d^3\mathbf{k}dk_4 = -id^4k\,, \tag{10.26}$$

where the index E denotes Euclidean momentum space. Now the Feynman propagator takes the form

$$\Delta_F(x) = -i\int \frac{d^4k_E}{(2\pi)^4}\frac{e^{-ikx}}{k_E^2 + m^2}\,. \tag{10.27}$$

Note that this expression[3], up to a factor of $-i$, coincides with the Ornstein–Zernike correlation function of the theory of critical phenomena in four-dimensional space [3, 36, 42], if we take $m^2 \sim T - T_c$, where T_c is the temperature of the second order phase transition (and for simplicity we are dealing with temperatures $T > T_c$). Here we meet for the first time the deep interconnections between quantum field theory and the modern theory of critical phenomena in statistical physics [3, 42]. From equation (10.7), taking into account $d^4x = -id^4x_E$ and $(\partial_\mu\varphi)^2 = -(\partial_E^\mu\varphi)^2$, we obtain the generating functional of Euclidean field theory as

$$Z_{0E}[J] = \int \mathcal{D}\varphi \exp\left\{-\int d^4x_E\left(\frac{1}{2}[(\partial_E^\mu\varphi)^2 + m^2\varphi^2] - \varphi J\right)\right\}\,, \tag{10.28}$$

which actually coincides with the partition function of the *Gaussian* model of the phase transition (i. e., the Landau theory [36] with no $\sim \varphi^4$ and higher-order terms in Landau expansion) for scalar-order parameter φ, interacting with the external field J [3, 42].

[3] There is no problem with encircling poles here: both are at the imaginary axis at points $k_4 = \pm i\sqrt{k_E^2 + m^2}$.

10.2 Functional integration

Now we present the formal discussion of functional integration. Let us start from the well-known expression for the Poisson–Gauss integral (9.28):

$$\int_{-\infty}^{\infty} dx e^{-\frac{1}{2}ax^2} = \sqrt{\frac{2\pi}{a}}\,. \tag{10.29}$$

In the following we assume the integration limits to be always from $-\infty$ to ∞ and do not write them explicitly. Let us take the product of n such integrals:

$$\int dx_1 dx_2 \dots dx_n \exp\left(-\frac{1}{2}\sum_n a_n x_n^2\right) = \frac{(2\pi)^{n/2}}{\prod_{i=1}^n a_i^{1/2}}\,. \tag{10.30}$$

Let A be the diagonal matrix with elements $a_1, a_2, \dots, a_n$ and x the n-dimensional vector (column) with components $x_1, x_2, \dots, x_n$. Then the expression in the exponent in (10.30) can be written as the scalar product:

$$(x, Ax) = \sum_n a_n x_n^2\,, \tag{10.31}$$

The determinant of the matrix A is

$$\mathrm{Det}\, A = a_1 a_2 \cdots a_n = \prod_{i=1}^n a_i\,. \tag{10.32}$$

Then (10.30) can be written as

$$\int d^n x e^{-\frac{1}{2}(x,Ax)} = (2\pi)^{n/2}(\mathrm{Det}\, A)^{-1/2}\,. \tag{10.33}$$

This expression is valid for any diagonal matrix; correspondingly it is also valid for any real symmetric matrix, as it can always be diagonalized by linear transformation. Let us define the integration measure as

$$[dx] = (2\pi)^{-n/2} d^n x\,. \tag{10.34}$$

Then (10.33) is rewritten as

$$\int [dx] e^{-\frac{1}{2}(x,Ax)} = (\mathrm{Det}\, A)^{-1/2}\,. \tag{10.35}$$

This relation is easily generalized to the case when there is a general quadratic form in the exponent:

$$Q(x) = \frac{1}{2}(x, Ax) + (b, x) + c\,. \tag{10.36}$$

We can proceed as during the derivation of (9.35). The form (10.36) reaches its minimum for $\bar{x} = -A^{-1}b$ and can be rewritten as

$$Q(x) = Q(\bar{x}) + \frac{1}{2}[x - \bar{x}, A(x - \bar{x})]\,. \tag{10.37}$$

Then we immediately obtain the analogue of (9.35) as

$$\int [dx] \exp\left[-\frac{1}{2}(x, Ax) - (b, x) - c\right] = \exp\left[\frac{1}{2}(b, A^{-1}b) - c\right] (\mathrm{Det}\, A)^{-1/2}, \quad (10.38)$$

where A^{-1} denotes the inverse matrix.

Consider now the case of Hermitian matrices. Taking the square of (10.29) we write

$$\int dx dy e^{-\frac{1}{2}a(x^2+y^2)} = \frac{2\pi}{a}. \quad (10.39)$$

Let us introduce $z = x + iy$ and $z^* = x - iy$, so that (calculating the Jacobian of transformation from x, y to z, z^*) we have $dxdy = -i dz^* dz/2$, so that (10.39) can be written as

$$\int \frac{dz^*}{(2\pi i)^{1/2}} \frac{dz}{(2\pi i)^{1/2}} e^{-az^* z} = \frac{1}{a} \quad (10.40)$$

We can generalize this expression, similarly to the transformation from (10.30) to (10.35) and (10.38), introducing the positive definite Hermitian matrix A, the complex vector b, and the integration measure

$$[dz] = (2\pi i)^{-n/2} d^n z. \quad (10.41)$$

Then we obtain

$$\int [dz^*][dz] e^{-(z^*, Az)} = (\mathrm{Det}\, A)^{-1}. \quad (10.42a)$$

$$\int [dz^*][dz] e^{-(z^*, Az) - (b^*, z) - (z^*, b) - c} = \exp\left[(b^*, A^{-1}b) - c\right] (\mathrm{Det}\, A)^{-1} \quad (10.42b)$$

All these expressions are quite rigorous and represent the direct generalization of "one-dimensional" integrals to the case of the vector space of finite dimensionality. Let us make the formal generalization to the case of infinite-dimensional functional space. Consider the space of real functions $\varphi(x^\mu)$. We can define the scalar product as

$$(\varphi, \varphi) = \int d^4x [\varphi(x)]^2. \quad (10.43)$$

The generalization of equation (10.35) is written as

$$\int \mathcal{D}\varphi(x) \exp\left\{-\frac{1}{2} \int dx \varphi(x) A \varphi(x)\right\} = (\mathrm{Det}\, A)^{-1/2}, \quad (10.44)$$

where A is some operator acting upon functions $\varphi(x)$:

$$A\varphi(x) = \int dy\, A(x, y) \varphi(y), \quad (10.45)$$

and its determinant is naturally defined as the the corresponding product of eigenvalues. Integration measure is $\mathcal{D}\varphi(x) = [d\varphi(x)]$. All these expressions should be understood as the limiting expression like (10.5). Expression (10.44) is usually called a Gaussian functional integral.

If $\varphi(x)$ is a complex function (field), we obtain the natural generalization of (10.42a)

$$\int \mathcal{D}\varphi^*(x)\mathcal{D}\varphi(x)\exp\left\{-\int dx\varphi^*(x)A\varphi(x)\right\} = (\operatorname{Det} A)^{-1}, \tag{10.46}$$

where A is the Hermitian operator.

Generalization of (10.38) for the case of real fields $\varphi(x)$ has the form

$$\int \mathcal{D}\varphi(x)\exp\left\{-\frac{1}{2}\int dx\int dy\varphi(x)A(x,y)\varphi(y) - \int dxB(x)\varphi(x) - c\right\} = \exp\left\{\frac{1}{2}\int dx\int dyB(x)A^{-1}(x,y)B(y) - c\right\}(\operatorname{Det} A)^{-1/2} \tag{10.47}$$

where $A^{-1}(x,y)$ denotes the inverse operator. A similar expression for integration over complex fields differs from (10.47) by the presence of integration over φ^* and φ, as well by the replacement of $(\operatorname{Det} A)^{-1/2}$ by $(\operatorname{Det} A)^{-1}$:

$$\int \mathcal{D}\varphi^*(x)\mathcal{D}\varphi(x)\exp\left\{-\int dx\int dy\varphi^*(x)A(x,y)\varphi(y) - \int dx\big[B^*(x)\varphi(x) + \varphi^*(x)B(x)\big] - c\right\} = \exp\left\{\int dx\int dyB^*(x)A^{-1}(x,y)B(y) - c\right\}(\operatorname{Det} A)^{-1}. \tag{10.48}$$

Let us return to the discussion of the general expression for generating the functional of a Klein–Gordon field (10.10):

$$Z_0[J] = \int \mathcal{D}\varphi\exp\left\{-i\int d^4x\left[\frac{1}{2}\varphi(\Box + m^2 - i\varepsilon)\varphi - \varphi J\right]\right\}. \tag{10.49}$$

Here we have precisely the Gaussian functional integral like (10.47) with $A(x,y) = i(\Box + m^2 - i\varepsilon)\delta(x-y)$, $B(x) = -iJ(x)$, $c = 0$. Then, from (10.47) we get

$$Z_0[J] = \exp\left\{\frac{i}{2}\int dxdyJ(x)(\Box + m^2 - i\varepsilon)^{-1}J(y)\right\}[i\operatorname{Det}(\Box + m^2 - i\varepsilon)]^{-1/2}. \tag{10.50}$$

The determinant here can be rewritten using (10.44) as

$$[i\operatorname{Det}(\Box + m^2 - i\varepsilon)]^{-1/2} = \int \mathcal{D}\varphi(x)\exp\left\{-\frac{i}{2}\int dx\varphi(x)(\Box + m^2 - i\varepsilon)\varphi(x)\right\}, \tag{10.51}$$

and the inverse operator is

$$(\Box + m^2 - i\varepsilon)^{-1} = -\Delta_F(x-y)\,, \tag{10.52}$$

which follows directly from (10.17). Then equation (10.50) reduces to

$$\begin{aligned} Z_0[J] = \exp\left\{-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right\} \\ \times \int \mathcal{D}\varphi \exp\left\{-\frac{i}{2}\int dx\varphi(\Box + m^2 - i\varepsilon)\varphi\right\}, \end{aligned} \tag{10.53}$$

which coincides with (10.19). Thus, the direct calculation using the rules of functional integration produces the same result obtained above via "indirect" arguments.

The expressions for Gaussian functional integrals will be widely used below.

10.3 Free particle Green's functions

Let us now show that $Z_0[J]$ is the generating functional for the Green's functions of free particles. We can expand (10.20) in the series:

$$\begin{aligned} Z_0[J] = \mathcal{N}\left\{1 - \frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right. \\ + \frac{1}{2!}\left(\frac{i}{2}\right)^2\left[\int dxdyJ(x)\Delta_F(x-y)J(y)\right]^2 \\ \left. -\frac{1}{3!}\left(\frac{i}{2}\right)^3\left[\int dxdyJ(x)\Delta_F(x-y)J(y)\right]^3 + \cdots\right\}. \end{aligned} \tag{10.54}$$

Introducing Fourier representation for the source

$$J(x) = \int d^4pJ(p)e^{-ipx} \tag{10.55}$$

and using (10.21) we easily obtain

$$-\frac{i}{2}\int d^4xd^4yJ(x)\Delta_F(x-y)J(y) = -\frac{i}{2}(2\pi)^4\int d^4p\frac{J(-p)J(p)}{p^2 - m^2 + i\varepsilon}\,. \tag{10.56}$$

We may associate analytic expressions in these series with graphic elements as shown in Figure 10.3. Then equation (10.56) corresponds to the diagram, shown in Figure 10.4. As a result, the expansion of the generating functional (vacuum–vacuum transition amplitude) (10.54) is represented by the diagrams shown in Figure 10.5[4]. We see that this series describes the propagation of 1, 2, 3, etc. "particles" between sources, so that we are dealing with a many particle theory. It is clear that $Z_0[J]$ is the generating functional for the Green's functions of our field theory.

[4] Normalization factor $\mathcal{N}$ is dropped here.

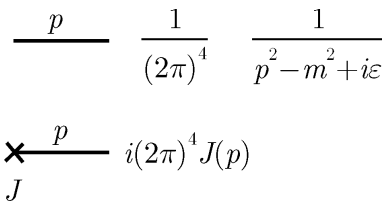


Figure 10.3

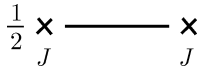


Figure 10.4

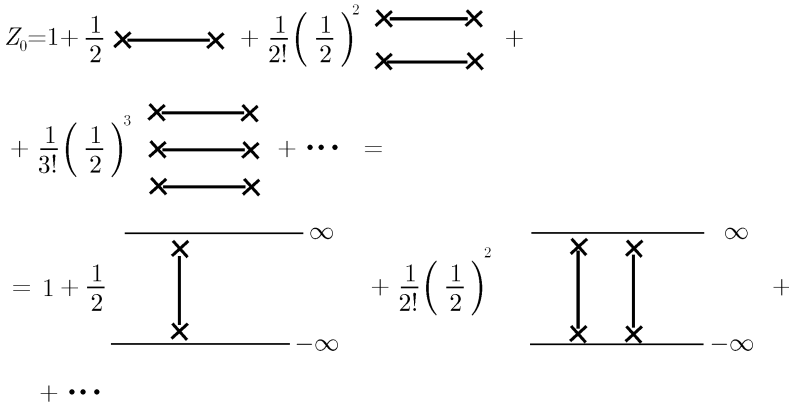


Figure 10.5

Let us comment on the formal aspects of this analysis. Consider, for example, Taylor expansion of some functions $F(y_1,\dots,y_k)$ of k variables $y_1,\dots,y_k$:

$$F\{y\} \equiv F(y_1,\dots,y_k) = \sum_{n=0}^{\infty}\sum_{i_1=1}^{k}\cdots\sum_{i_n=1}^{k}\frac{1}{n!}T_n(i_1,\dots,i_n)y_{i_1}\dots y_{i_n}\,, \tag{10.57}$$

where

$$T_n = \left.\frac{\partial^n F\{y\}}{\partial y_{i_1}\dots\partial y_{i_n}}\right|_{y=0}. \tag{10.58}$$

We may go to the intuitively clear limit, when variables form the continuum $i \to x, y_i (i = 1,\dots,k) \to y(x), \sum_i \to \int dx$ and obtain an expansion for the *functional*

$$F[y] = \sum_{n=0}^{\infty}\int dx_1\dots dx_n\frac{1}{n!}T_n(x_1,\dots,x_n)y(x_1)\dots y(x_n)\,, \tag{10.59}$$

where

$$T_n(x_1,\dots,x_n) = \left.\frac{\delta}{\delta y(x_1)}\cdots\frac{\delta}{\delta y(x_n)}F[y]\right|_{y=0}. \tag{10.60}$$

In such a case we call $F[y]$ the generating functional for functions $T_n(x_1,\dots x_n)$.

Our generating functional $Z[J]$ should be normalized. We have seen that it is, proportional to vacuum-vacuum transition amplitude in the presence of a source J Natural normalization is $Z[J=0]=1$. Then we can write

$$Z[J] = \langle 0, \infty | 0, -\infty \rangle^J , \tag{10.61}$$

so that $Z[0]=1$ is satisfied automatically. Thus, we have to rewrite both (10.10) and (10.20) as

$$Z_0[J] = \frac{\int \mathcal{D}\varphi \exp\left\{-i \int d^4x \left[\frac{1}{2}\varphi(\Box + m^2 - i\varepsilon)\varphi - \varphi J\right]\right\}}{\int \mathcal{D}\varphi \exp\left\{-i \int d^4x \frac{1}{2}\varphi(\Box + m^2 - i\varepsilon)\varphi\right\}}, \tag{10.62}$$

$$Z_0[J] = \exp\left\{-\frac{i}{2}\int dx dy J(x)\Delta_F(x-y)J(y)\right\} . \tag{10.63}$$

These new definitions obviously satisfy the condition of $Z[J=0]=1$, justifying our dropping of the irrelevant normalization factor $\mathcal{N}$. The functional $Z_0[J]$, defined by equation (10.63) in accordance with equation (10.60), is the generating functional of the functions

$$\tau(x_1, \ldots, x_n) = \frac{1}{i^n} \frac{\delta^n Z_0[J]}{\delta J(x_1) \ldots \delta J(x_n)}\bigg|_{J=0} . \tag{10.64}$$

Recalling equation (9.105), we understand that

$$\frac{\delta^n Z_0[J]}{\delta J(x_1) \ldots \delta J(x_n)}\bigg|_{J=0} = i^n \langle 0 | T\varphi(x_1) \ldots \varphi(x_n) | 0 \rangle , \tag{10.65}$$

so that

$$\tau(x_1, \ldots, x_n) = \langle 0 | T\varphi(x_1) \cdots \varphi(x_n) | 0 \rangle \tag{10.66}$$

represents the vacuum average of the chronological product of the field *operators*, i. e., n-point (number of coordinates) Green's functions of our theory. This definition coincides with our previous definition of the Green's functions in the operator formalism of quantum field theory. Generating the functional can now be written as

$$Z_0[J] = \sum_{n=0}^{\infty} \frac{i^n}{n!} \int dx_1 \ldots dx_n J(x_1) \ldots J(x_n) \tau(x_1, \ldots, x_n) , \tag{10.67}$$

which means that $Z_0[J]$ is the generating functional of the Green's functions $\tau(x_1, \ldots, x_n)$. This expansion is shown graphically in Figure 10.5.

Let us now calculate some simplest n-point Green's functions in our free scalar field theory. We start with a 2-point function

$$\tau(x, y) = -\frac{\delta^2 Z_0[J]}{\delta J(x) \delta J(y)}\bigg|_{J=0} . \tag{10.68}$$

Calculations can be done explicitly, using the general definition of the functional derivative. We have

$$\frac{1}{i}\frac{\delta Z_0[J]}{\delta J(x)} = \frac{1}{i}\frac{\delta}{\delta J(x)}\exp\left[-\frac{i}{2}\int dx_1 dx_2 J(x_1)\Delta_F(x_1-x_2)J(x_2)\right]$$
$$= -\int dx_1 \Delta_F(x-x_1)J(x_1)\exp\left[-\frac{i}{2}\int dx_1 dx_2 J(x_1)\Delta_F(x_1-x_2)J(x_2)\right] \tag{10.69}$$

$$\frac{1}{i}\frac{\delta}{\delta J(x)}\frac{1}{i}\frac{\delta}{\delta J(y)}Z_0[J] = i\Delta_F(x-y)\exp\left(-\frac{i}{2}\int J\Delta_F J\right)$$
$$+\int dx_1 \Delta_F(x-x_1)J(x_1)\int dx_2 \Delta_F(y-x_2)J(x_2)\exp\left(-\frac{i}{2}\int J\Delta_F J\right) \tag{10.70}$$

where we use the shortened notations in the exponent. Putting now $J = 0$ we get

$$\frac{1}{i}\frac{\delta}{\delta J(x)}\frac{1}{i}\frac{\delta}{\delta J(y)}Z_0[J]\bigg|_{J=0} = i\Delta_F(x-y) \tag{10.71}$$

or

$$\tau(x,y) = i\Delta_F(x-y)\,. \tag{10.72}$$

It is clear that the 2-point Green's function in fact coincides with the Feynman propagator for a scalar particle (single-particle Green's function of a free scalar particle). Let us once more consider its physical meaning. We start with operator formalism. By definition of the chronological product we have

$$\tau(x,y) = \langle 0|T\varphi(x)\varphi(y)|0\rangle$$
$$= \theta(x_0-y_0)\langle\varphi(x)\varphi(y)\rangle + \theta(y_0-x_0)\langle\varphi(y)\varphi(x)|0\rangle\,. \tag{10.73}$$

Here the first term represents the probability amplitude of particle creation at point y at time moment y_0, and its annihilation at point x at later moment x_0. The second term gives the probability amplitude of particle creation at point x at moment x_0 and its annihilation at point y at time moment y_0. These processes are graphically illustrated in Figure 10.6. The sum of these amplitudes gives the Feynman propagator. We know that in operator formalism the field φ can be written as the sum of the terms with positive and negative frequencies (cf. Chapter 3)

$$\varphi(x) = \varphi^{(+)}(x) + \varphi^{(-)}(x)\,, \tag{10.74}$$

where

$$\varphi^{(+)}(x) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_\mathbf{k}}}a_\mathbf{k}e^{-ikx}\,, \tag{10.75}$$

$$\varphi^{(-)}(x) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{\sqrt{2\omega_\mathbf{k}}}a^+_\mathbf{k}e^{ikx}\,, \tag{10.76}$$

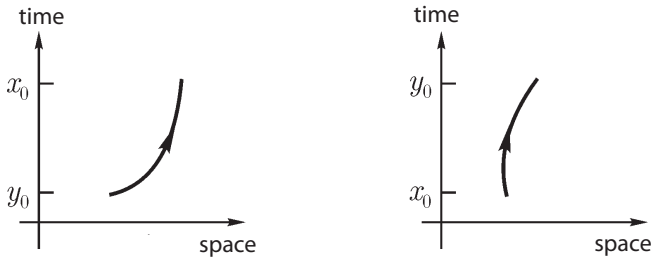


Figure 10.6

where $\omega_{\mathbf{k}} = \sqrt{\mathbf{k}^2 + m^2}$ and $a^+_{\mathbf{k}}$, $a_{\mathbf{k}}$ are the corresponding creation and annihilation operators. Taking into account the physical meaning of these operators, only terms like $\varphi^{(+)}\varphi^{(-)}$ remain in the vacuum average (10.73):

$$\tau(x,y) = \theta(x_0 - y_0)\langle 0|\varphi^{(+)}(x)\varphi^{(-)}(y)|0\rangle + \theta(y_0 - x_0)\langle 0|\varphi^{(+)}(y)\varphi^{(-)}(x)|0\rangle\,. \tag{10.77}$$

Substituting here (10.75) into (10.76) we get

$$\tau(x,y) = \int \frac{d^3\mathbf{k}d^3\mathbf{k}'}{(2\pi)^6\sqrt{2\omega_{\mathbf{k}}\omega_{\mathbf{k}'}}}[\theta(x_0 - y_0)e^{-i(kx-k'y)} + \theta(y_0 - x_0)e^{-i(ky-k'x)}]\langle 0|a_{\mathbf{k}}a^+_{\mathbf{k}'}|0\rangle\,, \tag{10.78}$$

so that interchanging operators in the vacuum average using commutation relations (to separate normal product of operators giving zero and nonzero contribution from δ-function) we obtain

$$\tau(x,y) = \int \frac{d^3\mathbf{k}}{(2\pi)^3 2\omega_{\mathbf{k}}}[\theta(x_0 - y_0)e^{-ik(x-y)} + \theta(y_0 - x_0)e^{ik(x-y)}]\,. \tag{10.79}$$

Actually, this expression can be shown to coincide with $i\,\Delta_F(x-y)$, where $\Delta_F(x-y)$ is given by (10.21). and equation (10.21) can be rewritten as

$$\begin{aligned}\Delta_F(x) &= \int \frac{d^4k}{(2\pi)^4}\frac{e^{-ikx}}{k^2 - m^2 + i\varepsilon} = \int \frac{d^3\mathbf{k}dk_0}{(2\pi)^4}\frac{e^{-ikx}}{k_0^2 - (\mathbf{k}^2 + m^2) + i\varepsilon} \\ &= \int \frac{d^3\mathbf{k}dk_0}{(2\pi)^4}\frac{e^{-ikx}}{2\omega_{\mathbf{k}}}\left\{\frac{1}{k_0 - \omega_{\mathbf{k}} + i\delta} - \frac{1}{k_0 + \omega_{\mathbf{k}} - i\delta}\right\}.\end{aligned} \tag{10.80}$$

The integral over k_0 can be calculated as usual by contour integration in a complex plane. Due to the exponential factor $e^{-ik_0x_0}$, for $x_0 > 0$ we are closing the integration contour in the lower half-plane of k_0, so that the integral is determined by the contribution of the pole at $k_0 = \omega_{\mathbf{k}} - i\delta$. For $x_0 < 0$ we close the integration contour in the upper half-plane, so that integral is determined by the pole at $k_0 = -\omega_{\mathbf{k}} + i\delta$. Then,

using the Cauchy theorem, we have

$$\Delta_F(x) = \int \frac{d^2\mathbf{k}}{(2\pi)^3}\frac{e^{i\mathbf{kx}}}{2\omega_\mathbf{k}}\big[\theta(x_0)(-i)e^{-i\omega_\mathbf{k}x_0} - \theta(-x_0)i\,e^{i\omega_\mathbf{k}x_0}\big]. \tag{10.81}$$

After the replacement $\mathbf{k} \to -\mathbf{k}$ in the second integral and variable transformation $x \to x - y$, we obtain

$$\Delta_F(x-y) = -i\int \frac{d^3\mathbf{k}}{(2\pi)^3 2\omega_\mathbf{k}}\big[\theta(x_0-y_0)e^{-ik(x-y)}+\theta(y_0-x_0)e^{ik(x-y)}\big], \tag{10.82}$$

which coincides with $-i\,\tau(x, y)$ from (10.79). Thus, the 2-point Green's function appearing in the functional approach coincides with single-particle propagator of the operator formulation of quantum field theory.

But what is the 1-point function? From (10.69) we obviously have

$$\begin{aligned}\tau(x) = \langle 0|T\varphi(x)|0\rangle = \langle 0|\varphi(x)|0\rangle &= \frac{1}{i}\frac{\delta Z_0[J]}{\delta J(x)}\bigg|_{J=0}\\ &= -\int dx_1 \Delta_F(x-x_1)J(x_1)\exp\left(-\frac{i}{2}\int J\Delta_F J\right)\bigg|_{J=0} = 0,\end{aligned} \tag{10.83}$$

i. e., the vacuum average of the field itself is just zero!

Let us now find the 3-point function. Differentiating (10.70) once more, we get

$$\begin{aligned}&\frac{1}{i}\frac{\delta}{\delta J(x_1)}\frac{1}{i}\frac{\delta}{\delta J(x_2)}\frac{1}{i}\frac{\delta}{\delta J(x_3)}Z_0[J] =\\ &\quad -i\,\Delta_F(x_2-x_3)\int dx\Delta_F(x_1-x)J(x)\exp\left(-\frac{i}{2}\int J\Delta_F J\right)\\ &\quad -i\,\Delta_F(x_2-x_1)\int dx\Delta_F(x_3-x)J(x)\exp\left(-\frac{i}{2}\int J\Delta_F J\right)\\ &\quad -i\,\Delta_F(x_3-x_1)\int dx\Delta_F(x_2-x)J(x)\exp\left(-\frac{i}{2}\int J\Delta_F J\right)\\ &\quad -\int dx\Delta_F(x_2-x)J(x)\int dy\Delta_F(x_3-y)J(y)\\ &\quad \times\int dz\Delta_F(x_1-z)J(z)\exp\left(-\frac{i}{2}\int J\Delta_F J\right),\end{aligned} \tag{10.84}$$

which for $J = 0$ obviously gives zero. Thus

$$\tau(x_1, x_2, x_3) = \langle 0|T\varphi(x_1)\varphi(x_2)\varphi(x_3)|0\rangle = 0\,. \tag{10.85}$$

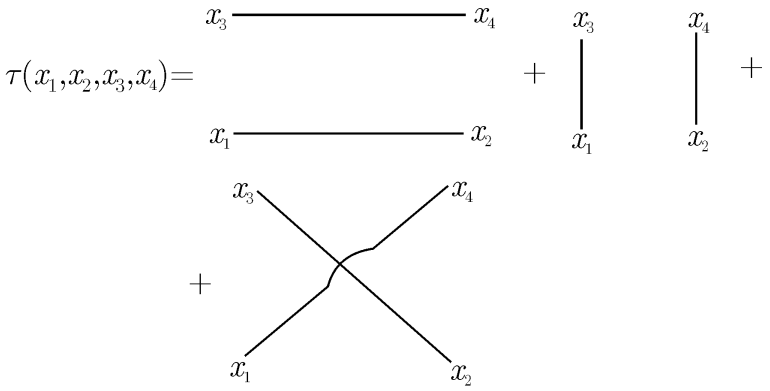


Figure 10.7

Similar calculations give

$$\frac{1}{i}\frac{\delta}{\delta J(x_1)}\cdots\frac{1}{i}\frac{\delta}{\delta J(x_4)}Z_0[J] = -\Delta_F(x_2-x_3)\Delta_F(x_1-x_4)\exp\left(-\frac{i}{2}\int J\Delta_F J\right)$$
$$-\Delta_F(x_2-x_1)\Delta_F(x_3-x_4)\exp\left(-\frac{i}{2}\int J\Delta_F J\right)$$
$$-\Delta_F(x_3-x_1)\Delta_F(x_2-x_4)\exp\left(-\frac{i}{2}\int J\Delta_F J\right)$$
$$+\cdots, \tag{10.86}$$

where multiple dots denote terms giving zero for $J = 0$. Accordingly, we obtain

$$\tau(x_1,x_2,x_3,x_4) = -\Delta_F(x_2-x_3)\Delta_F(x_1-x_4) - \Delta_F(x_2-x_1)\Delta_F(x_3-x_4)$$
$$-\Delta_F(x_3-x_1)\Delta_F(x_2-x_4), \tag{10.87}$$

which is graphically shown by the diagrams in Figure 10.7 and represents the propagation amplitude of two free particles. Here we have just four space-time points interconnected in all possible ways by the lines of free particles.

Going to n-point functions, we can easily be convinced that for the odd value of n they are just zero:

$$\tau(x_1,x_2,\ldots,x_{2n+1}) = 0. \tag{10.88}$$

For even n each n-point function is factorized into the sum of the products of the 2-point functions (i. e., the sum of all "pairings" (contractions), defined by all possible permutations of the coordinates, entering in pairs):

$$\tau(x_1,x_2,\ldots,x_{2n}) = \sum_P \tau(x_{p_1},x_{p_2})\cdots\tau(x_{p_{2k-1}},x_{p_{2k}}), \tag{10.89}$$

where

$$\tau(x_1, x_2) = i\Delta_F(x - y)\,. \tag{10.90}$$

This reduces to the Wick theorem we are familiar with, which is now proved in a functional formulation of quantum field theory.

10.4 Generating the functional for interacting fields

So far we have discussed the case of a free (noninteracting) field. How can this formalism be generalized to the interacting case? Consider the simplest case of interacting theory, taking the Lagrangian of a scalar field in the form

$$\mathcal{L} = \frac{1}{2}\partial^\mu\varphi\partial_\mu\varphi - \frac{m^2}{2}\varphi^2 - \frac{g}{4!}\varphi^4 = \mathcal{L}_0 + \mathcal{L}_{int}\,, \tag{10.91}$$

where g is some coupling constant. This is the so-called $g\varphi^4$-theory. The interaction Lagrangian is

$$\mathcal{L}_{int} = -\frac{g}{4!}\varphi^4\,. \tag{10.92}$$

Lagrange equations for such theory are nonlinear (containing the term $\sim g\varphi^3$), which reflects the presence of (self)interaction. In the general case the interaction Lagrangian is some function $V(\varphi)$. In principle, we could consider even nonpolynomial functions, but we shall limit ourselves here to the simplest models.

Remarks on the dimensionality of coupling constants

We have seen above that the action $S = \int d^4x\mathcal{L}$ is dimensionless (we are using natural system units with $\hbar = 1$). Correspondingly, the dimensionality of the Lagrangian $[\mathcal{L}] = l^{-4}$, where l is some length. The dimensionality of energy (mass) is: $[E] = [m] = l^{-1}$. From the explicit expression (10.91) for the Lagrangian it is clear that $[\varphi] = l^{-1}$. Then from equation (10.92) it is clear that in $g\varphi^4$-theory the interaction constant g is *dimensionless*. This is very important! Due to this property, this theory is renormalizable. Intuitively, this can be understood from the following elementary arguments. Consider a more general interaction Lagrangian:

$$\mathcal{L}_{int} = g_k\varphi^{4+k} \quad k > 0 \tag{10.93}$$

In this case the dimensionality of the coupling constant $[g_k] = l^k$. However, perturbation expansion should be always performed in powers of some dimensionless small parameter. In our case, such a parameter is given by

$$g_k l^{-k} \sim g_k m^k \sim g_k E^k \tag{10.94}$$

which *grows* with the growth of energy E (or at small distances). This is bad and actually reflect the nonrenormalizability of such a theory. Roughly speaking, we may say that the dimensionality of the coupling constant is a necessary (but not sufficient!) condition for the renormalizability of any theory of interacting particles. More precisely, it is necessary that interaction constant be dimensionless, or that it has the dimensionality of some negative power

of length: $g \sim l^{-a}$, $a > 0$. In this last case, the dimensionless parameter of perturbation theory is gE^{-a}, which is harmless at high energies. From this point of view, $g\varphi^3$-theory is also satisfactory, but it leads to other problems: it breaks the positive definiteness of energy (there is no stable ground state). Thus, the $g\varphi^4$-theory is actually the only "reasonable" theory of a scalar field in 4-dimensional space-time[5].

For the spinor field ($s = 1/2$) we argue in a similar way. Dirac's Lagrangian $\mathcal{L} \sim i\bar{\psi}\partial\psi - m\bar{\psi}\psi$, so that $[\psi] = [\bar{\psi}] = l^{-3/2}$. Correspondingly, if we write the interaction Lagrangian of Dirac field with scalar fields in the obvious form (so-called Yukawa interaction)

$$\mathcal{L}_{int} \sim g\bar{\psi}\psi\varphi\,, \tag{10.95}$$

the appropriate interaction constant g is again dimensionless, and the theory is renormalizable. However, if we take the 4-Fermion interaction (Fermi)

$$\mathcal{L}_{int} \sim G\bar{\psi}\psi\bar{\psi}\psi\,, \tag{10.96}$$

the coupling constant G is dimensional: $[G] = [m^{-2}] = l^2$. Such a theory has "bad" behavior at high energies and is nonrenormalizable.

Modern quantum field theory deals mainly with renormalizable theories. The dimensionality of the coupling constant is the crude criterion for choosing between different interaction Lagrangians, satisfying the general requirements of relativistic invariance.

Tje normalized generating functional for the theory with interactions is defined similarly to the case of noninteracting theory (cf.(10.1), (10.62)):

$$Z[J] = \frac{\int \mathcal{D}\varphi \exp\left(iS + i\int dx J\varphi\right)}{\int \mathcal{D}\varphi e^{iS}}\,, \tag{10.97}$$

where $S = \int d^4x\mathcal{L}$ is the action of our theory, including the contribution from the interaction Lagrangian. For $\mathcal{L}_{int} = 0$, (10.97) naturally reduces to the case of the free theory. In the general case we can write $S = S_0 + S_{int}$, where $S_{int} = \int d^4x\mathcal{L}_{int}$.

Performing functional differentiation explicitly, we have

$$\frac{1}{i}\frac{\delta Z}{\delta J(x)} = \frac{\int \mathcal{D}\varphi \exp\left(iS + i\int dx J\varphi\right)\varphi(x)}{\int \mathcal{D}\varphi e^{iS}}\,, \tag{10.98}$$

$$\frac{1}{i^2}\frac{\delta^2 Z}{\delta J(x)\delta J(y)} = \frac{\int \mathcal{D}\varphi \exp\left(iS + i\int dx J\varphi\right)\varphi(x)\varphi(y)}{\int \mathcal{D}\varphi e^{iS}}, \tag{10.99}$$

etc. Putting here $J = 0$ we generate all the Green's functions of our theory:

$$\langle 0|T\varphi(x)\varphi(y)|0\rangle = \frac{\int \mathcal{D}\varphi \exp(iS)\,\varphi(x)\varphi(y)}{\int \mathcal{D}\varphi e^{iS}}\,, \tag{10.100}$$

$$\langle 0|T\varphi(x_1)\varphi(x_2)\varphi(x_3)\varphi(x_4)|0\rangle = \frac{\int \mathcal{D}\varphi \exp(iS)\,\varphi(x_1)\varphi(x_2)\varphi(x_3)\varphi(x_4)}{\int \mathcal{D}\varphi e^{iS}}, \tag{10.101}$$

[5] These arguments, including the dependence on spatial dimensionality, will be discussed later in more detail.

etc. We see that Green's functions are represented by the functional "averages" of the products of an even number of fields, and "averaging" is performed with "weight" e^{iS}. If we write here $S = S_0 + S_{int}$ and perform an expansion of the exponent in powers of S_{int} (i. e., consider the perturbation series in powers of the coupling constant) and use Wick theorem (proven above), we can build a diagram technique for calculating arbitrary Green's functions, similarly to the case of operator formalism. The "averages" of the pairs of fields in different points will be "averaged" with e^{iS_0}. These "averages" are easily calculated (Gaussian integrals!) and reduced to the appropriate free Green's functions. However, below we shall use a more formal approach based on the analysis of the general relations for generating the functional of interacting theory (10.97).

Note that

$$\frac{1}{i}\frac{\delta}{\delta J(x)} e^{i\int dx J\varphi} = \varphi(x) e^{i\int dx J\varphi}. \tag{10.102}$$

As J and φ here are independent (functional) variables, a similar equality also is valid for an arbitrary function of φ:

$$V\left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right) e^{i\int dx J\varphi} = V\left(\varphi(x)\right) e^{i\int dx J\varphi}, \tag{10.103}$$

which is easily proven making Taylor expansion of $V(\varphi)$. Then we have

$$e^{-i\int dx V(\varphi)} e^{i\int dx J\varphi} = e^{-i\int dx V\left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)} e^{i\int dx J\varphi}. \tag{10.104}$$

Now, taking for $V(\varphi)$ the interaction Lagrangian $\mathcal{L}_{int}(\varphi)$, we can write the generating functional of interacting theory as

$$\begin{aligned} Z[J] &= \mathcal{N}\int \mathcal{D}\varphi \, \exp\left\{ i\int dx \left[\frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}(m^2 - i\varepsilon)\varphi^2 + \mathcal{L}_{int}(\varphi) + J\varphi\right]\right\} \\ &= \mathcal{N}\int \mathcal{D}\varphi \, \exp\left\{ i\int dx \mathcal{L}_{int}(\varphi)\right\} \\ &\quad \times \exp\left\{ i\int dx \left[\frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}(m^2 - i\varepsilon)\varphi^2 + J\varphi\right]\right\} \\ &= \mathcal{N}\exp\left\{ i\int dx \mathcal{L}_{int}\left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)\right\} Z_0[J] \end{aligned} \tag{10.105}$$

or, using (10.20),

$$Z[J] = \mathcal{N}\exp\left[i\int dx \mathcal{L}_{int}\left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)\right] \exp\left\{ -\frac{i}{2}\int dx dy J(x)\Delta_F(x-y)J(y)\right\}. \tag{10.106}$$

Thus we have obtained the general expression for generating functional of interacting theory, which will be used below to construct a diagram technique.

10.5 φ^4 theory

Let us return to the theory with interaction the Lagrangian

$$\mathcal{L}_{int} = -\frac{g}{4!}\varphi^4 . \tag{10.107}$$

The normalized generating functional for this theory is written as

$$Z[J] = \frac{\exp\left[i\int dz\mathcal{L}_{int}\left(\frac{1}{i}\frac{\delta}{\delta J(z)}\right)\right]\exp\left[-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right]}{\left\{\exp\left[i\int dz\mathcal{L}_{int}\left(\frac{1}{i}\frac{\delta}{\delta J(z)}\right)\right]\exp\left[-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right]\right\}\Big|_{J=0}}. \tag{10.108}$$

Perturbation theory is constructed by expanding this expression into the series in powers of the interaction constant g. Let us write the first terms of this expansion for the numerator:

$$\left[1-\frac{ig}{4!}\int dz\left(\frac{1}{i}\frac{\delta}{\delta J(z)}\right)^4 + O(g^2)\right]\exp\left[-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right]. \tag{10.109}$$

Making all the necessary differentiations we get

$$\frac{1}{i}\frac{\delta}{\delta J(z)}\exp\left[-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right] = -\int dx\Delta_F(z-x)J(x)\exp\left[-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right], \tag{10.110}$$

$$\left(\frac{1}{i}\frac{\delta}{\delta J(z)}\right)^2\exp\left[-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right] = \left\{i\Delta_F(0)+\left[\int dx\Delta_F(z-x)J(x)\right]^2\right\} \times\exp\left[-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right], \tag{10.111}$$

$$\left(\frac{1}{i}\frac{\delta}{\delta J(z)}\right)^3\exp\left[-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right] = \left\{3[-i\Delta_F(0)]\int dx\Delta_F(z-x)J(x)-\left[\int dx\Delta_F(z-x)J(x)\right]^3\right\} \times\exp\left[-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right], \tag{10.112}$$

$$\left(\frac{1}{i}\frac{\delta}{\delta J(z)}\right)^4 \exp\left[-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right] =$$
$$\left\{-3[\Delta_F(0)]^2 + 6i\,\Delta_F(0)\left[\int dx\Delta_F(z-x)J(x)\right]^2 + \left[\int dx\Delta_F(z-x)J(x)\right]^4\right\}\exp\left[-\frac{i}{2}\int dxdyJ(x)\Delta_F(x-y)J(y)\right]. \tag{10.113}$$

These expressions can be associated with diagrams. Let us draw function $-i\,\Delta_F(x-y)$ (propagator) by the straight line, connecting points x and y. The value of $-i\,\Delta_F(0) = -i\,\Delta_F(x-x)$ will be drawn as closed loop, connected with point x. Then, equation (10.113) is graphically represented as shown in Figure 10.8. The origin of coefficients 3, 6, 1 here can be understood from symmetry considerations. For example, coefficient 3 corresponds to three ways to connect two pairs of lines to draw diagram with two loops. Similarly, in the second term there are 6 ways to connect two lines to obtain the diagram shown in Figure 10.8. These coefficients are called symmetry factors, and we later shall discuss the general algorithm for finding them. Note that the first term in (10.113) and in Figure 10.8 represents the typical vacuum contribution (diagram) with no external lines.

Consider now the denominator of (10.108). We can simply put $J = 0$ in (10.113), which excludes the second and third terms in Figure 10.8. Thus, up to the terms of the order of g, generating the functional is expressed by the diagrams shown in Figure 10.9, where the second equality is obtained by expanding the denominator to the same accuracy, so that the vacuum diagram from the denominator is "lifted upwards" and cancels the vacuum diagram from the numerator. This reflects the general rule of cancellation of vacuum diagrams valid for *normalized* generating functionals in quantum field theory.

$$\left(\frac{1}{i}\frac{\delta}{\delta J(z)}\right)^4 \exp\left(-\frac{i}{2}\int J\Delta_F J\right) =$$

$$\left\{-3\,\bigcirc\!\!\bigcirc + 6i\,\text{—}\!\bigcirc\!\text{—} + \times\right\}\times$$

$$\times\ \exp\left(-\frac{i}{2}\int J\Delta_F J\right) =$$

Figure 10.8

$$Z[J]= \frac{\left[1-\frac{ig}{4!}\int\left(-3\ \bigcirc\!\!\!\!\bigcirc\ +6i\ \text{—}\!\bigcirc\!\text{—}\ +\ \times\right)dz\right]\ \exp\left(-\frac{i}{2}\int J\Delta_F J\right)}{1-\frac{ig}{4!}\int(-3\ \bigcirc\!\!\!\!\bigcirc\ )\,dz} =$$

$$= \left[1-\frac{ig}{4!}\int\left(6i\ \text{—}\!\bigcirc\!\text{—}\ +\ \times\right)dz\right]\ \exp\left(-\frac{i}{2}\int J\Delta_F J\right)$$

Figure 10.9

2-point function

2-point function is defined as

$$\tau(x_1,x_2) = -\left.\frac{\delta^2 Z[J]}{\delta J(x_1)\delta J(x_2)}\right|_{J=0}. \tag{10.114}$$

From Figure 10.9 it is clear, that contribution of the first term of $Z[J]$ into $\tau(x_1,x_2)$ is equal simply to $i\,\Delta_F(x_1-x_2)$, i. e., the free propagator. The diagram in Figure 10.9 with four "legs" contains four factors of J and does not contribute ($J=0$) to the 2-point function. The contribution of the diagram with lthe oop in $Z[J]$ is equal to

$$\frac{g}{4}\Delta_F(0)\int dxdy\Delta_F(z-x)J(x)\Delta_F(z-y)J(y)\exp\left(-\frac{i}{2}\int J\Delta_F J\right). \tag{10.115}$$

Differentiating this expression twice we get

$$\frac{1}{i}\frac{\delta}{\delta J(x_1)}(\cdots) = -\frac{ig}{4}\Delta_F(0)2\int dydz\Delta_F(z-x_1)\Delta_F(z-y)J(y)\exp\left(-\frac{i}{2}\int J\Delta_F J\right)+\cdots, \tag{10.116}$$

$$\frac{1}{i}\frac{\delta}{\delta J(x_2)}\frac{1}{i}\frac{\delta}{\delta J(x_1)}(\cdots) = -\frac{g}{2}\Delta_F(0)\int dz\Delta_F(z-x_1)\Delta_F(z-x_2)\exp\left(-\frac{i}{2}\int J\Delta_F J\right)+\cdots, \tag{10.117}$$

where we have dropped the terms giving zero for $J\to 0$. Finally, we have

$$\tau(x_1,x_2) = i\,\Delta_F(x_1-x_2) - \frac{g}{2}\Delta_F(0)\int dz\Delta_F(z-x_1)\Delta_F(z-x_2)+O(g^2), \tag{10.118}$$

which is shown by the diagrams in Figure 10.10. For the free particle we have

$$\tau(x) = i\,\Delta_F(x) = i\int\frac{d^4k}{(2\pi)^4}\frac{e^{-ikx}}{k^2-m^2+i\varepsilon} \tag{10.119}$$

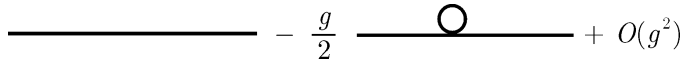


Figure 10.10

and the Fourier transform of the free propagator has the pole at $k^2 = m^2$, which determines the spectrum of the corresponding particle. It is easy to see that in the presence of interactions the particle mass becomes different form m. In fact, we can write the second term of Figure 10.10 as

$$\begin{aligned}
&-\frac{g}{2}\Delta_F(0)\int dz\Delta_F(x_1-z)\Delta_F(x_2-z)\\
&\quad=-\frac{g}{2}\Delta_F(0)\int\frac{d^4pd^4qdz}{(2\pi)^8}\frac{e^{-ip(x_1-z)}}{p^2-m^2+i\varepsilon}\frac{e^{-iq(x_2-z)}}{q^2-m^2+i\varepsilon}\\
&\quad=-\frac{g}{2}\Delta_F(0)\int\frac{d^4pd^4q}{(2\pi)^4}\frac{e^{-ip(x_1-x_2)}}{(p^2-m^2+i\varepsilon)^2}\delta(p+q)\\
&\quad=-\frac{g}{2}\Delta_F(0)\int\frac{d^4p}{(2\pi)^4}\frac{e^{-ip(x_1-x_2)}}{(p^2-m^2+i\varepsilon)^2}\,,
\end{aligned}\tag{10.120}$$

so that (10.118) reduces to

$$\tau(x_1,x_2)=i\int\frac{d^4p}{(2\pi)^4}\frac{e^{-ip(x_1-x_2)}}{p^2-m^2+i\varepsilon}\left\{1+\frac{i}{2}g\frac{\Delta_F(0)}{p^2-m^2+i\varepsilon}\right\}\,.\tag{10.121}$$

For $g \ll 1$ the term in the figure brackets in (10.121) can be rewritten (with the same accuracy) as

$$\left\{1-i\frac{g}{2}\frac{\Delta_F(0)}{p^2-m^2+i\varepsilon}\right\}^{-1}\,.\tag{10.122}$$

Then

$$\tau(x_1,x_2)=i\int\frac{d^4p}{(2\pi)^4}\frac{e^{-ip(x_1-x_2)}}{p^2-m^2-\frac{i}{2}g\Delta_F(0)+i\varepsilon}\,.\tag{10.123}$$

We see that now the Fourier transform of $\tau(x_1,x_2)$ has the pole at

$$p^2=m^2+\frac{i}{2}g\Delta_F(0)\equiv m^2+\delta m^2=m_r^2\,,\tag{10.124}$$

where

$$\delta m^2=\frac{i}{2}g\Delta_F(0)\,,\tag{10.125}$$

and m_r here represents the physical (or renormalized) mass of the particle. Thus, interaction changes the mass. Unfortunately, the value of δm^2 cannot be calculated, as it is formally infinite as $\Delta_F(0) \sim \int\frac{d^4k}{k^2} \sim \int dkk^3/k^2 \sim \int dkk$, and this integral is quadratically divergent at the upper limit. This is again a typical example of "ultraviolet" divergence in quantum field theory. The situation here is the same as in QED. The physical origin of divergence is the point-like nature of interaction in

local field theory. We do not know whether or not any "realistic" mechanism of the "cutoff" of these divergences exists. In the theory of condensed matter, in similar situations the upper limit of integration in momentum space is usually $\sim 1/a$, where a is some "minimal" length of the order of the average interatomic distance or lattice constant. There is no known analogue of such "minimal" length in quantum field theory. Its introduction (e. g., by assuming a kind of lattice structure of space-time at small distances) explicitly breaks the relativistic invariance of the theory. The problem is solved for renormalizable theories, where all such divergences can be "hidden" in the finite number of parameters to be determined from experiments. For renormalizable $g\varphi^4$-theory we shall return to the discussion of these problem later.

4-point function

We have

$$\tau(x_1, x_2, x_3, x_4) = \left. \frac{\delta^4 Z[J]}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)} \right|_{J=0} . \tag{10.126}$$

The term of the order of g^0 was considered above, and from (10.87) we have

$$\begin{aligned} \tau(x_1, x_2, x_3, x_4) = &-\Delta_F(x_2 - x_3)\Delta_F(x_1 - x_4) - \Delta_F(x_2 - x_1)\Delta_F(x_3 - x_4) \\ &- \Delta_F(x_3 - x_1)\Delta_F(x_2 - x_4) , \end{aligned} \tag{10.127}$$

which is shown diagrammatically in Figure 10.7 and corresponds to the free propagation of two particles without any scattering. Consider the contribution of the first order in g. From the diagrammatic form of generating a functional, shown in Figure 10.9, it is clear that one of the contributions of this type, which is due to differentiation of the loop graph in $Z[J]$, is shown in Figure 10.11 and is equal to

$$\begin{aligned} &\frac{g}{4}\frac{\delta^4}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)} \left\{ \Delta_F(0) \int dx \int dy \int dz \Delta_F(x - z)\Delta_F(y - z) \right. \\ &\qquad \left. \times J(y)J(x) \exp\left(-\frac{i}{2} \int J\Delta_F J \right) \right\} \Big|_{J=0} = \\ &\quad -\frac{ig}{2}\Delta_F(0) \int dz [\Delta_F(z - x_1)\Delta_F(z - x_2)\Delta_F(x_3 - x_4) \\ &\quad + \Delta_F(z - x_1)\Delta_F(z - x_3)\Delta_F(x_2 - x_4) \\ &\quad + \Delta_F(z - x_1)\Delta_F(z - x_4)\Delta_F(x_2 - x_3) \\ &\quad + \Delta_F(z - x_2)\Delta_F(z - x_3)\Delta_F(x_1 - x_4) \\ &\quad + \Delta_F(z - x_2)\Delta_F(z - x_4)\Delta_F(x_1 - x_2) \\ &\quad + \Delta_F(z - x_3)\Delta_F(z - x_4)\Delta_F(x_1 - x_2)] , \end{aligned} \tag{10.128}$$

which is shown by the diagram in Figure 10.12, which replaces six terms in this expression. The other contribution of the first order in g is obtained by differentiation

$$\frac{g}{4}\ \frac{g}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)}$$

$$\times\left\{ \bigcirc\!\!\!-\!\!\!- \exp\left[-\frac{i}{2}\int J(x)\Delta_F J(x-y)\,J(y)\,dx\,dy\right]\right\}\Bigg|_{J=0}$$

Figure 10.11

$$-3ig\left[\ \bigcirc\ \right]$$

Figure 10.12

of the "four leg" graph in $Z[J]$, which gives

$$-\frac{ig}{4!}\,\frac{\delta^4}{\delta J(x_1)\cdots\delta J(x_4)}\left\{\int dz\left[\int dx\Delta_F(z-x)J(x)\right]^4\exp\left(-\frac{i}{2}\int J\Delta_F J\right)\right\}\Bigg|_{J=0}$$
$$=-ig\int dz\Delta_F(x_1-z)\Delta_F(x_2-z)\Delta_F(x_3-z)\Delta_F(x_4-z)\,, \qquad (10.129)$$

which may be expressed graphically by a point with four "legs", where the point represents the elementary ("bare") interaction vertex.

Thus, the 4-point function up to terms of the order of g is expressed by the diagrams in Figure 10.13. Here the first term of the order g^0, as we noted above, does not contribute to scattering, the second term describes the self-interaction of one particles, and scattering itself is described only by the third term.

The numerical coefficients in Figure 10.13, as well as in other similar cases, can be understood from simple combinatorics. Consider an arbitrary diagram of the order g^n for a 4-point function. It contains n vertices, as shown in Figure 10.14. A 4-point function has 4 external "legs", as shown in Figure 10.15 ("prediagram"). Now we have to connect these "legs" in all possible ways with n vertices, using the rules of diagram technique. For example, in the first order in g there exist three topologically

$$\tau(x_1,x_2,x_3,x_4)=-3\left[\ \ \right]-3ig\left(\ \ \right)-ig\left(\ \ \right)=$$

$$=-3\left[\ \ \right]\frac{-ig}{4!}\left[\ 12\times 6\left(\ \ \right)+24\left(\ \ \right)\right]$$

Figure 10.13

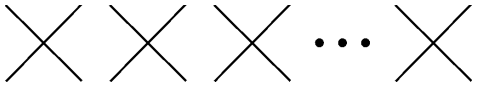


Figure 10.14

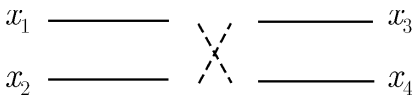


Figure 10.15

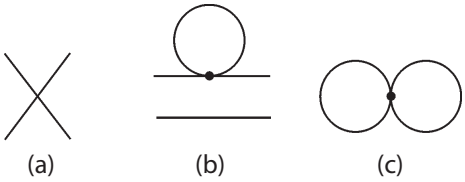


Figure 10.16

different types of Feynman diagrams, shown in Figure 10.16. To obtain the diagram of Figure 10.16(a) we have to connect x_1 in the prediagram of Figure 10.15 with one of the legs of the vertex; there are four ways to do this Afterwards, there remain only three ways to connect x_2 with one of the remaining legs, etc. In total, there are $4! = 24$ ways to obtain this diagram from the prediagram, leading to the corresponding coefficient in Figure 10.13. To obtain the diagram of Figure 10.16(b) we have to connect x_1 with one of the external legs x_2, x_3, x_4, which will produce a single line: there are three ways to do this. Then we take one of the vertex legs and connect it with one of remaining external points: this can be done in 4×2 ways. After that, we connect one of the three remaining legs of the dotted vertex to the last remaining point, which can be done in three ways. Finally, we connect the two remaining legs with each other. As a result, we obtain a multiplicity factor of $3 \times 4 \times 2 \times 3 = 12 \times 6$, which gives the coefficient before the diagram in Figure 10.13. It is clear that the multiplicity of the diagram in Figure 10.16(c) is equal to $3 \times 3 = 9$, but this (vacuum) diagram is not present in Figure 10.13, being cancelled by the corresponding contribution from the denominator of the normalized functional $Z[J]$.

Finally, we formulate the following diagram rules for $g\varphi^4$-theory (in coordinate representation).

- The free particle propagator $-i\,\Delta_F(x - y)$ is associated with the continuous line, connecting points x and y.
- The elementary interaction vertex is expressed by a point, connected with four continuous lines, and associated with factor $-i\,g$. There is integration over the vertex coordinates.

- Each diagram is multiplied by the corresponding symmetry factor $S(1/4!)^n$, where S is the number of ways to construct this diagram from the corresponding pre-diagram.

10.6 The generating functional for connected diagrams

We can introduce the generating functional $W[J]$, which generates only connected Feynman diagrams, i. e., diagrams which can not be represented by independent "blocks"[6]. Connected diagrams are important, because only these diagrams contribute to the nontrivial part of the S-matrix (scattering). The functional $W[J]$ is defined as

$$W[J] = -i \ln Z[J] , \tag{10.130}$$

so that

$$Z[J] = \exp\left(i W[J]\right) . \tag{10.131}$$

Let us consider, for example, 2-point and 4-point functions and show that $W[J]$ generates only connected diagrams. We have

$$\frac{\delta^2 W}{\delta J(x_1)\delta J(x_2)} = \frac{i}{Z^2}\frac{\delta Z}{\delta J(x_1)}\frac{\delta Z}{\delta J(x_2)} - \frac{i}{Z}\frac{\delta^2 Z}{\delta J(x_1)\delta J(x_2)} . \tag{10.132}$$

For $J = 0$ we have

$$\left.\frac{\delta Z[J]}{\delta J(x)}\right|_{J=0} = 0 , \quad Z[0] = 1 , \tag{10.133}$$

so that

$$\left.\frac{\delta^2 W}{\delta J(x_1)\delta J(x_2)}\right|_{J=0} = -i \left.\frac{\delta^2 Z}{\delta J(x_1)\delta J(x_2)}\right|_{J=0} = i\tau(x_1, x_2) . \tag{10.134}$$

We see that W determines the propagator in all orders of g.

Now for 4-point function: let us differentiate (10.132) two more times and put $J = 0$. Then we have

$$\begin{aligned}
&\frac{\delta^4 W}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)} \\
&= i\left[\frac{1}{Z^2}\frac{\delta^2 Z}{\delta J(x_1)\delta J(x_2)}\frac{\delta^2 Z}{\delta J(x_3)\delta J(x_4)} + \frac{1}{Z^2}\frac{\delta^2 Z}{\delta J(x_1)\delta J(x_3)}\frac{\delta^2 Z}{\delta J(x_2)\delta J(x_4)}\right. \\
&\qquad \left. + \frac{1}{Z^2}\frac{\delta^2 Z}{\delta J(x_1)\delta J(x_4)}\frac{\delta^2 Z}{\delta J(x_2)\delta J(x_3)} - \frac{1}{Z^2}\frac{\delta^4 Z}{\delta J(x_1)\delta J(x_2)\delta J(x_3)\delta J(x_4)}\right]\Bigg|_{J=0} \\
&= i\left[\tau(x_1, x_2)\tau(x_3, x_4) + \tau(x_1, x_3)\tau(x_2, x_4) + \tau(x_1, x_4)\tau(x_2, x_3) - \tau(x_1, x_2, x_3, x_4)\right] .
\end{aligned} \tag{10.135}$$

[6] An example of a nonconnected diagram is shown in Figure 10.16(b).

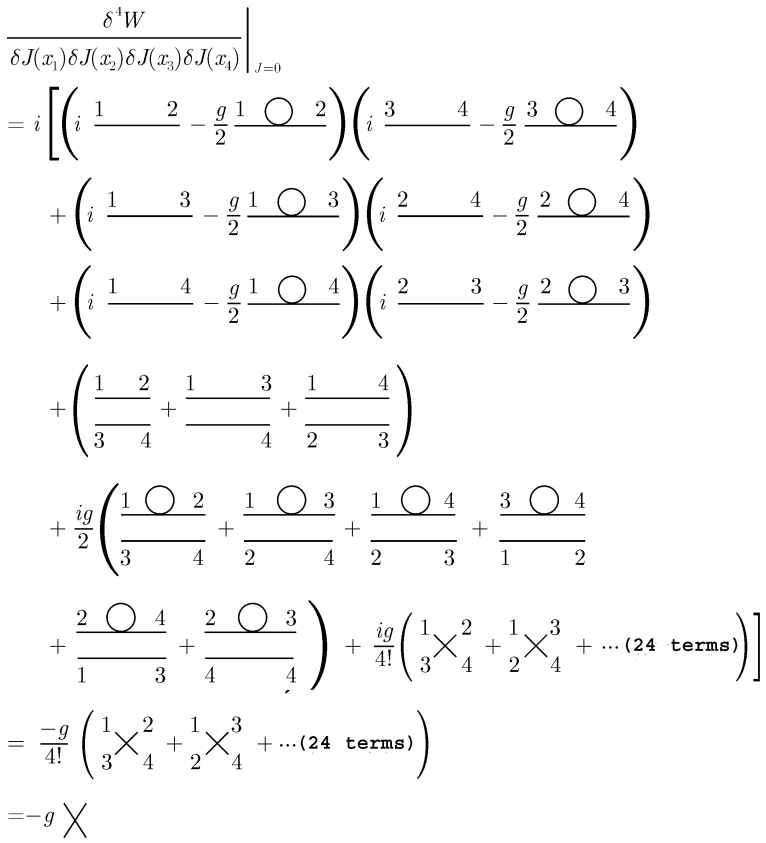


Figure 10.17

It is easily seen that this expression does not contain nonconnected diagrams. Substituting (10.118) and (10.129) into (10.135), with an accuracy up to the terms of the order g, we obtain Figure 10.17. We see that only connected diagrams contribute here.

Let us briefly discuss the n-point function

$$\tau(x_1,\ldots,x_n) = \frac{1}{i^n}\frac{\delta^n Z[J]}{\delta J(x_1)\ldots\delta J(x_n)}\bigg|_{J=0}. \tag{10.136}$$

The irreducible (connected) n-point function $\varphi(x_1,\ldots,x_n)$ can be defined as

$$\varphi(x_1,\ldots,x_n) = \frac{1}{i^n}\frac{\delta^n W[J]}{\delta J(x_1)\cdots\delta J(x_n)}\bigg|_{J=0}. \tag{10.137}$$

In fact, from Figure 10.13 and equation (10.136) directly follow the expressions, shown

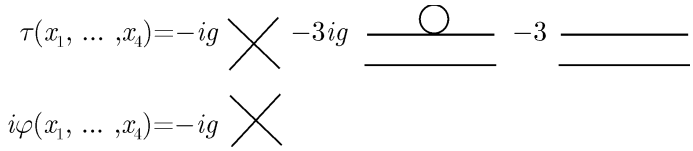


Figure 10.18

Figure 10.19

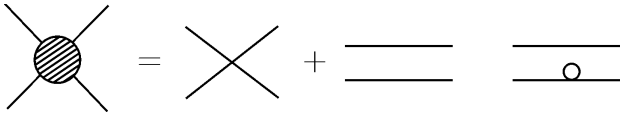


Figure 10.20

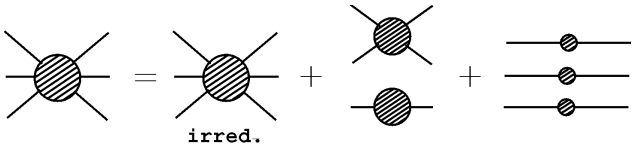


Figure 10.21

in Figure 10.18. From (10.135) it follows that

$$i\varphi(x_1,\ldots,x_4) = \tau(x_1,\ldots,x_4) - \tau(x_1,x_2)\tau(x_3,x_4) - \tau(x_1,x_3)\tau(x_2,x_4) - \tau(x_1,x_4)\tau(x_2,x_3)\,. \tag{10.138}$$

As $\tau(x_1,x_2) = i\varphi(x_1,x_2)$, we have:

$$\tau(x_1,\ldots,x_4) = i\varphi(x_1,\ldots,x_4) - \sum_p \varphi(x_{i_1},x_{i_2})\varphi(x_{i_3},x_{i_4})\,, \tag{10.139}$$

where the sum is taken over all the possible combinations of the indices $(1,\ldots,4)$ into the pairs $(i_1,i_2),(i_3,i_4)$. Thus, the 4-point function breaks into an "irreducible" (or connected) part and reducible parts, as shown in Figure 10.19. In the first order over g we have the diagrams shown in Figure 10.20. For the case of n-points functions, the appropriate generalization has the form shown in Figure 10.21.

10.7 Self-energy and vertex functions

Let us continue our discussion of the general structure of equations in quantum field theory in the functional formulation, limiting ourselves mainly to the $g\varphi^4$-theory. From the generating functional $Z[J]$ we can determine n-point functions $\tau(x_1,\ldots,x_n)$ (Green's functions $G^n(x_1,\ldots,x_n)$):

$$\tau(x_1,\ldots,x_n) = G^{(n)}(x_1,\ldots,x_n) = \frac{1}{i^n}\frac{\delta^n Z[J]}{\delta J(x_1)\ldots\delta J(x_n)}\bigg|_{J=0}. \qquad (10.140)$$

These functions contain both connected (irreducible) and nonconnected (reducible) parts, as shown, for example, for the case of $G^{(4)}$ in Figure 10.22. The scattering processes are determined only by connected diagrams, which are generated by functional $W = -i\ln Z$, so that the connected Green's functions are defined as

$$i\varphi(x_1,\ldots,x_n) = G_c^{(n)}(x_1,\ldots,x_n) = \frac{1}{i^{n-1}}\frac{\delta^n W[J]}{\delta J(x_1)\ldots\delta J(x_n)}\bigg|_{J=0}. \qquad (10.141)$$

Then, of all the graphs shown in Figure 10.22 only the third one remains, which determines $G_c^{(4)}$ in the first order over g.

Connected (irreducible) 2-point Green's function, up to the terms g^3, is determined by the diagrams shown in Figure 10.23. The complete sum of such diagrams gives

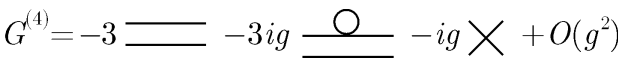


Figure 10.22

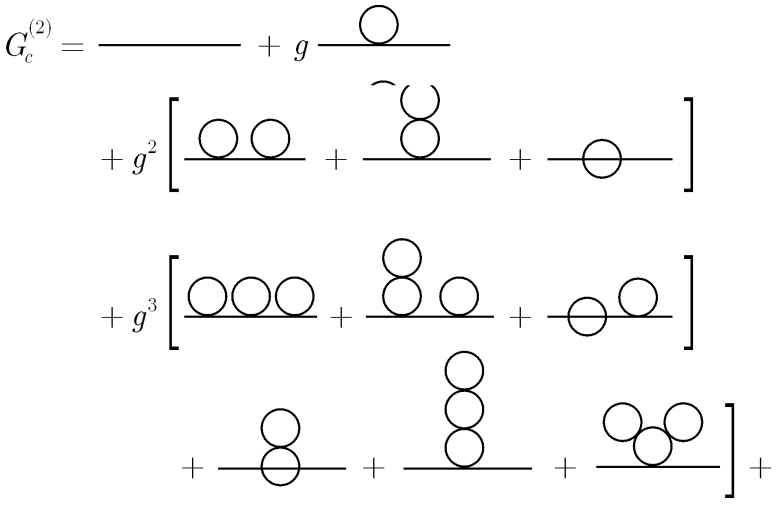


Figure 10.23

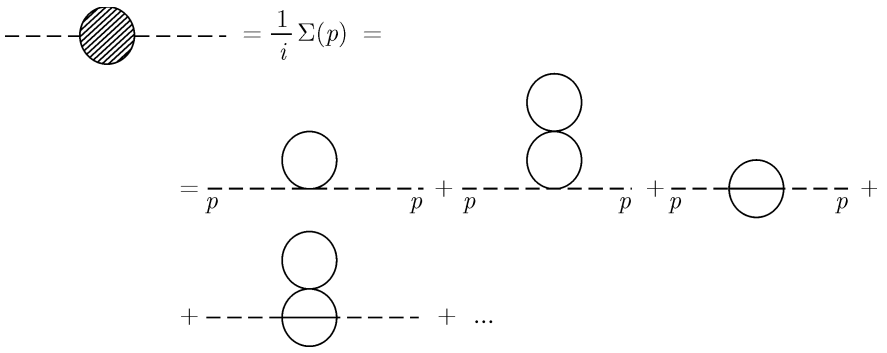


Figure 10.24

the "dressed" propagator $G_c^{(2)}(x, y)$, which is usually depicted by a "fat" line. We can perform the usual procedure, extracting *single-particle irreducible* diagrams (which cannot be cut over the single-particle line), and introduce their sum, as shown in Figure 10.24. This sum defines the irreducible self-energy part. The exact (dressed) propagator is now determined by the Dyson equation:

$$\begin{aligned} G_c^{(2)}(p) &= G_0(p) + G_0(p)\frac{1}{i}\Sigma(p)G_0(p) + G_0(p)\frac{1}{i}\Sigma(p)G_0(p)\frac{1}{i}\Sigma(p)G_0(p) + \cdots \\ &= G_0\left\{1 + \frac{1}{i}\Sigma G_0 + \frac{1}{i}\Sigma G_0\frac{1}{i}\Sigma G_0 + \ldots\right\} = \\ &= G_0\left[1 - \frac{1}{i}\Sigma G_0\right]^{-1} = \left[G_0^{-1}(p) - \frac{1}{i}\Sigma(p)\right]^{-1} \end{aligned} \tag{10.142}$$

or

$$G_c^{(2)}(p) = \frac{i}{p^2 - m^2 - \Sigma(p)}, \tag{10.143}$$

where we have taken into account that

$$G_0(p) = \frac{i}{p^2 - m^2}. \tag{10.144}$$

The Dyson equation is shown in diagrammatic form in Figure 10.25. Defining the

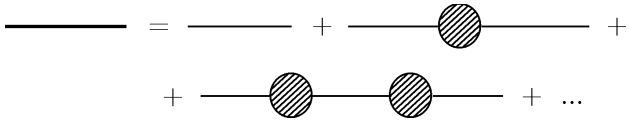


Figure 10.25

physical mass of a particle m_{phys} from the pole of the dressed propagator[7]

$$G_c^{(2)}(p) = \frac{i}{p^2 - m_{phys}^2}\,, \tag{10.145}$$

we obtain

$$m_{phys}^2 = m^2 + \Sigma(p^2 = m_{phys}^2)\,. \tag{10.146}$$

From equation (10.142) we have

$$[G_c^{(2)}(p)]^{-1} = G_0^{-1}(p) - \frac{1}{i}\Sigma(p)\,, \tag{10.147}$$

so that the inverse 2-point function contains (besides an inverse "bare" propagator) only single-particle irreducible diagrams. Formally, we can define the 2-point *vertex* function $\Gamma^{(2)}(p)$ by

$$G_c^{(2)}(p)\Gamma^{(2)}(p) = i\,, \tag{10.148}$$

which, according to (10.147), reduces to

$$\Gamma^{(2)}(p) = p^2 - m^2 - \Sigma(p)\,. \tag{10.149}$$

In fact, the nontrivial part of this construction reduces simply to $\Sigma(p)$, but this new notation is convenient within the framework of some universal notation system, which introduces the general notion of vertex functions[8].

We can also introduce the generating functional for n-point single-particle irreducible vertices Γ^n. This is denoted as $\Gamma[\varphi]$ and is usually called the *effective action*. This functional is defined by the so-called Legendre transformation of the functional $W[J]$:

$$W[J] = \Gamma[\varphi] + \int dx\, J(x)\varphi(x)\,. \tag{10.150}$$

Now we immediately obtain

$$\frac{\delta W[J]}{\delta J(x)} = \varphi(x)\,, \quad \frac{\delta\Gamma[\varphi]}{\delta\varphi(x)} = -J(x)\,. \tag{10.151}$$

Then as propagator we have

$$G(x,y) = -\frac{\delta^2 W[J]}{\delta J(x)\delta J(y)} = -\frac{\delta\varphi(x)}{\delta J(y)}\,. \tag{10.152}$$

Let us define $\Gamma(x,y)$ as

$$\Gamma(x,y) = \frac{\delta^2\Gamma[\varphi]}{\delta\varphi(x)\delta\varphi(y)} = -\frac{\delta J(x)}{\delta\varphi(y)}\,, \tag{10.153}$$

[7] The energy spectrum of a freely propagating "dressed" particle is determined from $p^2 = m_{phys}^2$.

[8] It is convenient to exclude the imaginary i from the definition of $G_c^{(2)}(p)$, so that the right-hand side of (10.148) is equal simply to 1. Correspondingly, in the Dyson equation we replace $\frac{1}{i}\Sigma \to \Sigma$. These notations are most common in the literature.

which is inverse to the propagator:

$$\int dx G(x,z)\Gamma(z,y) = -\int dz \frac{\delta^2 W[J]}{\delta J(x)\delta J(z)} \frac{\delta^2 \Gamma[\varphi]}{\delta\varphi(z)\delta\varphi(y)} = \int dz \frac{\delta\varphi(x)}{\delta J(z)} \frac{\delta J(z)}{\delta\varphi(y)} = \frac{\delta\varphi(x)}{\delta\varphi(y)} = \delta(x-y)\,, \qquad (10.154)$$

Differentiating both sides of (10.154) by $J(x'')$, replacing y by z, and using the relation

$$\frac{\delta}{\delta J(x'')} = \int dz'' \frac{\delta\varphi(z'')}{\delta J(x'')} \frac{\delta}{\delta\varphi(z'')} = -\int dz'' G(x'',z'') \frac{\delta}{\delta\varphi(z'')}\,; \qquad (10.155)$$

we then obtain

$$\int dz \frac{\delta^3 W}{\delta J(x)\delta J(x'')\delta J(z)} \frac{\delta^2\Gamma}{\delta\varphi(z)\delta\varphi(z')} - \int dz \times \int dz'' \frac{\delta^2 W}{\delta J(x)\delta J(z)} G(x'',z'') \frac{\delta^3\Gamma}{\delta\varphi(z)\delta\varphi(z')\delta\varphi(z'')} = 0\,, \qquad (10.156)$$

so that

$$\int dz \frac{\delta^3 W}{\delta J(x)\delta J(x'')\delta J(z)} \Gamma(z,z') + \int dz dz'' G(x,z) G(x'',z'') \frac{\delta^3\Gamma}{\delta\varphi(z)\delta\varphi(z')\delta\varphi(z'')} = 0\,. \qquad (10.157)$$

Multiplying both sides of the last equation by $G(x',z')$ and integrating by z', taking into account (10.154), we get

$$\frac{\delta^3 W}{\delta J(x)\delta J(x')\delta J(x'')} = -\int dz dz' dz'' G(x,z) G(x',z') G(x'',z'') \frac{\delta^3\Gamma}{\delta\varphi(z)\delta\varphi(z')\delta\varphi(z'')}\,. \qquad (10.158)$$

Thus, the connected 3-point function reduces to a single-particle irreducible 3-point vertex function, with external lines given by exact propagators. Correspondingly, $\frac{\delta^3\Gamma}{\delta\varphi(z)\delta\varphi(z')\delta\varphi(z'')}$ represents the complete three-leg vertex. All this is shown graphically in Figure 10.26. Equation (10.158) can be inverted with the help of (10.154), so that

$$\frac{\delta^3\Gamma}{\delta\varphi(y)\delta\varphi(y')\delta\varphi(y'')} = -\int dx dx' dx'' \Gamma(x,y)\Gamma(x',y')\Gamma(x'',y'') \frac{\delta^3 W}{\delta J(x)\delta J(x')\delta J(x'')}\,. \qquad (10.159)$$

In the right-hand side the external legs of (10.158) are "amputated".

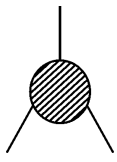


Figure 10.26

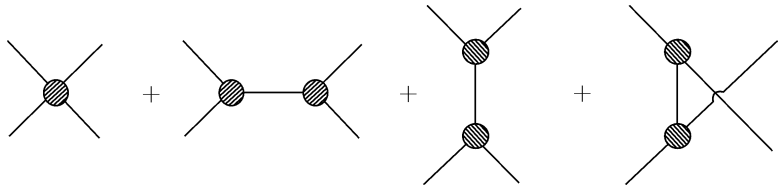


Figure 10.27

Differentiating (10.158) once more, we obtain the 4-point function represented by the diagrams in Figure 10.27, where a 4-leg irreducible vertex and a three single-particle reducible contributions appeared, corresponding to the three cross-channels of the reaction.

Thermodynamic analogy

There is a deep analogy between quantum field theory and statistical mechanics, which is expressed by the following table:

Quantum field theory	*Statistical mechanics*
Z - generating functional	Z - partition function
$Z = e^{iW}$	$Z = e^{-\frac{F}{T}}$
$W[J] = \Gamma[\varphi] + \int J\varphi$	F – free energy

In the following section we shall consider an explicit example of the application of the methods of quantum field theory to the theory of the critical phenomena at second-order phase transitions.

10.8 The theory of critical phenomena

Let us consider briefly one of most successful applications of quantum field theory methods to problems of statistical physics: the theory of critical phenomena in the vicinity of the critical temperature of the second order phase transitions. This problem remained unsolved by the traditional methods of statistical physics for a long time. The essence of the problem is well known: in the rather narrow (so-called critical) region near the phase transition temperature T_c, the critical exponent describing the singular behavior of physical properties at T_c is not satisfactory, as described by the general Landau theory of second-order phase transitions [36]. The reason for this deficiency is also quite clear; close to a phase transition point, strong *fluctuations* of the order parameter develop in the system which *strongly* interact with each other [36,42]. Significant progress in the theory was achieved with the introduction of the important concept of scale invariance, or *scaling* [42, 48]. However, the rigorous derivation of

this concept and the explicit calculations of the critical exponents became possible only after the development of the appropriate quantum field theory methods, which lead to the successes of the modern theory of critical phenomena [3, 42, 48]. Below we shall give a very short presentation of the main ideas and results of this theory, which illustrate a deep relationship between quantum field theory and statistical physics. Many important details of the calculations will, however, be skipped.

The Landau functional of free energy can be written in the standard form as [3, 42, 48][9]

$$\frac{1}{T}F[\phi(r)] = \int d^d r \left\{ \frac{1}{2}\sum_{j=1}^{n}\left[(\nabla\phi_j)^2 + \tau\phi_j^2\right] + \frac{1}{8}g\left(\sum_{j=1}^{n}\phi_j^2\right)^2 \right\}, \qquad (10.160)$$

where T is the temperature and the parameter $\tau = \frac{T-T_c}{T_c}$ determines the the size of the critical region close to the phase transition point. We shall limit ourselves to the temperature region $T > T_c$ (symmetric phase). The order parameter ϕ_j is represented by an n-component vector in some "isotopic" space with dimensionality n. Equation (10.160) is quite general. In fact, we are dealing with an $O(n)$-symmetric (isotropic) model of phase transition, which well describes a rather wide class of real systems. The case of $n = 1$ corresponds to the Ising model, $n = 2$ describes the so-called XY-model (superfluidity, superconductivity), $n = 3$ corresponds to isotropic Heisenberg ferromagnet, etc. [42, 48].

In the Landau theory, which completely neglects fluctuations of the order parameter (mean-field theory), $\phi = 0$ for $T > T_c$ [36]. However, even for $T > T_c$, fluctuations may lead to the appearance of regions in the system with $\phi(r) \neq 0$. The probability of such fluctuations is defined by [36, 42]

$$\mathcal{P}[\phi(r)] = \frac{1}{Z}\exp\left\{-\frac{1}{T}F[\phi(r)]\right\}, \qquad (10.161)$$

where partition function Z is determined by functional integral:

$$Z = \int \mathcal{D}\phi(r)\exp\left\{-\frac{1}{T}F[\phi(r)]\right\}. \qquad (10.162)$$

The free energy of the whole system is given by

$$F = -T\ln Z. \qquad (10.163)$$

The correlation function of the order parameter is defined as

$$G_{jl}(r,r') = Z^{-1}\int \mathcal{D}\phi(r)\phi_j(r)\phi_l(r')\exp\left\{-\frac{1}{T}F[\phi(r)]\right\} \equiv \langle\phi_j(r)\phi_l(r')\rangle\,. \qquad (10.164)$$

[9] From the very beginning we shall consider the space of dimension d because of the important dependence of critical phenomena on spatial dimensionality [36, 42].

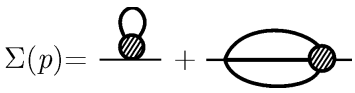


Figure 10.28

An analogy with the results of the previous sections is obvious: the theory of critical phenomena is equivalent to the *Euclidean* quantum theory of an n-component scalar field in d-dimensional space. Equation (10.164) is simply the propagator (Green's function or 2-point function) of such a field theory. In the simplest variant of the so-called Gaussian model of critical phenomena we have already met this theory in connection with equation (10.28).

The structure of the perturbation theory over the coupling constant g for fluctuations of the order parameter is quite similar to that in the $g\varphi^4$-theory with a single-component scalar field discussed above. The free Green's function coincides with Ornstein–Zernike correlator (cf.(10.27)):

$$G_{0jl}(p) = \frac{\delta_{jl}}{p^2 + \tau}\,. \tag{10.165}$$

The correlation function of interacting fluctuations is determined by the Dyson equation

$$G^{-1}(p) = G_0^{-1}(p) - \Sigma(p)\,, \tag{10.166}$$

where the self-energy part $\Sigma(p)$ is represented by the diagrams in Figure 10.28. The vertex part ("four-leg" vertex) determines the 4-point correlator $\langle\phi_i(r_1)\phi_j(r_2)\phi_l(r_3)\phi_m(r_4)\rangle$ etc.

Nontrivial physics of critical phenomena is connected with interaction of fluctuations. Let us consider the lowest-order perturbation theory corrections to the "bare" interaction, defined by the coupling constant g. In Figure 10.29 we show diagrams $\sim g^2$, corresponding to the three cross-channels of "two-particle scattering", determined by three sums of incoming momenta[10]:

1. $p_1 + p_2$,
2. $p_1 - p_3$,
3. $p_1 - p_4$.

Naturally, we have the conservation of the total momentum

$$p_1 + p_2 = p_3 + p_4\,. \tag{10.167}$$

In the problem with an n-component field it is convenient to use the symmetrized (over "isotopic" indices) form of the "bare" interaction

$$\bullet = g(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \equiv g I_{ijkl}\,. \tag{10.168}$$

[10] Arrows on lines define the directions of incoming and outgoing momenta.

Then the interaction term in (10.160) is $\sim I_{ijkl}\phi_i\phi_j\phi_k\phi_l$, where we assume summation over the repeating indices from 1 to n.

To find the full vertex part of a two-particle scattering, we need to perform the summation of all the diagrams like those shown in Figure 10.30. Obviously, in the general case this problem is unsolvable. However, we can introduce some topological classification of diagrams which allows us to write the general system of the so-called "parquet"(integral) equations, which determines this full vertex part [17]. It is clear that the full vertex Γ can be written as

$$\Gamma = R + \Gamma_1 + \Gamma_2 + \Gamma_3\,, \tag{10.169}$$

where the "blocks" $\Gamma_1, \Gamma_2, \Gamma_3$ are built of diagrams which can be cut over two lines in channels 1,2,3, while block R consists of all diagrams which cannot be cut in this way in either of these channels. Then for blocks $\Gamma_1, \Gamma_2, \Gamma_3$ we can construct diagrammatic equations, which are shown in Figure 10.31. Here we introduced blocks

$$I_i = R + \sum_{j\neq i} \Gamma_j\,, \tag{10.170}$$

which cannot be cut over two lines in channel i. The structure of the diagrams, determining block R, is clear from the diagrams shown in Figure 10.32.

This system of integral equations is very complicated. However, there is a case, where the solution is more or less simple. This is the so-called approximation of the "leading logarithms". To understand the main idea, let us estimate Diagram 1 in Figure 10.29[11]. Analytically, the contribution of this diagram is determined by the following integral:

$$g^2(n+8)\int \frac{d^d p}{(2\pi)^d}\,\frac{1}{p^2+\tau}\,\frac{1}{(p+k)^2+\tau}\,. \tag{10.171}$$

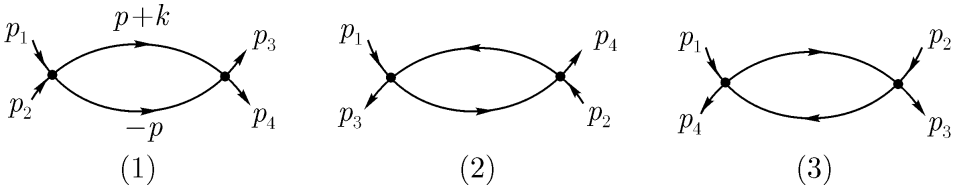


Figure 10.29

[11] In the following, in most cases we drop irrelevant numerical constants like symmetry factors etc.

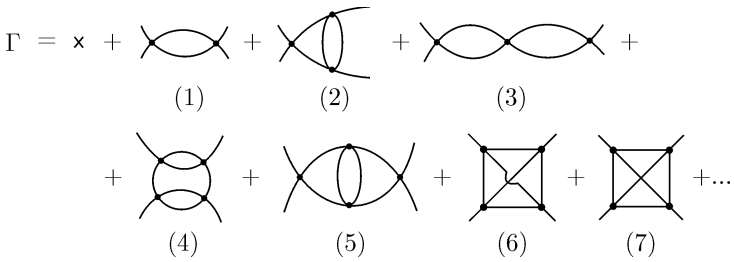


Figure 10.30

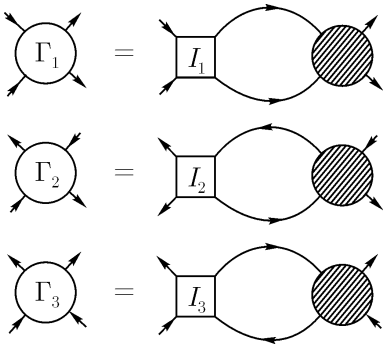


Figure 10.31

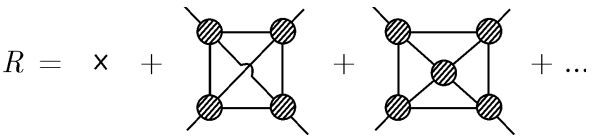


Figure 10.32

The factor $n + 8$ here originates from the product of two factors (10.168), standing at the vertices

$$I_{ijmn}I_{mnkl} + I_{ikmn}I_{mnjl} + I_{ilmn}I_{mnjk} = (n + 8)(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\,. \tag{10.172}$$

Now consider our theory in 4-dimensional space $d = 4$. Then we can estimate our integral as

$$\int d^4p \frac{1}{p^2+\tau}\frac{1}{(p+k)^2+\tau} \sim \int_{\mathrm{Max}(k,\sqrt{\tau})}^{\Lambda} \frac{dpp^3}{p^4} \sim \int_{\mathrm{Max}(k,\sqrt{\tau})}^{\Lambda} \frac{dp}{p}$$
$$\sim \ln \frac{\Lambda}{\mathrm{Max}(k,\sqrt{\tau})}\,, \tag{10.173}$$

where we have introduced upper limit cutoff $\Lambda \sim \frac{1}{a}$. Here, in contrast to quantum field theory, we do not have any problem with logarithmic divergence at the upper limit; Landau expansion (10.160) is valid only on the length scale, which is much larger than the interatomic distances a, and there is simply no fluctuation with shorter wavelengths. The value of a plays the role of "minimal length", which is absent in quantum field theory. However, the presence of logarithmic divergence in (10.173) is of prime importance. This logarithm becomes very large in the vicinity of a phase transition point when we are dealing with $k, \sqrt{\tau} \ll \Lambda$. Here we have "infrared" divergence for τ or k tending to zero! In fact, with an accuracy up to second order terms in g we now have

$$\Gamma(k) \approx g - g^2(n+8) \ln \frac{\Lambda}{\mathrm{Max}(k, \sqrt{\tau})} + \cdots . \tag{10.174}$$

We see that the first correction to the interaction vertex for $\tau \to 0$, $k \to 0$ can become much larger than the bare coupling constant g; fluctuations become strongly interacting as we move to the transition point. This is the essence of the problem: we need some everywhere relevant corrections, and this is not an easy problem to solve! For $d = 3$ this seems to be a hopeless task; however, for $d = 4$, due to the relatively weak logarithmic singularity, we can perform the summation of a certain set of diagrams, corresponding to the "leading logarithm" approximation. The logarithm appears from momentum integration in the loop graph. In higher orders higher powers of logarithm appear, with their powers determined by the number of loops in the corresponding diagram. For example, considering in a similar way Diagrams 2 and 3 in Figure 10.30, we estimate their contribution to be $\sim g^3 \ln^2 \frac{\Lambda}{\mathrm{Max}(k,\sqrt{\tau})}$, while for Diagram 4 we get $\sim g^4 \ln^3 \frac{\Lambda}{\mathrm{Max}(k,\sqrt{\tau})}$. At the same time, the estimate of Diagram 6 gives $\sim g^4 \ln \frac{\Lambda}{\mathrm{Max}(k,\sqrt{\tau})}$, while for Diagram 7 we have $\sim g^5 \ln \frac{\Lambda}{\mathrm{Max}(k,\sqrt{\tau})}$, which is much smaller than contributions of $2, 3, 4$ due to the assumption of weakness of the "bare" coupling $g \ll 1$. Thus, we can limit ourselves to the "leading logarithms", i. e., take into account only those diagrams which give the power of the logarithm as equal to the power of coupling constant g minus 1; in Figure 10.30 these are Diagrams 2, 3, and 4. The topology of these diagrams is quite clear: in the given order they contain the maximal number of loops, like Figure 10.29. This set of graphs is typically called "parquet" diagrams. "Parquet" takes into account all the vertex corrections of the order of $\sim g^n \ln^n$, but neglects contributions like $\sim g^{n+k} \ln^n$. Moreover, we can neglect the contributions without logarithms. In particular, block R introduced above now reduces to the first term in Figure 10.32, i. e., simply to the "bare" interaction g. In such an approximation the "parquet" equations of Figure 10.31 can be solved. However, the procedure of this solution is rather complicated, and we shall not discuss it. The correct answer for the full vertex can be obtained using a more "naïve" analysis, which will be used here. Consider the simple one-dimensional set of diagrams shown in Figure 10.33. This is a simple progression, which is easily summed (in contrast to

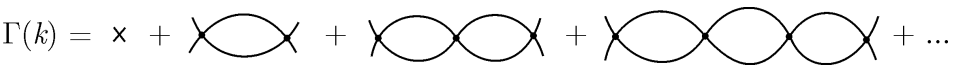


Figure 10.33

a two-dimensional "parquet"):

$$\Gamma(k) = g - g^2(n+8)\ln\frac{\Lambda}{\mathrm{Max}(k,\sqrt{\tau})} + g^3(n+8)^2\ln^2\frac{\Lambda}{\mathrm{Max}(k,\sqrt{\tau})} + \cdots$$
$$= \frac{g}{1+g(n+8)\ln\frac{\Lambda}{\mathrm{Max}(k,\sqrt{\tau})}}\,. \tag{10.175}$$

The rigorous solution of "parquet" equations gives the same answer (for external momenta of the same order of magnitude)[12]. "Parquet" equations for such vertices are reduced to the differential equation

$$\frac{d\Gamma(s)}{ds} = -(n+8)\Gamma^2(s) \tag{10.176}$$

with the boundary condition $\Gamma(s) \to g$ for $s \to 0$. Here we introduce the logarithmic variable

$$s = \ln\frac{\Lambda}{\mathrm{Max}(k,\sqrt{\tau})}\,. \tag{10.177}$$

Integration of (10.176) gives

$$\Gamma(s) = \frac{g}{1+g(n+8)s}\,, \tag{10.178}$$

which coincides with (10.175). In fact, this result is completely similar to expressions for physical charge obtained in Chapter 8 during our discussion on the asymptotic properties of QED which lead to the problem of the "Moscow zero" or the Landau "ghost pole"[13].

Consider the case of $k = 0$ (interaction of fluctuations with very long wavelengths). Then (10.175) reduces to

$$\Gamma(k=0) = \frac{g}{1+g(n+8)\ln\frac{\Lambda}{\sqrt{\tau}}} \to \frac{1}{(n+8)\ln\frac{\Lambda}{\sqrt{\tau}}} \quad \text{for } \tau \to 0\,. \tag{10.179}$$

Now, as we approach the point of phase transition, the dependence on the "bare" coupling constant g cancels, while the effective interaction *tends to zero* (typical "zero-charge" behavior!)[14]. But here, in contrast to relativistic field theory, this behavior

[12] This coincidence with the correct answer is pretty accidental; the rigorous solution was first obtained in [17].

[13] A result of the type of (10.175) was first obtained from the analysis of "parquet" equations during the studies of asymptotic properties of relativistic scalar field $g\varphi^4$-theory [17]. In the theory of critical phenomena for $d = 4$ it was obtained much later by Larkin and Khmelnitskii [38].

[14] Let us stress that in the theory of critical phenomena we have $g > 0$, so that here we have no problems like "ghost poles".

does not lead to any problem and, in fact, completely clarifies the situation. Equation (10.179) corresponds to the effective interaction of fluctuations becoming weaker as we approach the phase transition. We can explicitly calculate the influence of this weak interaction on all physical characteristics which are singular at the phase transition point, and we can see that it leads only to some insignificant (logarithmic) temperature corrections to the critical behavior described by the Landau theory. These logarithmic corrections do not change the *powers* of the temperature singularities, i. e., critical exponents. Thus, for $d = 4$ the critical indices (exponents) are simply equal to their values of the Landau theory!

Definitions of critical exponents

The theory of critical phenomena usually considers the following standard set of physical characteristics of the system and critical exponents (indices), determining sthe ingular behavior of these characteristics for $\tau = \frac{T-T_c}{T_c} \to 0$.

Order parameter:

$$\bar{\varphi} \sim |\tau|^{\beta}, \quad T \to T_c - 0, \tag{10.180}$$

$$\bar{\varphi} \sim h^{\frac{1}{\delta}}, \quad T = T_c, \tag{10.181}$$

where h is the external field, interacting with the order parameter.

Susceptibility:

$$\chi \sim \begin{cases} \tau^{-\gamma}, & T \to T_c + 0, \\ |\tau|^{-\gamma'}. & T \to T_c - 0. \end{cases} \tag{10.182}$$

Correlation function of the order parameter (d – spatial dimensionality):

$$G(r) \sim \frac{\exp(-r/\xi)}{r^{d-(2-\eta)}}, \tag{10.183}$$

where the correlation length is

$$\xi \sim \begin{cases} \tau^{-\nu}, & T \to T_c + 0, \\ |\tau|^{-\nu'}, & T \to T_c - 0. \end{cases} \tag{10.184}$$

At the critical point itself,

$$G(r) \sim \frac{1}{r^{d-(2-\eta)}}, \tag{10.185}$$

$$G(p) \sim \frac{1}{k^{2-\eta}}. \tag{10.186}$$

In a similar way we introduce the critical index of specific heat α:

$$C(\tau, h=0) = \frac{A^+}{\alpha}[\tau^{-\alpha} - 1] + B^+, \qquad T \to T_c + 0, \tag{10.187}$$

$$C(\tau, h=0) = \frac{A^-}{\alpha'}[|\tau|^{-\alpha'} - 1] + B^-, \qquad T \to T_c - 0, \tag{10.188}$$

so that $\alpha = 0$ corresponds to logarithmic singularity.

It is important to note that in Landau theory (mean-filed theory) the values of the critical indices are [36]

$$\begin{aligned} &\nu = \frac{1}{2}\,, \quad \gamma = 1\,, \quad \eta = 0\,, \\ &\alpha = 0\,, \quad \beta = \frac{1}{2}\,, \quad \delta = 3 \end{aligned} \tag{10.189}$$

and satisfy the standard scaling relations [36, 42]

$$\begin{aligned} \nu &= \frac{\gamma}{2-\eta}\,, \\ \alpha &= 2 - \nu d\,, \\ \beta &= \frac{1}{2}\nu(d-2+\eta)\,, \end{aligned} \tag{10.190}$$

in 4-dimensional space $d = 4$. In this sense we can say that Landau theory gives a correct description of critical phenomena for $d = 4$. The same statement is actually valid for all $d > 4$; it can be easily seen that corrections of the type of (10.171) do not lead to any divergences for $d > 4$ and are small due to the assumption of $g \ll 1$. Spatial dimensionality $d = 4$ is called the *upper critical dimension* of the theory.

For the physically most interesting case of $d = 3$ there is no possibility of choosing the "leading" (dominating) diagrams in a perturbation series; actually, all the diagrams are of the same order. This was the main obstacle for construction a rigorous theory of critical phenomena. Wilson has proposed an original method for calculating critical exponents, which is based on the idea of the introduction of an artificial small parameter of the perturbation theory $\varepsilon = 4 - d$: a deviation from the upper critical dimensionality $d = 4$, for which all critical indices coincide with the predictions of the mean-field theory (ε-expansion). The idea of "fractional" spatial dimensionality is rather simple. In all the Feynman integrals above we have dealt with integration over the volume of d-dimensional momentum space, with the volume element in spherical coordinates (for the integrand depending only on the absolute value of momentum) written as

$$d^d p = \Omega_d\, p^{d-1} dp\,, \tag{10.191}$$

where Ω_d is the surface of a d-dimensional sphere of the unit radius

$$\Omega_d = \frac{2\pi^{d/2}}{\Gamma\left(\frac{d}{2}\right)}\,, \tag{10.192}$$

where we use the usual definition of a Γ-function. In this expression we already can consider d as an arbitrary (noninteger) real parameter. Then we can write

$$\int \frac{d^d p}{(2\pi)^d} \cdots = \frac{\Omega_d}{(2\pi)^d} \int dp\; p^{d-1} \cdots = K_d \int dp\; p^{d-1} \cdots\,, \tag{10.193}$$

where we introduces the standard notation

$$K_d = 2^{-(d-1)}\pi^{-d/2}\left[\Gamma\left(\frac{d}{2}\right)\right]^{-1} . \tag{10.194}$$

In particular, $K_4 = (8\pi^2)^{-1}$. Previously, during our estimates of (10.173) this constant was just dropped. In the following we shall also drop it.

Let us estimate once again the contribution of Diagram 1 in Figure 10.29 for the space dimensionality $d = 4 - \varepsilon$. Instead of (10.173) we have

$$\begin{aligned} g^2(n+8)K_d \int_{\mathrm{Max}(k,\sqrt{\tau})}^{\Lambda} dp\, p^{d-1}\frac{1}{p^4} &\sim g^2(n+8)\int_{\mathrm{Max}(k,\sqrt{\tau})}^{\Lambda} dp\, p^{d-5} \\ &\sim g^2(n+8)\frac{1}{d-4}p^{d-4}\Big|_{\mathrm{Max}(k,\sqrt{\tau})}^{\Lambda} \\ &\sim g^2(n+8)\frac{1}{\varepsilon}\left\{\left[\mathrm{Max}(k,\sqrt{\tau})\right]^{-\varepsilon} - \Lambda^{-\varepsilon}\right\} . \end{aligned} \tag{10.195}$$

The changes in comparison to the case of $d = 4$ reduce the replacement of logarithm (10.173) by a "logarithmic variable"

$$s = \frac{1}{\varepsilon}\left\{\left[\mathrm{Max}(k,\sqrt{\tau})\right]^{-\varepsilon} - \Lambda^{-\varepsilon}\right\} , \tag{10.196}$$

which gives the same logarithm in the limit of $\varepsilon \to 0$. Thus, during the solution of "parquet" equations, we can again use the approximation of the "leading logarithms", and the differential equation for the vertex (10.176) conserves its form. Its solution (10.178) for the case of $k = 0$ in space with $d = 4 - \varepsilon$ can now be written as

$$\begin{aligned} \Gamma(k=0) = \frac{g}{1 + g(n+8)\frac{1}{\varepsilon}[\tau^{-\varepsilon/2} - \Lambda^{-\varepsilon}]} &\to \\ \frac{1}{(n+8)\frac{1}{\varepsilon\tau^{\varepsilon/2}}} = \frac{\varepsilon\tau^{\varepsilon/2}}{(n+8)} &\qquad \text{for } \tau \to 0 . \end{aligned} \tag{10.197}$$

We see that the effective interaction of fluctuations is small due to the assumed smallness of our artificial parameter $\varepsilon = 4 - d$.

Equation (10.176) can also be considered as a differential equation over *cutoff parameter* Λ, which enters the variable s (10.196), (10.177): $ds = \Lambda^{-(1+\varepsilon)}d\Lambda$. In this case this equation describes the renormalization of the vertex Γ under infinitesimal transformation of the cutoff parameter $\Lambda \to \Lambda' = \Lambda + d\Lambda$. Essentially this is the differential equation of the renormalization group introduced first by Gell-Mann and Low and already known to us in the case of QED. Renormalization group ideology is the basis of the modern theory of phase transitions [3, 42, 69].

Let us schematically present how critical exponents are calculated in ε-expansion. Consider a correlation function of the order parameter (Green's function) $G(p\tau)$. We

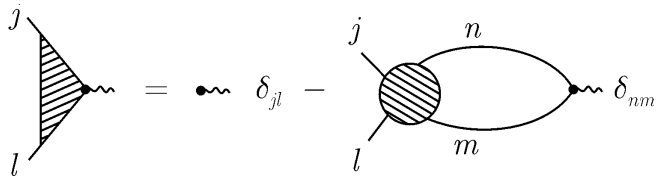


Figure 10.34

have, by definition:

$$G(p=0\tau)=\chi(\tau)\sim\tau^{-\gamma}\,,$$
$$G(p\tau=0)\sim p^{-2+\eta}\,. \tag{10.198}$$

Let us limit ourselves to indices γ aqndη, as all others can be determined from scaling relations like (10.190) [42,48].

In this theory we may prove two Ward identities:

$$\frac{\partial}{\partial p_\alpha}G^{-1}_{jl}(p\tau=0)=2p_\alpha\delta_{jl}-2\int\frac{d^dp'}{(2\pi)^d}p'_\alpha G^2_{mm}(p'0)\Gamma_{jlmm}(ppp'p')\,, \tag{10.199}$$

$$\frac{\partial}{\partial\tau}G^{-1}_{jl}(p\tau=0)=\delta_{jl}-\int\frac{d^dp'}{(2\pi)^d}G^2_{mm}(p'0)\Gamma_{jlmm}(ppp'p')\,. \tag{10.200}$$

Introducing the "triangular" vertex $\mathcal{T}_{jl}=\frac{\partial}{\partial\tau}G^{-1}_{jl}(p\tau=0)$, we can draw the second of these identities as shown in Figure 10.34. This identity can be derived by differentiation of diagrams for self-energy (inverse propagator), as is shown schematically in Figure 10.35. Differentiation of the inverse free propagator (10.165) (Figure 10.35(a)) gives the first term, while differentiation of the simplest contribution to self-energy (Figure 10.35(a)) gives lowest order contributions to the vertex with two linked "legs", i. e., the lowest-order contribution to the second term. The full series of "differentiated" graphs is summed to the full vertex. The identity (10.199) is derived in a similar way, differentiating by p_α.

Let us substitute into (10.199) the "parquet" solution for $\Gamma(ppp'p')$. We have not derived it explicitly, but it is sufficient to know that (similarly to $\Gamma(k)$ derived above) it depends only on the absolute values $|p|$ and $|p'|$, so that the integral in the right-hand

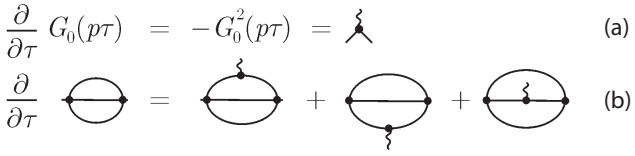


Figure 10.35

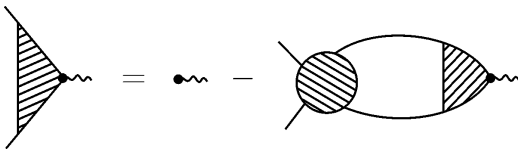


Figure 10.36

side of (10.199) gives zero after integration over the polar angle. Thus, we simply have

$$\frac{\partial G^{-1}(p\tau = 0)}{\partial p_\alpha} = 2p_\alpha \,, \tag{10.201}$$

so that

$$G(p\tau = 0) \sim \frac{1}{p^2} \,, \tag{10.202}$$

which gives the value of the critical exponent $\eta = 0$.

Let us use now the Ward identity (10.200). In a "parquet" approximation we can resume the diagrams in such a way that this identity reduces to an integral equation for a "triangular" vertex, shown in Figure 10.36. Using logarithmic variables we can rewrite this equation as

$$\mathcal{T}_{jl}(s) = \delta_{jl} - \int_0^s dt \Gamma_{jlmn}(t)\mathcal{T}_{mn}(t) \,. \tag{10.203}$$

Using $\mathcal{T}_{jl} = \mathcal{T}\delta_{jl}$ and (10.168) we obtain

$$I_{jlmn}\delta_{mn} = (n+2)\delta_{jl} \tag{10.204}$$

and (10.203) reduces to

$$\mathcal{T}(s) = 1 - (n+2)\int_0^s dt \Gamma(t)\mathcal{T}(t) \,. \tag{10.205}$$

Differentiating by s, we reduce this integral equation to a differential one:

$$\frac{d\mathcal{T}(s)}{ds} = -(n+2)\Gamma(s)\mathcal{T}(s) \tag{10.206}$$

with boundary condition $\mathcal{T}(s=0) = 1$. Then we find

$$\mathcal{T}(s) = \exp\left\{-(n+2)\int_0^s dt \Gamma(t)\right\} \,. \tag{10.207}$$

Using here (10.178), we obtain finally

$$\mathcal{T}(s) = [1 + g(n+8)s]^{-\frac{n+2}{n+8}} \,. \tag{10.208}$$

Then we have

$$\frac{\partial}{\partial\tau}G^{-1}(p=0\tau) = \frac{\partial\chi^{-1}(\tau)}{\partial\tau} = [1 + g(n+8)s]^{-\frac{n+2}{n+8}} \tag{10.209}$$

Integrating with the necessary accuracy, we get

$$\chi(\tau) \approx \frac{1}{\tau}\left\{1 + g(n+8)\frac{1}{\varepsilon}[\tau^{-\varepsilon/2} - \Lambda^{-\varepsilon}]\right\}^{\frac{n+2}{n+8}} \to \tau^{-(1+\frac{\varepsilon}{2}\frac{n+2}{n+8})} \tag{10.210}$$

for $\tau \to 0$. Then, for the susceptibility critical exponent we find

$$\gamma = 1 + \frac{n+2}{n+8}\frac{\varepsilon}{2} + \cdots . \tag{10.211}$$

This expression, as well as the previous result $\eta = 0$, are valid up to the terms of the first order in ε, and represent the first terms of the ε-expansion of critical indices. More tedious calculations allow the derivation of higher order corrections.

One remarkable result of the modern theory following from these expressions is the *universality* of critical behavior; the values of the critical exponents in quite different physical systems are determined only by the dimensionality of space (or system) and the number of components n of the order parameter (i. e., in fact, the type of the symmetry broken during the phase transition).

Expansion (10.160) may in principle contain higher powers of the order parameter. What is their role in critical behavior? Why have we limited ourselves only to $g\phi^4$? Consider a possible term like $\lambda\phi^6$ and the simplest diagram due to such an interaction, shown in Figure 10.37. By the order of magnitude it is determined by the integral:

$$\lambda^2 \int d^3p_1 \int d^3p_2 \frac{1}{p_1^2 p_2^2 (p_1^2 + p_2^2)} \sim \lambda^2 \int_{\sqrt{\tau}}^{\Lambda} dp_1 \int_{\sqrt{\tau}}^{\Lambda} dp_2 \frac{p_1^2 p_2^2}{p_1^2 p_2^2 (p_1^2 + p_2^2)}$$
$$\sim \lambda^2 \ln \frac{\Lambda}{\sqrt{\tau}} . \tag{10.212}$$

For $d > 3$ this correction just converges (at the lower limit, for $\tau \to 0$), so that for $d = 4 - \varepsilon$ an interaction of the type of $\lambda\phi^6$ is actually *irrelevant*. Quite analogous is the situation with the higher-order terms of the Landau expansion, which justifies the analysis made above.

In conclusion, let us quote the values of critical indices up to the terms of the order of $\sim \varepsilon^2$ for the theory with an n-component order parameter [3, 42]:

$$\gamma = 1 + \frac{n+2}{n+8}\frac{\varepsilon}{2} + \frac{n+2}{n+8}\frac{n^2+22n+52}{(n+8)^2}\frac{\varepsilon^2}{4} + \cdots ,$$

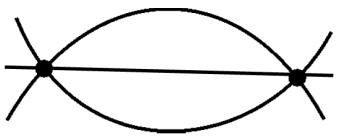


Figure 10.37

$$
\begin{aligned}
2\nu &= 1 + \frac{n+2}{n+8}\frac{\varepsilon}{2} + \frac{n+2}{n+8}\frac{n^2+23n+60}{(n+8)^2}\frac{\varepsilon^2}{4} + \cdots , \\
\eta &= \frac{n+2}{2(n+8)^2}\varepsilon^2 + \frac{n+2}{2(n+8)^2}\left[\frac{6(3n+14)}{(n+8)^2} - \frac{1}{4}\right]\varepsilon^3 + \cdots , \\
\delta &= 3 + \varepsilon + \left[\frac{1}{2} - \frac{n+2}{(n+8)^2}\right]\varepsilon^2 + \cdots , \\
\beta &= \frac{1}{2} - \frac{3}{n+8}\frac{\varepsilon}{2} + \frac{(n+2)(2n+1)}{2(n+8)^3}\varepsilon^2 + \cdots , \\
\alpha &= \frac{4-n}{n+8}\frac{\varepsilon}{2} + \cdots .
\end{aligned}
\tag{10.213}
$$

It is interesting to compare the values calculated from these expressions for $d = 3$ ($\varepsilon = 1$) and $n = 1$ (the Ising case), with the results of numerical calculations (high-temperature expansion) for the three-dimensional Ising model. In the table we also show the values derived from the mean-field theory (Landau). We see that ε-expansion gives rather a satisfactory agreement with the results of numerical analysis[15].

Table 10.1. Critical indices for the Ising model ($n = 1$).

Index	*Wilson*	*Numerical*	*Landau*
ν	0.626	0.642	0.5
η	0.037	0.055	0
γ	1.244	1.250	1
α	0.077	0.125	0
β	0.340	0.312	0.5
δ	4.460	5.15	3

Modern methods of the calculation of critical exponents significantly improve the results of simple ε-expansion; taking into account higher-order diagrams, they give the values of the indices, which practically coincide with the results of numerical calculations and experiments [69].

10.9 Functional methods for fermions

Generalization of the functional integral approach to quantization to fermions is not obvious. In the Bose case, functional integration is performed over all possible classical (c-number) field configurations. For fermion fields the classical limit is absent,

[15] Another effective method for calculating critical indices is based on the expansion in powers of the inverse number of components of the order parameter $1/n$ [3,42]; as for $n \to \infty$ it can be shown that the indices also reduce to their values in the mean-field approximation (Landau theory). Calculations are based on the summation of loop diagrams, as each loop contribution is $\sim n$.

and it is not clear what kind of field configurations we can introduce at all. The classical limit is achieved as $\hbar \to 0$. In this case, the nontrivial right-hand side of all Bose field operators, as considered in Chapter 2, tend to zero, and the operators become c-numbers. For fermion fields, quantization is done with *anticommutators*, so that for $\hbar \to 0$ in the Fermi case we just get some anticommuting variables, with no obvious "common-sense" meaning. However, it happens that these variables lead to a correct solution of our problem. Such variables were introduced in mathematics by Grassmann in the middle of 19th century and are called *Grassmann variables*. A functional formulation in quantum field theory for fermions, using Grassmann variables, was proposed by Berezin, who introduced the notion of integration over these variables [7].

Consider first the mathematical definitions. Generators C_i of n-dimensional Grassmann algebra satisfy the anticommutation relations

$$\{C_i, C_j\} \equiv C_i C_j + C_j C_i = 0\,, \tag{10.214}$$

where $i = 1, 2, \ldots, n$. In particular,

$$C_i^2 = 0\,. \tag{10.215}$$

Thus, the series expansion of an arbitrary function $f(C_i)$ can contain only a finite number of terms. For example, in the case of one-dimensional algebra we have

$$f(C) = a + bC\,, \tag{10.216}$$

where a and b are usual numbers. The quadratic and higher-power terms of this expansion are equal to zero.

For the general n-dimensional case the analogue of (10.216) takes the form

$$f(C) = P_0 + P_1^i C_i + P_2^{ij} C_i C_j + \cdots + P_n C_1 C_2 \cdots C_n\,, \tag{10.217}$$

where each summation index takes values from 1 to n, and coefficients P are antisymmetric with respect to the permutation of any pair of indices $i, j, \ldots$. Expansion is cut a finite number of terms because of (10.214).

Consider the notion of differentiation over Grassmann variables. We can introduce two types of derivatives, left and right. The left derivative of the product $C_1 C_2$ is defined as

$$\frac{\partial^L}{\partial C_i}(C_1 C_2) = \delta_{i1} C_2 - \delta_{i2} C_1\,. \tag{10.218}$$

Ccorrespondingly, the right derivative is given by

$$\frac{\partial^R}{\partial C_i}(C_1 C_2) = \delta_{i2} C_1 - \delta_{i1} C_2\,. \tag{10.219}$$

Then we have the following equalities:

$$\left\{\frac{\partial}{\partial C_i}, C_j\right\} = \delta_{ij}\,, \tag{10.220}$$

$$\left\{\frac{\partial}{\partial C_i}, \frac{\partial}{\partial C_j}\right\} = 0\,. \tag{10.221}$$

In particular, for one-dimensional algebra,

$$\left\{\frac{d}{dC}, C\right\} = 1 \tag{10.222}$$

and we always have

$$\left(\frac{\partial}{\partial C_i}\right)^2 = 0\,. \tag{10.223}$$

All these relations are rather natural.

In contrast, the definition of an integral over Grassmann variables is rather formal. In particular, it is impossible to introduce it as the inverse operation to differentiation. However, it can be defined in such a way that it possesses some general properties characteristic of the usual integral. For example, we can require our integral to be invariant towards the shift of the integration variable by a constant:

$$\int dC f(C) = \int dC f(C+\alpha)\,. \tag{10.224}$$

This is always so for the usual integral with infinite limits of integration, but here our new definition of integration has nothing in common with the usual definition (except the notation $\int$), and there is no limits of integration here in the usual sense. Using the explicit form of $f(C)$ (10.216), we obtain

$$\int dC(a+bC) = \int dC\,[a+b(C+\alpha)]\,, \qquad \text{so that}$$

$$\int dC\, bC = \int dC\, b(C+\alpha)\,, \tag{10.225}$$

from which it follows that

$$\int dC\, b\alpha = 0 \tag{10.226}$$

or, due to arbitrariness of $b\alpha$,

$$\int dC = 0\,. \tag{10.227}$$

Here α is another element of Grassmann algebra, independent of and anticommuting with C. The remaining integral $\int dC\, C$ can be just defined by the condition

$$\int dC\, C = 1\,. \tag{10.228}$$

Equations (10.227) and (10.228) completely determine integration over Grassmann variables.

Naturally, integration thus defined has nothing in common with the usual notion of an integral. Moreover, in the case of one-dimensional Grassmann algebra we have $\frac{df}{dC} = b$ and $\int dC f(C) = b$, so that the operation of integration acts upon a function in the same way as differentiation!

In the n-dimensional case we assume

$$\int dC_i = 0\,, \quad \int dC_i C_i = 1\,. \tag{10.229}$$

Let η and $\bar{\eta}$ be independent Grassmann variables, so that

$$\int d\eta = \int d\bar{\eta} = 0\,, \quad \int d\eta\eta = \int d\bar{\eta}\bar{\eta} = 1\,. \tag{10.230}$$

As $\eta^2 = \bar{\eta}^2 = 0$, we have

$$e^{-\bar{\eta}\eta} = 1 - \bar{\eta}\eta\,, \tag{10.231}$$

so that

$$\int d\bar{\eta} d\eta\, e^{-\bar{\eta}\eta} = \int d\bar{\eta} d\eta - \int d\bar{\eta} d\eta\, \bar{\eta}\eta = 0 + \int d\bar{\eta} d\eta\, \eta\bar{\eta} = 1\,. \tag{10.232}$$

Let us find the generalization of this expression for the case of a larger number of dimensions. Consider the two-dimensional case, introducing for convenience the new notations

$$\eta = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}, \quad \bar{\eta} = \begin{pmatrix} \bar{\eta}_1 \\ \bar{\eta}_2 \end{pmatrix}. \tag{10.233}$$

The exponent $\bar{\eta}\eta$ (or more precisely $\bar{\eta}^T\eta$) has the form

$$\bar{\eta}\eta = \bar{\eta}_1\eta_1 + \bar{\eta}_2\eta_2\,. \tag{10.234}$$

Then,

$$\begin{aligned}(\bar{\eta}\eta)^2 &= (\bar{\eta}_1\eta_1 + \bar{\eta}_2\eta_2)(\bar{\eta}_1\eta_1 + \bar{\eta}_2\eta_2) \\ &= \bar{\eta}_1\eta_1\bar{\eta}_2\eta_2 + \bar{\eta}_2\eta_2\bar{\eta}_1\eta_1 = 2\bar{\eta}_1\eta_1\bar{\eta}_2\eta_2\,,\end{aligned} \tag{10.235}$$

where we have taken into account that $\eta_1^2 = \eta_2^2 = \bar{\eta}_1^2\bar{\eta}_2^2 = 0$. Higher powers of $\bar{\eta}\eta$ are equal to zero, and we get

$$e^{-\bar{\eta}\eta} = 1 - (\bar{\eta}_1\eta_1 + \bar{\eta}_2\eta_2) + \bar{\eta}_1\eta_1\bar{\eta}_2\eta_2\,. \tag{10.236}$$

Applying our integration rules, we see that

$$\int d\bar{\eta} d\eta e^{-\bar{\eta}\eta} = \int d\bar{\eta}_1 d\eta_1 d\bar{\eta}_2 d\eta_2 \bar{\eta}_1\eta_1\bar{\eta}_2\eta_2 = 1\,, \tag{10.237}$$

as in the one-dimensional case. Let us transform the integration variables as

$$\eta = M\alpha\,, \quad \bar{\eta} = N\bar{\alpha}\,, \tag{10.238}$$

where M and N are 2×2-matrices, while α and $\bar{\alpha}$ are new Grassmann variables. Then we have

$$\begin{aligned}\eta_1\eta_2 &= (M_{11}\alpha_1 + M_{12}\alpha_2)(M_{21}\alpha_1 + M_{22}\alpha_2)\\ &= (M_{11}M_{22} - M_{12}M_{21})\alpha_1\alpha_2 = (\mathrm{Det}\, M)\alpha_1\alpha_2\,,\end{aligned} \tag{10.239}$$

where we have taken into account the anticommutativity of Grassmann variables. To conserve the integration rules

$$\int d\eta_1 d\eta_2 \eta_1\eta_2 = \int d\alpha_1 d\alpha_2 \alpha_1\alpha_2\,, \tag{10.240}$$

we have to require that

$$d\eta_1 d\eta_2 = (\mathrm{Det}\, M)^{-1} d\alpha_1 d\alpha_2\,, \tag{10.241}$$

which differs from the usual rule for an integration variable change by the power of the determinant. Taking into account that

$$\bar{\eta}\eta = N\bar{\alpha}M\alpha = N\bar{\alpha}\alpha M^T = -\alpha M^T N\bar{\alpha} = \bar{\alpha}M^T N\alpha\,, \tag{10.242}$$

we write (10.237) as

$$(\mathrm{Det}\, MN)^{-1}\int d\bar{\alpha}d\alpha e^{-\bar{\alpha}M^T N\alpha} = 1\,. \tag{10.243}$$

As $\mathrm{Det}\, MN = \mathrm{Det}\, M^T N$ we obtain the general result

$$\int d\bar{\alpha}d\alpha e^{-\bar{\alpha}A\alpha} = \mathrm{Det}\, A\,, \tag{10.244}$$

which represents the Gaussian integral over Grassmann variables.

To describe fermion *fields* we make a transition to Grassmann algebra with infinite dimensions, with the appropriate generators denoted as $C(x)$:

$$\{C(x), C(y)\} = 0\,, \tag{10.245}$$

$$\frac{\partial^{L,R}C(x)}{\partial C(y)} = \delta(x-y)\,, \tag{10.246}$$

$$\int dC(x) = 0\,, \quad \int dC(x)C(x) = 1\,. \tag{10.247}$$

As a result we obtain functional integrals over Grassmann (Fermion) fields.

As we already know, Dirac's Lagrangian has the form

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi\,. \tag{10.248}$$

Then, the normalized generating functional for the free Dirac field can be written as

$$\begin{aligned}Z_0[\eta,\bar{\eta}] &= \frac{1}{\mathcal{N}}\int \mathcal{D}\bar{\psi}\mathcal{D}\psi\\ &\times\exp\left\{i\int dx[\bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x) + \bar{\eta}(x)\psi(x) + \bar{\psi}(x)\eta(x)]\right\},\end{aligned} \tag{10.249}$$

where the normalization factor is

$$\mathcal{N} = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi \exp\left[i \int dx \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x)\right]. \tag{10.250}$$

Here we introduced the Grassmannian source $\bar{\eta}(x)$ for field $\psi(x)$ and $\eta(x)$ for field $\bar{\psi}(x)$.

To shorten the notations we introduce

$$S^{-1} = i\gamma^\mu \partial_\mu - m\,. \tag{10.251}$$

Then,

$$Z_0[\eta, \bar{\eta}] = \frac{1}{\mathcal{N}} \int \mathcal{D}\bar{\psi}\mathcal{D}\psi \exp\left[i \int dx (\bar{\psi} S^{-1}\psi + \bar{\eta}\psi + \bar{\psi}\eta)\right]. \tag{10.252}$$

Consider the quadratic form

$$Q(\psi, \bar{\psi}) = \bar{\psi} S^{-1}\psi + \bar{\eta}\psi + \bar{\psi}\eta\,. \tag{10.253}$$

Let us find the value of ψ "minimizing" this form from the condition

$$\frac{\partial^L Q}{\partial \bar{\psi}} = S^{-1}\psi + \eta = 0\,, \quad \frac{\partial^R Q}{\partial \psi} = \bar{\psi} S^{-1} + \bar{\eta} = 0\,, \tag{10.254}$$

which gives

$$\psi_m = -S\eta\,, \quad \bar{\psi}_m = -\bar{\eta}S\,, \tag{10.255}$$

where we have assumed the existence of the inverse operator S^{-1}. At the "minimum" we have

$$Q = Q_m = Q(\psi_m, \bar{\psi}_m) = -\bar{\eta}S\eta\,. \tag{10.256}$$

Then our quadratic form can be written as

$$Q = Q_m + (\bar{\psi} - \bar{\psi}_m)S^{-1}(\psi - \psi_m)\,. \tag{10.257}$$

Correspondingly[16],

$$\begin{aligned} Z_0[\eta, \bar{\eta}] &= \frac{1}{\mathcal{N}} \int \mathcal{D}\bar{\psi}\mathcal{D}\psi \exp\left\{i \int dx [Q_m + (\bar{\psi} - \bar{\psi}_m)S^{-1}(\psi - \psi_m)]\right\} \\ &= \frac{1}{\mathcal{N}} \exp\left[-i \int dx \int dy \bar{\eta}(x) S(x-y)\eta(y)\right] \mathrm{Det}(-iS^{-1})\,, \end{aligned} \tag{10.258}$$

where the last expression was derived extracting the factor $\exp\left[i \int Q_m\right]$ outside the integral, as Q_m does not depend on ψ and $\bar{\psi}$ and used the obvious functional generalization of (10.244)

$$\int \mathcal{D}\bar{\psi}\mathcal{D}\psi e^{-\bar{\psi}A\psi} = \mathrm{Det}\, A\,. \tag{10.259}$$

[16] The second equality in (10.258) is the fermion analogue of equation (10.48) for a complex boson field.

In a similar way we may show that $\mathcal{N} = \mathrm{Det}(-iS^{-1})$, so that finally we obtain the generating functional of the free Dirac field as

$$Z_0[\eta,\bar{\eta}] = \exp\left[-i\int dx\int dy\,\bar{\eta}(x)S(x-y)\eta(y)\right]. \tag{10.260}$$

It is easy to see that operator S really exists. It has the form

$$S(x) = (i\gamma^\mu\partial_\mu + m)\Delta_F(x)\,, \tag{10.261}$$

where $\Delta_F(x)$ is the well-known Feynman propagator of a scalar field. In fact, using (10.251) we have

$$S^{-1}S = (i\gamma^\mu\partial_\mu - m)(i\gamma^\mu\partial_\mu + m)\Delta_F(x) = (-\Box - m^2)\Delta_F(x) = \delta(x)\,. \tag{10.262}$$

Now we can define the free propagator of the Dirac field as

$$\begin{aligned}\tau(x,y) &= -\left.\frac{\delta^2 Z_0[\eta,\bar{\eta}]}{\delta\eta(x)\delta\bar{\eta}(y)}\right|_{\eta=\bar{\eta}=0}\\ &= -\frac{\delta}{\delta\eta(x)}\frac{\delta}{\delta\bar{\eta}(y)}\left\{-i\int dx\int dy\,\bar{\eta}(x)S(x-y)\eta(y)\right\}\Big|_{\eta=\bar{\eta}=0}\\ &= iS(x-y)\,,\end{aligned} \tag{10.263}$$

where we have used $\exp(-\bar{\eta}S\eta) = 1 - \bar{\eta}S\eta$.

Let us summarize the main expressions related to the free scalar and spinor fields. For the scalar field we have

$$\mathcal{L}_0 = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 = -\frac{1}{2}\varphi(\Box + m^2)\varphi\,. \tag{10.264}$$

We have seen above that

$$\tau(x,y) = i\,\Delta_F(x-y)\,, \tag{10.265}$$

where Δ_F is the Feynman propagator satisfying the equation

$$(\Box + m^2)\Delta_F(x-y) = -\delta(x-y)\,. \tag{10.266}$$

For the spinor (Dirac) field we have

$$\mathcal{L}_0 = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi = \bar{\psi}S^{-1}\psi\,, \tag{10.267}$$

$$\tau(x,y) = iS(x-y)\,. \tag{10.268}$$

In both cases we see that the propagator is inverse (operator) to the coefficient before the quadratic term in the Lagrangian. We can take this as the *definition* of a free propagator in the functional formulation for an arbitrary field.

Generating the functional of interacting Dirac fields can be defined similarly to the case of scalar fields:

$$Z[\eta,\bar{\eta}] = \exp\left[i\int dx\,\mathcal{L}_{int}\left(\frac{1}{i}\frac{\delta}{\delta\eta},\frac{1}{i}\frac{\delta}{\delta\bar{\eta}}\right)\right] Z_0[\eta,\bar{\eta}]\,. \tag{10.269}$$

From this expression we can derive all the rules of diagram technique for Fermi fields, in the same way how this was done above for the scalar field. The only important difference, due to the Grassmann nature of fermion fields, is the necessity of associating an additional factor of (-1) with each fermion loop[17]. We shall not give details of the diagrammatic rules for purely fermion models of particle interaction, as all such theories are nonrenormalizable in 4-dimensional space-time.

As an example of a fermion interaction model which is really applicable to particle interactions, we only mention the so-called Fermi (4-fermion) interaction. This is quite successful as a description of low-energy interactions of leptons. The corresponding interaction Lagrangian (for two lower generations of leptons) is written in the standard form [40]

$$\mathcal{L}_{int} = \frac{G}{\sqrt{2}} j_w^+ j_w\,, \tag{10.270}$$

where j_w is a lepton-weak current operator:

$$\begin{aligned} j_w^\alpha &= \bar{\psi}_e\Gamma^\alpha\psi_{\nu e} + \bar{\psi}_\mu\Gamma^\alpha\psi_{\nu\mu}\,,\\ j_w^{\alpha+} &= \bar{\psi}_{\nu e}\Gamma^\alpha\psi_e + \bar{\psi}_{\nu\mu}\Gamma^\alpha\psi_\mu\,, \end{aligned} \tag{10.271}$$

where

$$\Gamma^\alpha = \frac{1}{2}(1-\gamma^5)\gamma^\alpha(1+\gamma^5)\,, \tag{10.272}$$

and the lower indices of the field operators denote the corresponding particles (electron e, muon μ, electron neutrino ν_e, muon neutrino ν_μ).

From simple dimensional analysis it becomes clear that the Lagrangian corresponds to nonrenormalizable theory: the coupling constant G is dimensional, with the dimensionality of squared length or the inverse square of the mass. Its numerical value is well known from experimental data on low-energy processes (well-described by first order of perturbation theory over G), such as muon decay, and is written usually as

$$G = 1.0\cdot 10^{-5}\frac{\hbar^3}{m_p^2 c} = 1.43\cdot 10^{-49}\,\mathrm{erg}\cdot\mathrm{cm}^3\,, \tag{10.273}$$

where m_p is the proton mass, introduced here just as a dimensional parameter. Its appearance in (10.273) is rather artificial, and later we shall see how such interaction appears as effective in the modern theory of weak and electromagnetic interactions, and which mass scale is actually at work here.

Due to the nonrenormalizability of field theory with the interaction Lagrangian (10.270), it cannot be considered as fundamental, and it is rather senseless to write higher-order corrections of perturbation theory over G.

[17] It can be shown [56], that the origin of this factor is related to the functional generalization of (10.221), which has the form $\frac{\delta^2}{\delta\eta(x)\delta\eta(y)} = -\frac{\delta^2}{\delta\eta(y)\delta\eta(x)}$.

10.10 Propagators and gauge conditions in QED

In QED we can write down the generating functional of Maxwell field as

$$Z[J] = \int \mathcal{D}A_\mu \exp\left\{ i \int dx (\mathcal{L} + J^\mu A_\mu) \right\} , \tag{10.274}$$

where J^μ is an external source current,

$$\mathcal{L} = -\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu} . \tag{10.275}$$

Making partial integration over dx and dropping the surface integrals, we can rewrite this Lagrangian as

$$\mathcal{L} = \frac{1}{8\pi} A^\mu [g_{\mu\nu} \Box - \partial_\mu \partial_\nu] A^\nu . \tag{10.276}$$

The Lagrangian of the electromagnetic field is invariant to gradient (gauge) transformations $A_\mu \to A_\mu + \partial_\mu \Lambda$. At the same time, the functional integral in (10.274) is taken over *all* A_μ, including those connected with each other by gauge transformations. Obviously this leads to the appearance of an *infinite* contribution to Z and to Green's functions. It is clear that it is necessary to fix some gauge, so that the integral over A_μ *is not calculated* over the field configurations, which are obtained from each other by gauge transformations. Physically these configurations are just equivalent! Here we meet a problem which becomes especially difficult for non-Abelian gauge theories. In fact, this problem can be rigorously solved, as it will be shown in the next chapter. Here we just limit ourselves to several technical remarks.

If we use the Lorentz gauge $\partial_\mu A^\mu = 0$, the Lagrangian (10.276) becomes

$$\mathcal{L} = \frac{1}{8\pi} A^\mu g_{\mu\nu} \Box A^\nu . \tag{10.277}$$

The inverse operator for $g^{\mu\nu}\Box$ is represented by the Feynman propagator (see e. g., Chapter 4)

$$D_{F\mu\nu}(x, y) = g_{\mu\nu} 4\pi \Delta_F (x, y; m = 0) . \tag{10.278}$$

In momentum representation, the operator $-g_{\mu\nu}k^2$, originating from (10.277), has an inverse operator written as $-g^{\nu\lambda}\frac{1}{k^2}$, so that the Feynman propagator of the electromagnetic field in the Lorentz gauge has the form

$$D_{F\mu\nu}(k) = g_{\mu\nu} \frac{4\pi}{k^2} . \tag{10.279}$$

In the general case we can add to the Lagrangian an extra term, fixing the gauge, with an arbitrary coefficient α:

$$\mathcal{L} = -\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu} - \frac{1}{8\pi\alpha} (\partial_\mu A^\mu)^2 = \frac{1}{8\pi} A^\mu \left[g_{\mu\nu} \Box + \left(\frac{1}{\alpha} - 1 \right) \partial_\mu \partial_\nu \right] A^\nu . \tag{10.280}$$

In momentum representation, the coefficient before the square of the field is

$$-k^2 g_{\mu\nu} + \left(1 - \frac{1}{\alpha}\right) k_\mu k_\nu \,, \tag{10.281}$$

with the corresponding inverse operator produces propagator as

$$D_{\mu\nu}(k) = \frac{4\pi}{k^2}\left[g_{\mu\nu} + (\alpha - 1)\frac{k_\mu k_\nu}{k^2}\right] . \tag{10.282}$$

For $\alpha \to 1$ we obtain a Feynman propagator (Lorentz-Feynman gauge), while for $\alpha \to 0$ we obtain the propagator in a Landau gauge.

Chapter 11

Functional integrals: gauge fields

11.1 Non-Abelian gauge fields and Faddeev–Popov quantization

Let us consider the quantum theory of non-Abelian gauge fields. For a long time, quantization of Yang-Mills fields remained an unsolved problem due to difficulties related to the necessity of a correct account of gauge invariance. In particular, attempts to quantize this theory along the lines of the traditional (operator) approach in quantum field theory were mostly unsuccessful, despite the successes of Abelian QED. The complete solution of the problem was achieved by Faddeev and Popov, who used functional methods. In the following we mainly follow the presentation of [13].

The heuristic idea

We have seen above that the value of the generating functional Z, defined in the usual way, in case of gauge theory (even QED) is. in general, infinite, as it contains integration over *all* fields A_μ, including those connected with each other by gauge transformations, which leave an integrand invariant.

Before we start the analysis, allowing separation of the corresponding infinite "volume" factor from the (infinite-dimensional) functional integral over the gauge field, we shall consider a simple illustration of the main idea of our general method for the case of the usual two-dimensional integral:

$$W = \int dx \int dy e^{iS(x,y)} = \int d\mathbf{r} e^{iS(\mathbf{r})}, \tag{11.1}$$

where $\mathbf{r} = (r, \theta)$ defines the polar coordinates of a point on the plane. Assume that function $S(\mathbf{r})$ (analogue to the action!) is invariant with respect to rotations in two-dimensional space:

$$S(\mathbf{r}) = S(\mathbf{r}_\phi) \tag{11.2}$$

as $\mathbf{r} = (r, \theta) \to \mathbf{r}_\phi = (r, \theta + \phi)$. This means, that $S(\mathbf{r})$ is constant at the circles ("orbits") in the (x, y)-plane, as shown in Figure 11.1(a). In this trivial example, if we want to take into account only contributions from nonequivalent values of $S(\mathbf{r})$, we need to extract a "volume factor", corresponding to integration over angular variable[1]

[1] Angular integration is assumed to be done from 0 to 2π, and we just drop these integration limits below.

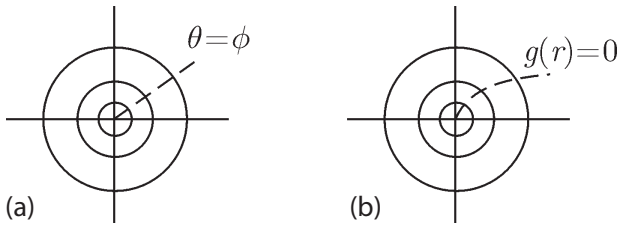


Figure 11.1

$\int d\theta = 2\pi$. To formalize this, we shall use the following "recipe", to be generalized for more complicated cases later. Let us put inside our integral a factor 1, written in the form[2]

$$1 = \int d\phi \delta(\theta - \phi)\,. \tag{11.3}$$

Then we have

$$W = \int d\phi \int d\mathbf{r} e^{iS(\mathbf{r})} \delta(\theta - \phi) = \int d\phi W_\phi\,, \tag{11.4}$$

where

$$W_\phi = \int d\mathbf{r} \delta(\theta - \phi) e^{iS(\mathbf{r})} \tag{11.5}$$

is calculated for the given value of $\phi = \theta$. Thus, first of all we calculate W at the fixed value of $\phi = \theta$ (constraint!), and afterwards integrate over all values of ϕ (see Figure 11.1(a)). Using the invariance of S (11.2), we have

$$W_\phi = W_{\phi'}\,. \tag{11.6}$$

We see that the "volume" of the orbit can be extracted as a factor:

$$W = \int d\phi W_\phi = W_\phi \int d\phi = 2\pi W_\phi\,. \tag{11.7}$$

In the eneral case, we can use a more complicated constraint (instead of $\phi = \theta$), which can be represented by some curve $g(\mathbf{r}) = 0$, crossing each orbit only once, as shown in Figure 11.1(b), so that the equation $g(\mathbf{r}_\phi) = 0$ has the unique solution for ϕ at a fixed value of r. Taking such a general constraint, we define, instead of simple equation (11.3), the "representation of unity" of the form

$$1 = \Delta_g(\mathbf{r}) \int d\phi \delta[g(\mathbf{r}_\phi)]\,. \tag{11.8}$$

In other words, we define the function $\Delta_g(\mathbf{r})$ as

$$[\Delta_g(\mathbf{r})]^{-1} = \int d\phi \delta[g(\mathbf{r}_\phi)]\,. \tag{11.9}$$

[2] Here it is assumed that θ is within the interval $(0, 2\pi)$.

Using the general rule

$$\int dx\delta[f(x)] = \int df \frac{1}{df/dx}\delta(f) = \left.\frac{1}{df/dx}\right|_{f=0}, \tag{11.10}$$

we obtain

$$\Delta_g(\mathbf{r}) = \left.\frac{\partial g(\mathbf{r}_\theta)}{\partial\theta}\right|_{g=0}, \tag{11.11}$$

and $\Delta_g(\mathbf{r})$ here is invariant with respect to two-dimensional rotations:

$$[\Delta_g(\mathbf{r}_{\phi'})]^{-1} = \int d\phi\delta[g(\mathbf{r}_{\phi+\phi'})] = \int d\phi''\delta[g(\mathbf{r}_{\phi''})] = [\Delta_g(\mathbf{r})]^{-1}. \tag{11.12}$$

Then, repeating arguments similar to those used during the transformation from (11.4) to (11.7), we can again extract from our integral the "volume factor" 2π:

$$W = \int d\phi \int d\mathbf{r}\Delta_g(\mathbf{r})\delta[g(\mathbf{r}_\phi)]e^{iS(\mathbf{r})} = \int d\phi W_\phi, \tag{11.13}$$

where

$$W_\phi = \int d\mathbf{r}e^{iS(\mathbf{r})}\Delta_g(\mathbf{r})\delta[g(\mathbf{r}_\phi)]. \tag{11.14}$$

The entire nontrivial part of the integral is here! The "volume factor" is equal just to 2π, which is the formal outcome of the invariance of W_ϕ with respect to the rotations:

$$W_{\phi'} = \int d\mathbf{r}e^{iS(\mathbf{r})}\Delta_g(\mathbf{r})\delta[g(\mathbf{r}_{\phi'})] = \int d\mathbf{r}'e^{iS(\mathbf{r}')}\Delta_g(\mathbf{r}'_\phi) = W_\phi, \tag{11.15}$$

where we have introduced the variable $\mathbf{r}' = (r, \phi')$ and used the rotational invariance of $S(\mathbf{r})$, $\Delta_g(\mathbf{r})$ and the integration measure $d\mathbf{r}$. Thus, our "recipe" for extracting the "volume factor" is to introduce into the integrand the constraining δ-function, which is multiplied by Δ_g, defined by (11.9).

Extracting the "volume factor" in a functional integral

Now let us discuss non-Abelian gauge fields. To be concrete, we consider here the case of Yang–Mills fields for a $SU(2)$ gauge group. The Lagrangian of this theory is written as

$$\mathcal{L} = -\frac{1}{16\pi}F^a_{\mu\nu}F^{a\mu\nu}, \quad a = 1, 2, 3, \tag{11.16}$$

where (cf. equation (2.112))

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g\varepsilon^{abc}A^b_\mu A^c_\nu. \tag{11.17}$$

Here g is the Yang–Mills coupling constant. We define the generating functional as usual:

$$Z[\vec{J}] = \int \mathcal{D}\vec{A}_\mu \exp\left\{i\int dx[\mathcal{L}(x) + \vec{J}_\mu \cdot \vec{A}^\mu(x)]\right\}. \tag{11.18}$$

The action is invariant with respect to gauge transformations

$$\vec{A}_\mu \to \vec{A}^\theta_\mu \,, \tag{11.19}$$

where (cf. equations (2.156), (2.159), and (2.162))

$$\vec{A}^\theta_\mu \cdot \frac{\vec{\tau}}{2} = U(\theta)\left[\vec{A}_\mu \cdot \frac{\vec{\tau}}{2} + \frac{1}{ig} U^{-1}(\theta)\partial_\mu U(\theta)\right] U^{-1}(\theta)\,. \tag{11.20}$$

Here

$$U(\theta) = \exp\left[i\vec{\theta}(x)\cdot\frac{\vec{\tau}}{2}\right] \tag{11.21}$$

is the spinor transformation of $SU(2)$. Near the unit transformation, we can write $U(\theta)$ as

$$U(\theta) = 1 + i\vec{\theta}\cdot\frac{\vec{\tau}}{2} + O(\theta^2)\,. \tag{11.22}$$

The values of $\vec{\theta}(x)$ represent the group parameters, depending on the point of space-time, while $\vec{\tau}$ are Pauli matrices in isotopic space.

The action of our theory is constant (invariant) on the *orbit* of the gauge group, consisting of all $\vec{A}^\theta_\mu$, obtained from some fixed field configuration $\vec{A}_\mu$ by transformation $U(\theta)$, which encompasses all elements of the group $SU(2)$. For the correct quantization procedure, functional integration should be done over the "hypersurface" in functional space, which crosses each orbit only once. Thus, if we write the equation for this hypersurface as

$$f_a(\vec{A}_\mu) = 0\,, \quad a = 1,2,3\,; \tag{11.23}$$

equation

$$f_a(\vec{A}^\theta_\mu) = 0 \tag{11.24}$$

should have the unique solution $\vec{\theta}$ for the given field configuration $\vec{A}_\mu$. This condition fixes the gauge.

Let us define integration over the gauge group parameters as

$$[d\vec{\theta}] = \prod_{a=1}^{3} d\theta_a\,. \tag{11.25}$$

If we make two gauge transformations $\vec{\theta}$ and $\vec{\theta}'$, the corresponding matrix is $U(\theta)$ $U(\theta')$, and the transformation parameters are summed: $\vec{\theta} + \vec{\theta}'$. Thus, the integration measure defined as in (11.25) is gauge invariant; if θ angles encompass all possible values, the shift by a constant θ' is irrelevant. Symbolically we write this as $d(\vec{\theta}\vec{\theta}') = d\vec{\theta}'' = d\vec{\theta}$.

Now we can act as above (in the case of usual integral) and introduce $\Delta_f[\vec{A}_\mu]$ as follows:

$$\Delta_f^{-1}[\vec{A}_\mu] = \int [d\vec{\theta}(x)]\delta[f_a(\vec{A}^\theta_\mu)]\,. \tag{11.26}$$

Then we have

$$\Delta_f[\vec{A}_\mu] = \operatorname{Det} M_f \tag{11.27}$$

which is called the Faddeev–Popov determinant, where

$$(M_f)_{ab} = \frac{\delta f_a}{\delta\theta_b}\,. \tag{11.28}$$

In more detail, making the usual discretization of space (followed by the continuous limit), we can write

$$\begin{aligned}\Delta_f^{-1}[\vec{A}_\mu] &= \int \prod_x \prod_a d\theta_a(x)\delta[f_a(x)] = \prod_x \prod_a \int df_a(x)\delta[f_a(x)]\frac{\partial(\theta_1(x),\theta_2(x),\theta_3(x))}{\partial(f_1(x),f_2(x),f_3(x))}\\ &= \prod_x \operatorname{Det}\left[\frac{\partial\theta_a(x)}{\partial f_b(x)}\right]_{f=0} = \operatorname{Det}\left[\frac{\delta\theta_a(x)}{\delta f_b(x)}\right]_{f=0}\,.\end{aligned} \tag{11.29}$$

In the last equality (after transformation to continuous x) the functional determinant (Jacobian) of the matrix with continuous indices $\frac{\delta\theta_a(x)}{\delta f_b(y)}$, appearked which is defined as the product of eigenstates of this matrix.

The matrix M_f is related to the infinitesimal gauge transformations of the function $f_a[\vec{A}_\mu]$:

$$\begin{aligned}f_a[\vec{A}^\theta_\mu(x)] &= f_a[\vec{A}_\mu(x)] + \int dy\frac{\delta f_a(x)}{\delta\theta_b(y)}\theta_b(y) + O(\theta^2)\\ &= f_a[\vec{A}_\mu(x)] + \int dy[M_f(x,y)]_{ab}\theta_b(y) + O(\theta^2)\,.\end{aligned} \tag{11.30}$$

Then, demanding the uniqueness of the solution of equation (11.24) $f_a(\vec{A}^\theta_\mu) = 0$ with respect to $\vec{\theta}$, we conclude that $\operatorname{Det} M_f$ should be nonzero. The explicit form of M_f is naturally dependent on the choice of specific gauge condition (the form of function f_a); concrete expressions for the case for Lorentz gauge will be presented below.

The Faddeev–Popov determinant $\Delta_f[\vec{A}_\mu]$ is gauge invariant. In fact, we can write

$$\Delta_f^{-1}[\vec{A}_\mu] = \int [d\vec{\theta}'(x)]\delta\big[f_a(\vec{A}^{\theta'}_\mu)\big]\,. \tag{11.31}$$

Then

$$\begin{aligned}\Delta_f^{-1}[\vec{A}^\theta_\mu] = \int [d\vec{\theta}'(x)]\delta[f_a(\vec{A}^{\theta\theta'}_\mu)] &= \int [d\vec{\theta}(x)\vec{\theta}'(x)]\delta[f_a(\vec{A}^{\theta\theta'}_\mu)]\\ &= \int [d\vec{\theta}''(x)]\delta[f_a(\vec{A}^{\theta''}_\mu)] = \Delta_f^{-1}[\vec{A}_\mu]\,,\end{aligned} \tag{11.32}$$

which completes the proof. Actually, this situation is similar to (11.12)).

Let us substitute the "unity representation" following from (11.26)

$$1 = \int [d\vec{\theta}(x)]\Delta_f[\vec{A}_\mu]\delta[f_a(\vec{A}^\theta_\mu)] \tag{11.33}$$

into (11.18). Then, denoting the integration measure over the Yang–Mills fields as $[d\vec{A}_\mu(x)]$, we obtain

$$\begin{aligned}\int [d\vec{A}_\mu(x)] \exp\left\{i\int dx\mathcal{L}(x)\right\} \\ = \int [d\vec{\theta}(x)] \int [d\vec{A}_\mu(x)]\Delta_f[\vec{A}_\mu]\delta[f_a(\vec{A}^\theta_\mu)] \exp\left\{i\int dx\mathcal{L}(x)\right\} \\ = \int [d\theta(x)] \int [d\vec{A}_\mu(x)]\Delta_f[\vec{A}_\mu(x)]\delta[f_a(\vec{A}_\mu)] \exp\left\{i\int dx\mathcal{L}(x)\right\} . \end{aligned} \tag{11.34}$$

To obtain the last equality we used the invariance of $\Delta_f[\vec{A}_\mu]$ and $\exp\{i\int dx\mathcal{L}(x)\}$ to gauge transformations $\vec{A}^\theta_\mu \to \vec{A}_\mu$. Then we see that the integrand does not depend on $\vec{\theta}(x)$, and $\int[d\vec{\theta}(x)] = \int\prod_x d\vec{\theta}(x)$ simply gives infinite "volume" of the orbit, which we wanted to separate! Thus, dropping this irrelevant factor, we can write the generating functional for gauge field $\vec{A}_\mu$ as

$$\begin{aligned} Z[\vec{J}] = \int [d\vec{A}_\mu]\Delta_f[\vec{A}_\mu]\delta[f_a(\vec{A}_\mu)] \exp\left\{i\int dx[\mathcal{L}(x) + \vec{J}^\mu\cdot\vec{A}_\mu]\right\} \\ = \int [d\vec{A}_\mu](Det M_f)\delta[f_a(\vec{A}_\mu)] \exp\left\{i\int dx[\mathcal{L}(x) + \vec{J}^\mu\cdot\vec{A}_\mu]\right\} . \end{aligned} \tag{11.35}$$

This is the essence of the so-called *Faddeev–Popov Ansatz* – we isolate and cancel all the irrelevant integrations, introducing into integration measure the factor of Det $M_f\delta[f_a(\vec{A}_\mu)]$.

Abelian gauge theory (QED)

Consider the simplest example – QED. In this case infinitesimal gauge transformation is written as

$$A^\theta_\mu = A_\mu(x) - \frac{1}{g}\partial_\mu\theta(x) . \tag{11.36}$$

For any choice of the gauge condition (11.23) linear over field $A_\mu(x)$, the matrix M_f (11.28) is *independent of field* $A_\mu(x)$. Then, Faddeev-Popov determinant is unimportant from physical point of view and can be moved outside functional integral over $A_\mu(x)$ and just dropped[3]. Then we can write generating functional as:

$$Z[J] = \int [dA_\mu]\delta[f(A_\mu)] \exp\left\{i\int dx[\mathcal{L}(x) + J_\mu(x)A^\mu(x)]\right\} , \tag{11.37}$$

where $\delta[f(A_\mu)]$ fixes the gauge and we obtain the usual formulation of QED.

[3] In terms of general case, discussed below, we can say, that in QED Faddeev-Popov "ghosts" do not interact with field A_μ and are irrelevant.

11.2 Feynman diagrams for non-Abelian theory

Consider now the details of the diagram technique for non-Abelian theory. Let us rewrite the generating functional (11.35) as

$$Z[\vec{J}] = \int [d\,\vec{A}_\mu] \exp\left\{ i S_{eff} + i \int dx \vec{J}^\mu \cdot \vec{A}_\mu \right\}, \tag{11.38}$$

where we have rewritten the factor $\mathrm{Det}\, M_f \delta[f_a(\vec{A}_\mu)]$ as $\exp \ln(\mathrm{Det}\, M_f \delta[f_a(\vec{A}_\mu)])$, and included $-i \ln(\mathrm{Det}\, M_f \delta[f_a(\vec{A}_\mu)])$ into the definition of effective action S_{eff}. Naturally, the presence of such a term in effective action complicates the construction of the diagram technique. First we shall try to write this term in a more natural and convenient form.

Faddeev–Popov "ghosts"

We can write $\mathrm{Det}\, M_f$ as an exponential, using the expression

$$\mathrm{Det}\, M_f = \exp[Sp \ln M_f]\,. \tag{11.39}$$

The proof of (11.39) is trivial. The equality $\ln \mathrm{Det}\, M_f = Sp \ln M_f$ is obvious for any matrix: $\mathrm{Det}\, M_f$ is represented by the product of eigenvalues of M_f, so that $\ln \mathrm{Det}\, M_f$ gives the sum of the logarithms of all eigenvalues of M_f, i. e., precisely $Sp \ln M_f$.

Writing the matrix M_f as

$$M_f = 1 + L \tag{11.40}$$

and expanding the logarithms, we have

$$\begin{aligned} \exp[Sp \ln M_f] &= \exp\left[SpL - \frac{1}{2} SpL^2 + \cdots + \frac{(-1)^{n+1}}{n} SpL^n + \cdots \right] \\ &= \exp\left\{ \int dx L_{aa}(x,x) - \frac{1}{2} \int dx \int dy L_{ab}(x,y) L_{ba}(y,x) + \cdots \right\}. \end{aligned} \tag{11.41}$$

We see that the Faddeev–Popov determinant can be represented as a *loop* expansion[4], as shown in Figure 11.2, where lines denote the propagators of some fictitious particles (Faddeev–Popov "ghosts"), forming a triplet of complex scalar (spinless) fields $\vec{c}(x)$. These fields and their interactions can be described by the generating functional

$$\mathrm{Det}\, M_f = \int [d\vec{c}][d\vec{c}^{+}] \exp\left\{ i \int dx dy \sum_{ab} c_a^{+}(x) [M_f(x,y)]_{ab} c_b(y) \right\}. \tag{11.42}$$

[4] This expansion is similar to the loop expansion of free energy in the theory of condensed matter [1]

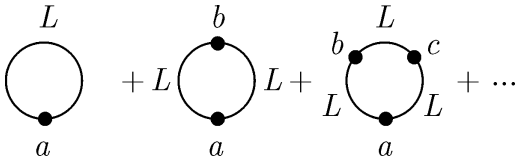


Figure 11.2

Here integration is done over *Grassmannian* $\vec{c}(x), \vec{c}^{+}(x)$, as the use of common c-number fields will lead to $(\mathrm{Det}\, M_f)^{-1}$! Thus, our *scalar* fields $\vec{c}(x), \vec{c}^{+}(x)$ obey *Fermi statistics*, and the Faddeev–Popov "ghosts" are fermions with spin)! There is no contradiction with the spin and statistics theorem here, because these "ghosts" are purely fictitious particles which are introduced to the theory just "for convenience". As their contribution to the generating functional reduces to the loop series (11.41), there are no diagrams with external "ghost" lines.

Gauge-fixing terms

Now we shall transform to the exponential form the term $\delta[f_a(\vec{A}_\mu)]$. First of all, we generalize the gauge-fixing condition, writing it as

$$f_a[\vec{A}_\mu] = B_a(x)\,, \quad a = 1,2,3\,, \tag{11.43}$$

where $B_a(x)$ is some arbitrary function of space-time point, *independent* of gauge field $\vec{A}_\mu$. Correspondingly we define Δ_f by the condition

$$\Delta_f[\vec{A}_\mu] \int [d\,\vec{\theta}(x)]\delta[f_a(\vec{A}^{\theta}_\mu) - B_a(x)] = 1\,. \tag{11.44}$$

Obviously, because of the independence of $B_a(x)$ from $\vec{A}_\mu$, this is the *same* function Δ_f defined in (11.26)[5]; in fact there is no dependence on $B_a(x)$ here at all! Thus the generating functional (11.35) can be rewritten as

$$\begin{aligned} Z[\vec{J}] = &\int [d\,\vec{A}_\mu][d\,\vec{B}](\mathrm{Det}\, M_f)\delta[f_a(\vec{A}_\mu) - B_a] \\ &\times \exp\left\{ i \int dx \left[\mathcal{L}(x) - \vec{J}^{\mu} \cdot \vec{A}_\mu - \frac{1}{8\pi\xi}\vec{B}^2(x) \right] \right\}\,, \end{aligned} \tag{11.45}$$

where we have included in the integrand the *constant* term like

$$\int [d\,\vec{B}] \exp\left\{ -\frac{i}{8\pi\xi} \int dx\, \vec{B}^2(x) \right\}\,, \tag{11.46}$$

[5] This is simply the analogue of (11.6) in the case of usual integration.

where ξ is an arbitrary constant coefficient which is usually called the gauge parameter. As a result, the generating functional (11.45) differs from (11.35) by an irrelevant constant factor which can be hidden into normalization. But now, using the δ-function, entering (11.45), we can lift integration over $[d\,\vec{B}(x)]$. Finally, taking into account also (11.42), we obtain

$$Z[\vec{J}] = \int [d\,\vec{A}_\mu][d\,\vec{c}][d\,\vec{c}^{\,+}] \exp(i S_{eff}[\vec{J}])\,, \tag{11.47}$$

where

$$S_{eff}[\vec{J}] = S[\vec{J}] + S_{fix} + S_{ghost}\,, \tag{11.48}$$

where $S[\vec{J}] = \int dx[\mathcal{L}(x) + \vec{J}^\mu \cdot \vec{A}_\mu]$ is the usual action of our theory,

$$S_{fix} = -\frac{1}{8\pi\xi} \int dx\{f_a[\vec{A}_\mu(x)]\}^2 \tag{11.49}$$

is the so-called gauge fixing term, and

$$S_{ghost} = \int dx dy \sum_{ab} c_a^+(x)[M_f(x,y)]_{ab} c_b(y) \tag{11.50}$$

is "ghosts" action.

The Lorentz gauge

In the Lorentz gauge we have

$$f_a(\vec{A}_\mu) \equiv \partial^\mu A_\mu^a = 0\,, \quad a = 1,2,3\,. \tag{11.51}$$

Under infinitesimal gauge transformations

$$U(\theta) = 1 + i\,\vec{\theta}(x) \cdot \frac{\vec{\tau}}{2} + O(\theta^2)\,, \tag{11.52}$$

so that

$$A_\mu^{a\theta} = A_\mu^a(x) - \varepsilon^{abc}\theta^b(x) A_\mu^c(x) + \frac{1}{g}\partial_\mu \theta^a(x)\,. \tag{11.53}$$

Substituting (11.53) into (11.51) we have

$$\begin{aligned} f_a(\vec{A}_\mu^\theta) &= f_a(\vec{A}_\mu) - \partial^\mu \left[\varepsilon^{abc}\theta^b(x) A_\mu^c(x) - \frac{1}{g}\partial_\mu \theta^a(x)\right] \\ &= f_a(\vec{A}_\mu) + \int dy [M_f(x,y)]_{ab}\theta^b(y)\,, \end{aligned} \tag{11.54}$$

where in the last equality we have used (11.30). Then we see that in this case

$$[M_f(x,y)]_{ab} = \frac{1}{g}\partial^\mu[\delta^{ab}\partial_\mu - g\varepsilon^{abc} A_\mu^c]\delta(x-y)\,. \tag{11.55}$$

Now, substituting everything into (11.49) and (11.50), we obtain

$$S_{fix} = -\frac{1}{8\pi\xi}\int dx(\partial^\mu \vec{A}_\mu)^2\,, \tag{11.56}$$

$$S_{ghost} = \frac{1}{g}\int dx \sum_{ab} c_a^+(x)\partial^\mu[\delta_{ab}\partial_\mu - g\varepsilon_{abc}A_\mu^c]c_b(x)\,. \tag{11.57}$$

We see that now "ghosts" are interacting with the gauge field $\vec{A}_\mu$, which is described by the second term in square brackets in (11.57). In similar expressions in QED such a term was just absent[6].

Let us also introduce the sources η_a^+, η_a for "ghost" fields c_a, c_a^+ and write the generating functional of gauge theory as

$$Z[\vec{J},\vec{\eta},\vec{\eta}^+] = \int [d\vec{A}_\mu d\vec{c} d\vec{c}^+]\exp\left\{i\int dx\Big[\mathcal{L}(x) - \frac{1}{8\pi\xi}(\partial^\mu \vec{A}_\mu)^2 + c_a^+\partial^\mu(\delta_{ab}\partial_\mu - g\varepsilon_{abc}A_\mu^c)c_b + J_\mu^a A^{a\mu} + \eta^{a+}c^a + \eta^a c^{a+}\Big]\right\}, \tag{11.58}$$

where we redefined fields c_a, c_a^+ in an obvious way by including the factor $1/g$.

Perturbation expansion

Let us write the action of our theory as

$$S_{eff} = S_0 + S_I\,, \tag{11.59}$$

where

$$S_0 = \int dx\Big[-\frac{1}{16\pi}(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a)^2 - \frac{1}{8\pi\xi}(\partial^\mu A_\mu^a)^2 + c_a^+\partial^2 c_a + J_\mu^a A^{a\mu} + \eta^{a+}c^a + \eta^a c^{a+}\Big], \tag{11.60}$$

and the interaction term, containing fields in powers higher than two, has the form

$$S_I = \int dx\Big[-\frac{1}{8\pi}(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a)g\varepsilon^{abc}A^{b\mu}A^{c\nu} - \frac{1}{16\pi}g^2\varepsilon^{abc}\varepsilon^{ade}A_\mu^b A_\nu^c A^{d\mu}A^{e\nu} - gc^{a+}\partial^\mu\varepsilon^{abc}A_\mu^c c^b\Big]. \tag{11.61}$$

[6] In the non-Abelian case it is also possible to choose a special, so-called axial, gauge, where "ghosts" are completely excluded [13], but this gauge is rather inconvenient for practical calculations due to a very complicated form of the gauge field propagator.

Generating functional can now be written as

$$Z[\vec{J},\vec{\eta},\vec{\eta}^{+}] = \exp\left\{ i S_I\left[\frac{\delta}{i\delta\vec{J}_\mu}, \frac{\delta}{i\delta\vec{\eta}}, \frac{\delta}{i\delta\vec{\eta}^{+}}\right]\right\} Z_A^0[\vec{J}] Z_c^0[\vec{\eta},\vec{\eta}^{+}], \tag{11.62}$$

where

$$Z_A^0[\vec{J}] = \int [d\vec{A}_\mu] \exp\left\{ i \int dx \left[-\frac{1}{16\pi}(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a)^2 - \frac{1}{8\pi\xi}(\partial^\mu A_\mu^a)^2 + J_\mu^a A^{a\mu}\right]\right\}, \tag{11.63}$$

$$Z_c^0[\vec{\eta},\vec{\eta}^{+}] = \int [d\vec{c}^{+}][d\vec{c}] \exp\left\{ i \int dx \left[c^{a+}\partial^2 c^a - \eta^{a+} c^a - \eta^a c^{a+}\right]\right\}. \tag{11.64}$$

These expressions allow the direct derivation of perturbation theory. We shall not give the detailed derivation here, but limit ourselves to the summary of the main rules of the diagram technique. Readers interested in details can find everything in a number of books , e.g., [13, 25, 28, 53, 56].

Propagators

To find the propagator of field $\vec{A}_\mu$ we rewrite Z_A^0 as

$$\begin{aligned} Z_A^0[\vec{J}] &= \int [d\vec{A}_\mu] \exp\left\{ i \int dx \left[\frac{1}{8\pi}\delta_{ab} A_\mu^a \left(g^{\mu\nu}\partial^2 - \frac{\xi-1}{\xi}\partial^\mu\partial^\nu\right) A_\nu^b + J_\mu^a A^{a\mu}\right]\right\} \\ &= \int [d\vec{A}_\mu] \exp\left\{ i \int dx \left[\frac{1}{2} A_\mu^a K_{ab}^{\mu\nu} A_\nu^b + J_\mu^a A^{a\mu}\right]\right\}, \end{aligned} \tag{11.65}$$

where

$$K_{ab}^{\mu\nu} = \left[g^{\mu\nu}\partial^2 - \left(1 - \frac{1}{\xi}\right)\partial^\mu\partial^\nu\right]\delta_{ab}. \tag{11.66}$$

Integration over $[d\vec{A}_\mu]$ can be performed using the well-known Gaussian integral (10.47), which in this case can be written as

$$\int [d\varphi] \exp\left[-\frac{1}{2}\langle\varphi K\varphi\rangle + \langle J\varphi\rangle\right] \sim (\mathrm{Det}\, K)^{-1/2} \exp\langle JK^{-1}J\rangle, \tag{11.67}$$

where angular brackets denote the appropriate integrals. Application of this expression to (11.65) gives

$$Z_A^0[\vec{J}] = \exp\left\{ -\frac{i}{2}\int dx dy J_\mu^a(x) G_{ab}^{\mu\nu}(x-y) J_\nu^b(y)\right\}, \tag{11.68}$$

where

$$G_{ab}^{\mu\nu}(x-y) = \delta^{ab}\int \frac{d^4k}{(2\pi)^4} e^{-ik(x-y)}\left[-\left(g^{\mu\nu} - \frac{k^\mu k^\nu}{k^2}\right) - \xi\frac{k^\mu k^\nu}{k^2}\right]\frac{4\pi}{k^2 + i\varepsilon}. \tag{11.69}$$

It is easy to check that

$$\int dy K^{\mu\nu}_{ab}(x-y) G^{bc}_{\nu\lambda}(y-z) = g^{\mu}_{\lambda}\delta^{c}_{a}\delta(x-z)\,, \tag{11.70}$$

so that the propagator G is the inverse of K.

Similarly we find

$$Z^{0}_{c}[\vec{\eta},\vec{\eta}^{+}] = \exp\left\{-i\int dx dy \eta^{a+}(x) G^{ab}(x-y)\eta^{a}(y)\right\}\,, \tag{11.71}$$

where

$$G^{ab}(x-y) = -\int \frac{d^4k}{(2\pi)^4}\frac{e^{-ik(x-y)}}{k^2+i\varepsilon}\delta_{ab}\,. \tag{11.72}$$

This directly corresponds to the fact, that "ghosts" are scalar particles with zero mass (obeying Fermi statistics).

Finally, we have

1. the propagator of massless vector bosons

$$i\Delta^{ab}_{\mu\nu}(k) = -i\delta_{ab}\left[g_{\mu\nu} - (1-\xi)\frac{k_\mu k_\nu}{k^2}\right]\frac{4\pi}{k^2+i\varepsilon}\,, \tag{11.73}$$

   denoted in diagrams by a wave-like line;

2. the propagator of Faddeev-Popov "ghosts"

$$i\Delta^{ab}(k) = -i\delta_{ab}\frac{4\pi}{k^2+i\varepsilon}\,, \tag{11.74}$$

   which is denoted by a dashed line with an arrow (a "ghost" is different from an "anti-ghost"!).

Elementary vertices

In non-Abelian gauge theories there are two types of self-interactions, which can be written as

$$\varepsilon^{\mu}_{a}(k_1)\varepsilon^{\nu}_{b}(k_2)\varepsilon^{\lambda}_{c}(k_3)\Gamma^{abc}_{\mu\lambda\nu}(k_1,k_2,k_3)\,, \tag{11.75}$$

$$\varepsilon^{\mu}_{a}(k_1)\varepsilon^{\nu}_{b}(k_2)\varepsilon^{\lambda}_{c}(k_3)\varepsilon^{\rho}_{d}(k_4)\Gamma^{abcd}_{\mu\lambda\nu\rho}(k_1,k_2,k_3,k_4)\,, \tag{11.76}$$

where we also explicitly show the corresponding polarization vectors. Feynman rules are derived directly from (11.61), (11.62). In momentum representation the first term in (11.61) can be written as

$$\frac{1}{3!}A^{a\mu}(k_1)A^{b\nu}(k_2)A^{c\lambda}(k_3)\Gamma^{abc}_{\mu\nu\lambda}(k_1,k_2,k_3)\,. \tag{11.77}$$

The vertex part $\Gamma^{abc}_{\mu\nu\lambda}$ shouldbe completely antisymmetric with respect to permutations

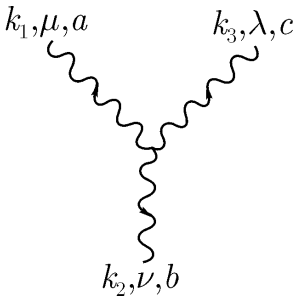


Figure 11.3

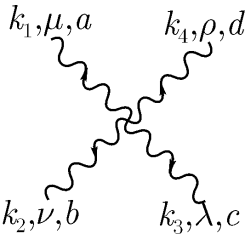


Figure 11.4

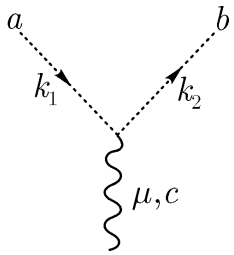


Figure 11.5

of fields A. The structure related to $SU(2)$ gauge group is already fixed:

$$\Gamma^{abc}_{\mu\nu\lambda}(k_1,k_2,k_3) = \varepsilon^{abc}\Gamma_{\mu\nu\lambda}(k_1,k_2,k_3)\,, \tag{11.78}$$

while the Lorentzian (relativistic) structure of this function can be obtained as follows. From (11.61) it is clear that $\Gamma_{\mu\nu\lambda}(k_1,k_2,k_3)$ consists of terms like $k_\mu g_{\nu\lambda}$. A precise combination of these terms can be established from the requirement of antisymmetry of $\Gamma_{\mu\nu\lambda}(k_1,k_2,k_3)$ with respect to permutations of the indices: $\mu,\nu,1,2$ etc., taking into account the total antisymmetry of the tensor ε^{abc}. Thus, we find

$$i\,\Gamma^{abc}_{\mu\nu\lambda} = ig\varepsilon^{abc}[(k_1-k_2)_\lambda g_{\mu\nu} + (k_2-k_3)_\mu g_{\nu\lambda} + (k_3-k_1)_\nu g_{\mu\lambda}] \tag{11.79}$$

where $k_1+k_2+k_3=0$. The corresponding diagram for a "triple" vertex is shown in Figure 11.3.

Similarly we can find the vertex of "quartic" interaction of the gauge field, corresponding to the second term in (11.61):

$$\begin{aligned} i\,\Gamma^{abcd}_{\mu\nu\lambda\rho} = ig^2[&\varepsilon^{abe}\varepsilon^{cde}(g_{\mu\lambda}g_{\nu\rho} - g_{\nu\lambda}g_{\mu\rho}) + \varepsilon^{ace}\varepsilon^{bde}(g_{\mu\nu}g_{\lambda\rho} - g_{\lambda\nu}g_{\mu\rho}) \\ &+ \varepsilon^{ade}\varepsilon^{cbe}(g_{\mu\lambda}g_{\rho\nu} - g_{\rho\lambda}g_{\mu\nu})] \end{aligned} \tag{11.80}$$

which is expressed by the diagram in Figure 11.4. Here $k_1+k_2+k_3+k_4=0$.

For the vertex connecting the "ghosts" and gauge fields with polarization vector $\varepsilon^\mu(q)$ we have

$$i\,\Gamma^{abc}_{\mu} = ig\varepsilon^{abc}k_{2\mu} \tag{11.81}$$

where $k_2 = k_1 + q$. This vertex is shown in Figure 11.5; it is antisymmetric over the isospin indices. Let us recall that "ghost" lines enter diagrams only in loops. Besides each diagram containing a closed loop of the gauge field, there is a corresponding diagram with a closed "ghost" line. As in the case of the usual fermion fields, each "ghost" loop can be multiplied by an additional (-1).

The ropagator of gauge field (11.69) depends on the gauge parameter ξ. Its value is chosen for the convenience of explicit calculations while solving concrete problems.

For example, $\xi = 1$ corresponds to the so-called t'Hooft–Feynman gauge, while $\xi = 0$ gives the Landau gauge.

The introduction of fermions into the Yang–Mills theory is not difficult: it is sufficient to add to the Lagrangian gauge invariant terms such as

$$\mathcal{L}_f = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi \tag{11.82}$$

where

$$D_\mu \psi = \partial_\mu \psi - igT^a A^a_\mu \psi \tag{11.83}$$

Here T^a is the gauge group generator in the given representation. For example, if ψ is a $SU(2)$ doublet, we have $T^a = \tau^a/2$. Thus we obtain additional Feynman rules for fermions (with group indices $n, m, \ldots$):

1. the fermion propagator has the standard form

$$i\Delta_{mn}(k) = \delta_{nm}\frac{1}{\gamma_\mu k^\mu - m + i\varepsilon} \tag{11.84}$$

   and is expressed by continuous line;

2. the fermion-gauge field interaction vertex has the form

$$i\Gamma^{\alpha\mu}_{nm} = ig(T^a)_{nm}\gamma^\mu. \tag{11.85}$$

This is shown diagrammatically in Figure 11.6.

The structure of the diagram technique described above is also conserved for the other gauge groups, such as the very important $SU(3)$ describing the color symmetry of quarks. The only difference is in the dimensionality of the corresponding irreducible representations and the explicit form of generator matrices.

So far we have already studied the basics of the modern theory of quantum gauge fields, which forms the foundations of the standard model of elementary particles. Now we will begin out discussion on specific models of interactions. However, some conceptual problems, which we have discussed from the beginning, still remain. In particular, so far it is still unclear how we should deal with the problem of the massless nature of Yang–Mills fields, which is in striking contrast to experiments, which clearly demonstrate that the only long-range interaction in nature (except gravitation)

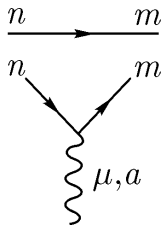


Figure 11.6

is electromagnetism. In the next chapter we shall see how this problem is solved in the unified theory of weak and electromagnetic interactions. The remarkable fact is that this solution is completely based on the ideas and methods originating from the modern theory of condensed matter.

Chapter 12

The Weinberg–Salam model

12.1 Spontaneous symmetry-breaking and the Goldstone theorem

As we have already noted, the significant progress in modern theory of elementary particles was achieved using some fundamental concepts of the modern theory of the condensed state. Most important was the introduction into quantum field theory of the idea of the possibility of *phase transitions* when the symmetry of the ground state becomes lower than the symmetry of the Lagrangian. This allowed the effective solution of the problem of mass generation for gauge fields without breaking local gauge invariance, directly leadto to quite rich and nontrivial foundations of the Standard Model. Moreover, the picture of possible "vacuum" phase transitions form the basis of modern cosmology and physics of matter at very high densities and temperatures. Here we shall limit ourselves to a presentation of some of the main ideas which played a decisive role during the construction of the unified theory of weak and electromagnetic interactions[1].

Let us again begin with the simplest example of the real scalar self-interacting field $\varphi(x)$, described by the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \varphi)^2 - V(\varphi) = \frac{1}{2}(\partial_t \varphi)^2 - \frac{1}{2}(\nabla \varphi)^2 - V(\varphi)\,, \tag{12.1}$$

where $V(\varphi)$ is some function of field invariants. The first term in this expression may be considered to be kinetic energy density, and all the others represent potential energy density.

From (12.1) we obtain the equations of motion

$$\partial_\mu^2 \varphi = -\frac{\partial V(\varphi)}{\partial \varphi} \quad \text{or}$$

$$\partial_t^2 \varphi - \nabla^2 \varphi = -\frac{\partial V(\varphi)}{\partial \varphi}\,. \tag{12.2}$$

The character of the solutions of these field equations depends essentially on the form of "potential energy" of self-interaction $V(\varphi)$. Consider first the case of the traditional field theory, which was analyzed before. Let $V(\varphi)$ be of the form shown in Figure 12.1(a). Then our system has the "stable equilibrium" state with $\varphi = 0$ and can

[1] In this chapter we mainly follow [5].

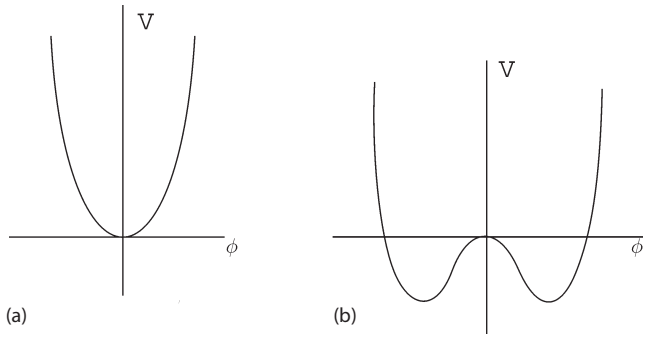


Figure 12.1

oscillate around it. Close enough to this equilibrium state we can always write

$$V(\varphi) \approx \frac{\mu^2}{2}\varphi^2 , \tag{12.3}$$

where $\mu^2 = \left(\frac{\partial^2 V}{\partial \varphi^2}\right)_{\varphi=0}$, so that (12.2) reduces to:

$$\partial_\mu^2 \varphi + \mu^2 \varphi = 0 , \tag{12.4}$$

i. e., to the Klein–Gordon equation. If we are looking for the plane-wave solution of this equation $\varphi \sim e^{ikx}$, from (12.4), the usual relativistic dispersion of a particle with mass μ: $k_0^2 = \mathbf{k}^2 + \mu^2$ immediately follows. The higher-order terms of expansion of $V(\varphi)$ lead to nonlinear terms in field equations, describing interactions of these plane waves or particle scattering. Let us limit ourselves to

$$V(\varphi) = \frac{1}{2}\mu^2\varphi^2 + \frac{1}{4}\lambda\varphi^4 . \tag{12.5}$$

There is no cubic term here, as $V(\varphi)$ should be symmetric with respect to $\varphi \to -\varphi$, so that there is always a minimum of $V(\varphi)$ at $\varphi = 0$. Limitation to powers not higher than $\sim \varphi^4$ is of conceptual importance, as the coupling constant $\lambda > 0$ is dimensionless and the theory is renormalizable. Thus, in this case we are dealing with the well-known φ^4-theory.

Consider now the case of $\mu^2 < 0$. This can seem strange, as from the naïve point of view we are dealing with imaginary mass. However, we need to be more accurate. Now $\varphi = 0$ is no longer a stable equilibrium, as the potential energy has the form shown in Figure 12.1(b)[2]. We now see *two* stable equilibrium states, corresponding to

$$\varphi = \eta = \pm\sqrt{\frac{|\mu|^2}{\lambda}} . \tag{12.6}$$

[2] The situation here is completely analogous to the Landau theory of phase transitions, where $\mu^2 \sim T - T_c$, so that $\mu^2 < 0$ for $T < T_c$, i. e., below the phase transition temperature.

Expanding $V(\varphi)$ around points (12.6) up to the quadratic terms, we have

$$V(\varphi) = \frac{\mu^2\eta^2}{4} - \mu^2(\varphi - \eta)^2 = \frac{\mu^2\eta^2}{4} - \mu^2(\delta\varphi)^2\,, \tag{12.7}$$

where $\delta\varphi = \varphi - \eta$ and $-\mu^2(\delta\varphi)^2 > 0$ because of $\mu^2 < 0$. Now we see that field equations (12.2) will have plane-wave solutions for $\delta\varphi$ with wave-vector k, satisfying the condition (determining dispersion) $k^2 = 2|\mu|^2$, so that these waves will correspond to particles with real *positive* mass $m = \sqrt{2|\mu|^2}$.

In fact, here we are dealing with a phase transition in quantum field theory. The system chooses *one* of two equilibrium positions in Figure 12.1(b), and the field oscillates close to this *new ground state*.

In quantum mechanics the system with two such minima of potential energy does not only oscillates around the single minimum, because of the possibility of quantum tunneling between these two equilibrium positions. The quantum state is split in two: the symmetric and the antisymmetric (with respect to these minima) states. The ground state corresponds to the symmetric state [35]. Thus, in quantum mechanics the symmetry of the ground state is in complete agreement with the symmetry of the Lagrange function (in our case even in φ). Quantum field theory in this sense is reminiscent of classical mechanics. Actually, the probability of quantum tunneling transition becomes less with the growth of the number of degrees of freedom, and becomes zero in the case of their number being infinite. In fact, let us consider the field in a finite volume Ω. Then the Lagrange function is $L = \int d^3x\mathcal{L} \sim \mathcal{L}\Omega$, so that the corresponding kinetic energy $\sim \Omega\dot{\varphi}^2$, while the potential is $\sim \Omega V(\varphi)$. Thus, our problem is equivalent to the tunneling of a particle with mass $M \sim \Omega$ through the potential barrier of width $|x| \sim \eta$ and height $\mathcal{V} \sim \Omega m^2\eta^2$. The probability of such tunnelling transition [35] is of the order of $\exp(-\sqrt{2M\mathcal{V}}|x|) \sim \exp(-\Omega m\eta^2) \to 0$ for $\Omega \to \infty$. We can say that our field in the ground state is represented by a macroscopic "string" or "rope" of infinite length, lying in the left or right valley of the potential in Figure 12.1(b), along the whole valley, which is perpendicular to the plane in the figure. Naturally, such an object cannot tunnel between the valleys of the potential.

In quantum field theory the ground state is called a vacuum. Thus, we have to choose the single ground state: one definite vacuum. The presence of another vacuum (physically equivalent to the first one) is now irrelevant. Two minima of $V(\varphi)$ correspond to two separate and vacuums of the theory which are orthogonal to each other, two orthogonal Hilbert spaces of excited states, two separate "worlds".

Traditional quantum field theory, corresponding to the potential $V(\varphi)$ shown in Figure 12.1(a), is constructed, as we know, as follows. The field is represented by the sum of the oscillators which are described by the creation and annihilation operators $a^\dagger$ and a, and the vacuum is the state without the particles $a|0\rangle = 0$, so that

$$\langle 0|\varphi|0\rangle = 0\,. \tag{12.8}$$

In the case of potential $V(\varphi)$, shown in Figure 12.1(b), the sum over the oscillators represents not the field φ itself, but its deviation from the equilibrium $\delta\varphi = \varphi - \eta$. In this case

$$\langle 0|\varphi|0\rangle = \eta\,, \tag{12.9}$$

i. e., the vacuum average of the field operator is nonzero: the system acquires the *Bose-condensate*[3] of particles corresponding to field φ. The initial Lagrangian (12.1), (12.5) is symmetric with respect to $\varphi \to -\varphi$. However, for $\mu^2 < 0$ it leads to a nonsymmetric ground state (vacuum), which is expressed by (12.9). Excitations above this vacuum also do not possess the symmetry of the initial Lagrangian, as $V(\varphi)$ from Figure 12.1(b) is nonsymmetric with respect to $\varphi = \eta$. This phenomenon in quantum field theory is called *spontaneous symmetry-breaking*, while in condensed matter theory this is the well-known situation of *phase transition* into the state with lower symmetry.

The mechanism of mass generation

The existence of the nonzero vacuum average of a scalar field can automatically lead to the generation of *mass* of an initially massless particle, which interacts with this field. Consider as an example Dirac's field of massless particles with spin 1/2. The Lagrangian of this field has the form

$$\mathcal{L} = i\bar{\psi}_L\hat{\partial}\psi_L + i\bar{\psi}_R\hat{\partial}\psi_R\,, \tag{12.10}$$

where $\hat{\partial} = \gamma^\mu\partial_\mu$ and we introduced "left" and "right" components of bispinor ψ:

$$\psi_R = \frac{1}{2}(1+\gamma^5)\psi\,, \quad \psi_L = \frac{1}{2}(1-\gamma^5)\psi\,, \quad \psi_L + \psi_R = \psi\,. \tag{12.11}$$

Now we can introduce the interaction of fields ψ_L, ψ_R with our scalar field φ breaking the symmetry of the ground state. Let us add to the Lagrangian (12.10) the term

$$\mathcal{L}_{int} = -\varkappa[\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L]\varphi\,, \tag{12.12}$$

where an expression in square brackets represents the only scalar which can be constructed from ψ_L and ψ_R, while $\varkappa$ is a dimensionless coupling constant (so that this interaction is renormalizable). Let us replace field φ in (12.12) by its vacuum average η; this means that we are not taking into account the particle creation processes for field φ. Then we have

$$\mathcal{L}_{int} = -\varkappa\eta(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L) = -\varkappa\eta\bar{\psi}\psi\,, \tag{12.13}$$

so that the sum of (12.10) and (12.13) gives

$$\mathcal{L} = i\bar{\psi}\hat{\partial}\psi - m\bar{\psi}\psi\,, \tag{12.14}$$

[3] Compare with Bogolyubov's approach to nonideal Bose gas!

which corresponds to the Dirac Lagrangian for fermions with mass

$$m = \varkappa\eta\,. \tag{12.15}$$

Thus, we can start with the model of initially massless "left" and "right" fermions, which interact with scalar field φ, undergoing the phase transition and transforming "left" particles into "right" ones and vice versa, and leading to the generation of mass.

Above we considered the simplest example of a Lagrangian with discrete symmetry with respect to $\varphi \to -\varphi$. Let us now consider the case of continuous symmetry-breaking. To do this we introduce the complex scalar field φ, which is equivalent to two real fields φ_1, φ_2, related by (see Chapter 2)

$$\varphi(x) = \frac{1}{\sqrt{2}}[\varphi_1(x) + i\varphi_2(x)] \tag{12.16}$$

the Lagrangian of this field can be written as

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\varphi_1)^2 + \frac{1}{2}(\partial_\mu\varphi_2)^2 - V(\varphi_1,\varphi_2) = (\partial_\mu\varphi)(\partial^\mu\varphi^*) - V(\varphi_1,\varphi_2)\,. \tag{12.17}$$

Let us assume that the potential $V(\varphi_1,\varphi_2)$ depends only on the absolute value of φ, i.i. on $\rho^2 = \varphi_1^2 + \varphi_2^2 = 2\varphi^*\varphi$, so that $V = V(\rho)$. This is equivalent to the requirement of an additional ("internal") symmetry of the theory with respect to transformations of group $U(1)$:

$$\varphi \to e^{i\alpha}\varphi \tag{12.18}$$

or, which is just the same, the invariance of the Lagrangian with respect to rotations in an "isotopic" plane:

$$\begin{aligned} &\varphi_1 \to \varphi_1\cos\alpha - \varphi_2\sin\alpha\,,\\ &\varphi_2 \to \varphi_1\sin\alpha + \varphi_2\cos\alpha\,. \end{aligned} \tag{12.19}$$

We have seen (see Chapter 2), that this symmetry determines the conservation of some charge (electric, barion, etc.). Fields φ and φ^* have the opposite values of this charge.

Consider now the potential $V(\rho)$, shown in Figure 12.2. which can be modelled, for example, by

$$V(\rho) = \frac{1}{2}\mu^2\rho^2 + \frac{1}{4}\lambda\rho^4 \tag{12.20}$$

with $\mu^2 < 0$. Writing the field as (modulus-phase representation),

$$\varphi(x) = \frac{1}{\sqrt{2}}\rho(x)e^{i\vartheta(x)}\,, \tag{12.21}$$

where $\rho(x)$ and $\vartheta(x)$ are real functions, we can see that $V(\rho)$ has the minimum at $\rho = \eta = \sqrt{\frac{|\mu^2|}{\lambda}}$, i. e., for field values

$$\varphi = \frac{1}{\sqrt{2}}\eta e^{i\alpha} \tag{12.22}$$

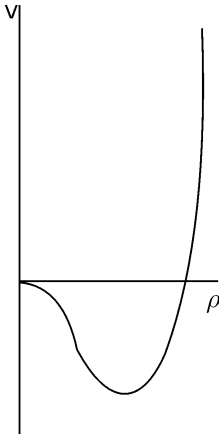


Figure 12.2

with *arbitrary* α! Here we have continuous degeneracy of the ground state with different values of α. Each value of α corresponds to its own vacuum (ground state) with the same (minimal) energy $V(\eta)$. All these vacuums are physically equivalent, but we have to choose the only one, e. g., corresponding to $\alpha = 0$, and the Hilbert space of states associated with this single vacuum, where we already have no $U(1)$ symmetry (12.18), (12.19).

Let us see which particles are describe by the Lagrangian (12.17). Using (12.21) we can rewrite the Lagrangian as

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \rho)^2 - V(\rho) + \frac{\rho^2}{2}(\partial_\mu \vartheta)^2 . \tag{12.23}$$

If we limit ourselves in (12.23) to terms which are quadratic in a field, we have to expand $V(\rho)$ around $\rho = \eta$ in powers of $\rho' = \rho - \eta$, and in third term of (12.23) replace ρ by η. Then we obtain the free particle Lagrangian as

$$\mathcal{L} = \text{const} + \frac{1}{2}(\partial_\mu \rho')^2 - \frac{m^2}{2}\rho'^2 + \frac{\eta^2}{2}(\partial_\mu \vartheta)^2 , \tag{12.24}$$

where $m^2 = 2|\mu^2|$. From here we immediately obtain the equations of motion

$$(\partial_\mu^2 + m^2)\rho' = 0 , \quad \partial_\mu^2 \vartheta = 0 . \tag{12.25}$$

Thus, we obtained two neutral (real) fields ρ' and ϑ, where the first one describes particles with mass m, while the second one corresponds to *massless particles*. In equation (12.25) we have dropped the terms of higher orders describing the interactions of these particles.

The appearance of massless particles due to the spontaneous breaking of continuous symmetry is the essence of the *Goldstone theorem*; these particles are called *Goldstones*[4].

It is not difficult to generalize the Goldstone theorem to the case of higher symmetries. Consider $\varphi(x)$ with n components. Then, group transformations can be written as

$$\vec{\Phi} = S\vec{\Phi}', \tag{12.26}$$

where $\vec{\Phi}$ and $\vec{\Phi}'$ are columns with n components $(\varphi_1, \ldots .\varphi_n)$, abd S is an $n\times n$-matrix. Let the potential $V(\vec{\Phi})$ be dependent only on $\rho^2 = \varphi_1^2 + \ldots + \varphi_n^2$, and there is no other invariants. Then

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \vec{\Phi})^2 - V(\rho)\,. \tag{12.27}$$

In this case we can again make the transformation to "polar" coordinates for field $\vec{\Phi}$ when the field is determined by the modulus $\rho(x)$ and $n-1$ "angular" variables (phases) $\alpha_i(x)(i = 1, 2, \ldots, n-1)$. Then the Lagrangian is written as

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \rho)^2 + \frac{\rho^2}{2}\sum_{i,k=1}^{n-1} \theta_{ik}(\alpha_i)\partial_\mu\alpha_i\,\partial_\mu\alpha_k - V(\rho)\,. \tag{12.28}$$

Assume $V(\rho)$ having its minimum at $\rho = \eta$, i. e., $\langle 0|\rho|0\rangle = \eta$. Angular components α_i can be fixed by the condition $\langle 0|\alpha_i|0\rangle = 0$ (the choice of vacuum!) and by θ_{ik} for $\alpha_i = 0$ having the form $\theta_{ik}(0) = \delta_{ik}$. Then, again introducing $\rho' = \rho - \eta$, we have

$$\mathcal{L} = \text{ const } + \frac{1}{2}(\partial_\mu \rho')^2 - \frac{m^2}{2}\rho'^2 + \frac{1}{2}\eta^2\sum_{i=1}^{n-1}(\partial_\mu\alpha_i)^2\,. \tag{12.29}$$

We see that particles corresponding to fields α_i have zero masses, so that there are now $n-1$ Goldstones. This is the general form of the Goldstone theorem.

12.2 Gauge fields and the Higgs phenomenon

It may seem that the appearance of Goldstone particles with zero mass creates additional difficulties, as our main task is actually to solve the problem of zero mass of gauge gosons. But this is not the case! The unification of the main idea of gauge theories with the concept of spontaneous symmetry-breaking allows us to formulate the natural strategy for the construction of realistic models of interacting particles.

Consider the interaction of scalar field φ, breaking symmetry, with gauge field A_μ in its simplest Abelian (Maxwell) variant. The Lagrangian invariant with respect to

[4] In condensed matter theory the situation is just the same. For example, phase transition into the ferromagnetic state breaks the continuous symmetry of a rotation group – the Heisenberg exchange Hamiltonian is invariant to rotations (it contains scalar products of spins on lattice sites), while in the ground state we have a special direction: that of the vector of spontaneous magnetization (symmetry is lower!). Analogue of Goldstones in this case are *acoustic* spin waves.

local transformations of $U(1)$, has the form[5]

$$\mathcal{L} = \big[(\partial_\mu - ieA_\mu)\varphi^*\big]\big[(\partial^\mu + ieA^\mu)\varphi\big] - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - V(\varphi,\varphi^*)\,, \tag{12.30}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$,

$$V(\varphi,\varphi^*) = \mu^2\varphi^*\varphi + \lambda(\varphi^*\varphi)^2\,, \qquad \mu^2 < 0\,. \tag{12.31}$$

Let us again introduce modulus-phase representation of field φ:

$$\varphi(x) = \frac{1}{\sqrt{2}}\rho(x)e^{i\vartheta(x)}\,. \tag{12.32}$$

But now we can consider (12.32) as a *local gauge transformation* of $U(1)$:

$$\varphi(x) = e^{ie\chi(x)}\varphi'(x)\,, \tag{12.33}$$

where

$$\chi(x) = \frac{1}{e}\vartheta(x)\,, \quad \varphi'(x) = \frac{1}{\sqrt{2}}\rho(x)\,. \tag{12.34}$$

Then, the covariant derivative, entering (12.30), is transformed as follows:

$$D_\mu\varphi = (\partial_\mu + ieA_\mu)e^{ie\chi}\varphi' = e^{ie\chi}(\partial_\mu + ie\partial_\mu\chi + ieA_\mu)\varphi' = e^{ie\chi}(\partial_\mu + ieA'_\mu)\varphi'\,, \tag{12.35}$$

where

$$A'_\mu = A_\mu + \partial_\mu\chi \tag{12.36}$$

or, taking into account (12.33), (12.34),

$$(\partial_\mu + ieA_\mu)\varphi = \frac{1}{\sqrt{2}}e^{i\vartheta}(\partial_\mu + ieA'_\mu)\rho\,, \tag{12.37}$$

where

$$A'_\mu = A_\mu + \frac{1}{e}\partial_\mu\vartheta\,. \tag{12.38}$$

As a result, our Lagrangian is rewritten as

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}[(\partial_\mu - ieA'_\mu)\rho][(\partial^\mu + ieA'^\mu)\rho] - V(\rho^2) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &= \frac{1}{2}(\partial_\mu\rho)^2 + \frac{e^2}{2}\rho^2A'^\mu A'_\mu - V(\rho^2) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,.\end{aligned} \tag{12.39}$$

We see that phase component ϑ of field φ disappeared from the Lagrangian (so that the possibility of Goldstone also disappeared!); it just "gauged-out" into a redefined vector-potential.

[5] Later in this chapter we use the Heaviside system of units.

Let us expand (12.39) in powers of deviation $\rho' = \rho - \eta$ from vacuum average η, limiting ourselves only to quadratic terms. Then we get

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \rho')^2 - \frac{m^2}{2}\rho'^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}e^2\eta^2 A'^\mu A'_\mu + \text{const}\,, \tag{12.40}$$

where $m^2 = 2|\mu^2|$. This Lagrangian describes a pair of free fields: the field ρ' of particles with mass m and the vector field A'_μ *with mass*

$$m_A = e\eta\,, \tag{12.41}$$

which is due to the presence of a nonzero vacuum average of the scalar field. The equations of motion for these fields are

$$\partial_\mu^2 \rho' + m^2\rho' = 0\,, \quad \partial_\mu F^{\mu\nu} + m_A^2 A'^\nu = 0\,. \tag{12.42}$$

The second equation here has the form of a Proca equation.

Thus, in the initial Lagrangian we had the two-component field φ and the vector Maxwell (massless) field A_μ. For $\mu^2 > 0$, retaining only terms quadratic over fields, we obtain the Lagrangian of two free fields, one describing the charged particles with spin 0, while another corresponding to a photon with zero mass and two polarizations, i. e., four particles in total. For $\mu^2 < 0$ the total number of particles remains the same (conservation of the degrees of freedom), but their character has changed: now we have one neutral scalar filed with spin 0 and three independent components of a massive vector boson with spin 1. Initially we had QED of a scalar field, while after the reconstruction of the field we have a "completely different" theory. However, it should be stressed that all transformations were done exactly, and the initial gauge invariance of the theory is conserved (and was used during the derivation!), despite the appearance of gauge field mass! Renormalizability of the theory is also conserved.

The appearance of the vector gauge field mass due to its interaction with scalar field, breaking the symmetry of the ground state, is called the *Higgs phenomenon*, while field ρ is usually called a Higgs field (and the corresponding scalar particles are called *Higgs bosons*).

Remarks on the Ginzburg–Landau theory

Let us show that our theory is the precise analogue of the Ginzburg–Landau theory of superconducting transition, which was formulated long before the discovery of the Higgs phenomenon.

Consider the static case of the Higgs model, when $\partial_0\varphi = 0$, $\partial_0 A^\mu = 0$. The electromagnetic field will be considered in a Coulomb gauge: $A^\mu = (\varphi = 0, \mathbf{A})$, $\nabla \cdot \mathbf{A} = 0$. Then the Lagrangian (12.30) is written as

$$\mathcal{L} = -\frac{1}{2}\big[(\nabla - ie\mathbf{A})\varphi\big]\big[(\nabla + ie\mathbf{A})\varphi^*\big] - \frac{1}{2}m^2|\varphi|^2 - \frac{1}{4}\lambda|\varphi|^4 - \frac{1}{4}(\nabla \times \mathbf{A})^2\,. \tag{12.43}$$

Then

$$F = -\mathcal{L} = \frac{1}{4}(\nabla \times \mathbf{A})^2 + \frac{1}{2}|(\nabla - ie\mathbf{A})\varphi|^2 + \frac{1}{2}m^2|\varphi|^2 + \frac{1}{4}\lambda|\varphi|^4 \tag{12.44}$$

is precisely the free energy density of Ginzburg–Landau theory [37] if we put $m^2 = a(T - T_c)$, where T_c is the temperature of the superconducting transition[6]. In this case we have $m^2 > 0$ for temperatures $T > T_c$ and $m^2 < 0$ for $T < T_c$. For $T < T_c$ the minimum of F is at

$$|\varphi|^2 = -\frac{m^2}{\lambda} > 0\,, \tag{12.45}$$

which defines the equilibrium value of the superconducting *order parameter*, which is the precise analogue of the vacuum average of a Higgs field introduced above (ground state of **the** Higgs field, $T = 0$.).

Ginzburg–Landau free energy is invariant with respect to a gauge transformation:

$$\varphi \to e^{i\Lambda(x)}\varphi\,, \quad \mathbf{A} \to \mathbf{A} + \frac{1}{e}\nabla\Lambda(x)\,, \tag{12.46}$$

and the corresponding conserved current is

$$\mathbf{j} = -\frac{ie}{2}(\varphi^*\nabla\varphi - \varphi\nabla\varphi^*) - e^2|\varphi|^2\mathbf{A}\,. \tag{12.47}$$

For $T < T_c$ and for the spatially homogeneous order parameter φ, only the second term in (12.47) contributes

$$\mathbf{j} = \frac{e^2m^2}{\lambda}\mathbf{A}\,, \tag{12.48}$$

which is the so-called London equation. If we also take into account Maxwell equations $\nabla \times \mathbf{H} = 4\pi\mathbf{j}$, $\nabla \cdot \mathbf{H} = 0$ and calculatethe curl of both sides of equation (12.48), we obtain the equation for the magnetic field inside the superconductor:

$$\nabla^2\mathbf{H} = k^2\mathbf{H}\,, \quad k^2 = -\frac{e^2m^2}{\lambda} > 0\,, \tag{12.49}$$

which describes the *Meissner effect*: the exclusion of the magnetic field from interior of the superconductor. The field is exponentially decreasing inside the superconductor on characteristic length $|k|^{-1}$ (penetration depth) [37].

Finally, from (12.49) it follows that $\nabla^2\mathbf{A} = k^2\mathbf{A}$, which is an analogue of the relativistic equation $\Box A_\mu = -k^2A_\mu$: the "photon" inside the superconductor acquires "mass" $|k|$, which is equivalent to the Higgs effect. Thus, the Higgs model is the relativistic analogue of the Ginzburg-Landau theory, and the Higgs vacuum is similar to the ground state of a superconductor.

12.3 Yang–Mills fields and spontaneous symmetry-breaking

Let us now consider the Higgs mechanism in non-Abelian gauge theories. First of all, we shall recall the main facts related to Yang–Mills fields, using the example of the $SU(2)$ gauge group.

[6] In contrast with standard notations [37], her ewe put the electron mass and velocity of light equal to 1. More importantly, in the Ginzburg–Landau theory $e \to 2e$ in (12.44), in accordance with the value of the Cooper pair charge. But these slight differences are irrelevant for our discussion.

Interaction of scalar field φ with Yang–Mills field $\vec{A}_\mu$ (the arrow denotes a vector in isotopic space) is described by the replacement of the usual derivative $\partial_\mu\varphi$ by the covariant derivative

$$D_\mu\varphi = (\partial_\mu - ig\vec{T}\cdot\vec{A}_\mu)\varphi\,, \tag{12.50}$$

where $\vec{T}$ is the gauge group generator; for $SU(2)$ we have $\vec{T} = \frac{1}{2}\vec{\tau}$.

The Gauge invariance of the Lagrangian puts constraints on the field $\vec{A}_\mu$. If φ corresponds to some isotopic multiplet, its transformation under rotations in isotopic space can be written as

$$\varphi = S\varphi'\,, \tag{12.51}$$

where the operator S depends on the three parameters (angles) of the rotation vector $\vec{\omega}(x)$. Then we write the covariant derivative as

$$\begin{aligned} D_\mu\varphi &= S\partial_\mu\varphi' + (\partial_\mu S)\varphi' - ig\vec{T}\cdot\vec{A}_\mu S\varphi' \\ &= S(\partial_\mu + S^{-1}\partial_\mu S - igS^{-1}\vec{T}\cdot\vec{A}_\mu S)\varphi'\,. \end{aligned} \tag{12.52}$$

This expression should be the same as

$$D_\mu\varphi = S(\partial_\mu - ig\vec{T}\cdot\vec{A}'_\mu)\varphi'\,, \tag{12.53}$$

so that we have to require

$$\vec{T}\cdot\vec{A}'_\mu = S^{-1}(\vec{T}\cdot\vec{A}_\mu)S + \frac{i}{g}S^{-1}\partial_\mu S\,. \tag{12.54}$$

For small $\vec{\omega}$ we have

$$S = 1 + i\vec{T}\cdot\vec{\omega}\,. \tag{12.55}$$

Then

$$\begin{aligned} S^{-1}(\vec{T}\cdot\vec{A}_\mu)S &= (1 - i\vec{T}\cdot\vec{\omega})\vec{T}\cdot\vec{A}_\mu(1 + i\vec{T}\cdot\vec{\omega}) \\ &= \vec{T}\cdot\vec{A}_\mu - i[\vec{T}\cdot\vec{\omega}, \vec{T}\cdot\vec{A}_\mu] = \vec{T}\cdot A_\mu + [\vec{\omega}\times\vec{A}_\mu]\cdot\vec{T}\,, \end{aligned} \tag{12.56}$$

where we have used $[T_i, T_j] = i\varepsilon_{ijk}T_k$: the commutation relations for generators of $SU(2)$. Taking into account $S^{-1}\partial_\mu S = i\vec{T}\cdot\partial_\mu\vec{\omega}$ (12.54) and (12.55) give the general transformation rule

$$\vec{A}'_\mu = \vec{A}_\mu + [\vec{\omega}\times\vec{A}_\mu] - \frac{1}{g}\partial_\mu\vec{\omega}\,, \tag{12.57}$$

so that, besides gradient transformation, the Yang–Mills field is rotated in isotopic space.

The tensor of the Yang–Mills fields has the form

$$\vec{F}_{\mu\nu} = \partial_\mu\vec{A}_\nu - \partial_\nu\vec{A}_\mu + g[\vec{A}_\mu\times\vec{A}_\nu]\,, \tag{12.58}$$

Using (12.57) we can easily show that under infinitesimal rotations in isotopic space $\vec{F}_{\mu\nu}$ is transformed like an isovector:

$$\vec{F}'_{\mu\nu} = \vec{F}_{\mu\nu} + [\vec{\omega} \times \vec{F}_{\mu\nu}]\,. \tag{12.59}$$

The Yang–Mills Lagrangian is written as

$$\mathcal{L}_{YM} = -\frac{1}{4}\vec{F}_{\mu\nu} \cdot \vec{F}^{\mu\nu}\,, \tag{12.60}$$

which is invariant under local gauge group transformations.

Consider now the Yang–Mills field interacting with scalar Higgs field breaking symmetry. Let the Higgs field ϕ be an isospinor having two complex (four real) components

$$\phi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}, \tag{12.61}$$

which transform under rotations in isotopic space as

$$\phi = S\phi'\,, \quad S = e^{\frac{i}{2}g\vec{\tau}\vec{\omega}(x)}\,. \tag{12.62}$$

For small $\vec{\omega}$ we have $S = 1 + ig\vec{\omega}\vec{\tau}/2$.

The Lagrangian for fields ϕ and $\vec{A}_\mu$ is written as

$$\mathcal{L} = (D_\mu\phi)(D^\mu\phi)^* - V(\phi) - \frac{1}{4}\vec{F}_{\mu\nu}\vec{F}^{\mu\nu}\,, \tag{12.63}$$

where

$$D_\mu = \partial_\mu - ig\frac{\vec{\tau}}{2}\vec{A}_\mu\,, \tag{12.64}$$

$$V(\phi) = \mu^2\phi^*\phi + \lambda(\phi^*\phi)^2\,. \tag{12.65}$$

Then, for $\mu^2 < 0$ (12.65) has a minimum at

$$\phi^*\phi = \frac{1}{2}\eta^2\,, \qquad \eta^2 = \frac{|\mu^2|}{\lambda}\,. \tag{12.66}$$

The vacuum state has infinite degeneracy, but we have to choose a single definite vacuum (break symmetry!), e. g., taking

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \eta \end{pmatrix}, \tag{12.67}$$

where η is real and positive. Let us transform to "polar" coordinates:

$$\phi(x) = e^{i\frac{\vec{\tau}}{2}\vec{\vartheta}(x)}\phi'(x) = \left(\cos\frac{\vartheta}{2} + i(\vec{n}\cdot\vec{\tau})\sin\frac{\vartheta}{2}\right)\phi'(x)\,, \tag{12.68}$$

where

$$\phi'(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \rho(x) \end{pmatrix}, \quad \vec{\vartheta} = \vec{n}\,\vartheta \tag{12.69}$$

and $\vec{n}$ is unity vector in the direction of the rotation axis in isospace.

Equations (12.68) and (12.69) define the parametrization of isospinor $\phi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}$ by the four real functions $\rho, \vartheta, \theta, \varphi$, where $\vartheta, \theta, \varphi$ are polar angles determining the direction of vector $\vec{n}$ in isospace:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} i\rho \sin\frac{\vartheta}{2} \cos\theta e^{i\varphi} \\ \rho\left(\cos\frac{\vartheta}{2} - i \sin\frac{\vartheta}{2}\cos\theta\right) \end{pmatrix}, \tag{12.70}$$

so that $\phi^*\phi = \rho^2/2$, $\langle 0|\rho|0\rangle = \eta$, $\langle 0|\vartheta|0\rangle = \langle 0|\theta|0\rangle = \langle 0|\varphi|0\rangle = 0$.

Note that (12.68) is equivalent to (12.62) if we put $\vec{\omega} = \vec{\vartheta}/g$. The Lagrangian is invariant with respect to such transformations and has the form

$$\mathcal{L} = \frac{1}{2}(D'_\mu\rho)(D'^\mu\rho)^* - V(\rho) - \frac{1}{4}\vec{F}'_{\mu\nu}\vec{F}'^{\mu\nu}, \tag{12.71}$$

where in D'_μ and $\vec{F}'_{\mu\nu}$ we replaced $\vec{A}_\mu \to \vec{A}'_\mu$, which corresponds to the gauge transformation (12.57). We see that only one of the four components of the field ϕ that is ρ remain in the Lagrangian; the other three have gauged out!

Taking into account the form of the spinor $\phi' = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \rho \end{pmatrix}$, we rewrite the Lagrangian as (the prime over $\vec{A}_\mu$ is now dropped)

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\rho)^2 + \frac{g^2}{8}\rho^2\vec{A}^\mu \cdot \vec{A}_\mu - V(\rho) - \frac{1}{4}\vec{F}_{\mu\nu}\vec{F}^{\mu\nu}, \tag{12.72}$$

where

$$V(\rho) = \frac{1}{2}\mu^2\rho^2 + \frac{1}{4}\lambda\rho^4. \tag{12.73}$$

For small deviations from the vacuum state, making again an expansion of $V(\rho)$ in powers of $\rho' = \rho - \eta$ and retaining only quadratic terms we obtain

$$\mathcal{L} = \text{const} + \frac{1}{2}(\partial_\mu\rho')^2 - \frac{m^2}{2}\rho'^2 - \frac{1}{4}\vec{F}^0_{\mu\nu}\vec{F}^{0\mu\nu} + \frac{1}{8}g^2\eta^2\vec{A}^\mu \cdot \vec{A}_\mu, \tag{12.74}$$

where $m^2 = 2|\mu^2|$ and $\vec{F}^0_{\mu\nu} = \partial_\mu\vec{A}_\nu - \partial_\nu\vec{A}_\mu$. This Lagrangian describes four free fields: the real scalar field ρ and the triplet vector fields $\vec{A}_\mu$. The scalar field describes particles with mass m, the vector fields particles with mass

$$m_A = \frac{g\eta}{2}. \tag{12.75}$$

Thus, symmetry-breaking has again created masses for particles described by vector (gauge) field $\vec{A}_\mu$. Gauge invariance of the theory is conserved, despite the appearance of these masses! The total number of degrees of freedom has not changed: instead of the three components of field ϕ (Goldstones) which "disappeared", we got longitudinal polarization components of $\vec{A}_\mu$. In this model all the components of the Yang–Mills field acquired mass.

However, to construct the realistic unified theory of weak and electromagnetic interactions we have to guarantee the massiveness of vector bosons, responsible for the

weak interactions (short range forces!), while the electromagnetic field should remain massless. This can be done by some generalization of our $SU(2)$-model. We note that the invariant $\phi^*\phi$ of the scalar field automatically satisfies an additional symmetry, different from $SU(2)$ used above. We can multiply ϕ by an arbitrary phase factor like $\exp\left[i\frac{f}{2}\lambda(x)\right]$, and nothing will change. This is the transformation of $U(1)$, so that we shall now consider the theory with $SU(2)\otimes U(1)$ symmetry. This additional Abelian symmetry $U(1)$ allows us to associate with the particles of the field ϕ, except isospin, some new "hypercharge", which leads to the introduction of an additional (Abelian) gauge field, which will be denoted as B_μ. As a result, the full symmetry of our model corresponds to the invariance to local gauge transformations

$$\phi = S\phi', \tag{12.76}$$

where

$$S = \exp\left(ig\vec{\omega}(x)\cdot\frac{\vec{\tau}}{2} + if\frac{\lambda(x)}{2}\right), \tag{12.77}$$

with the Lagrangian of the model having the form

$$\mathcal{L} = (D_\mu\phi)(D^\mu\phi)^* - V(\rho) - \frac{1}{4}\vec{F}_{\mu\nu}\vec{F}^{\mu\nu} - \frac{1}{4}G_{\mu\nu}G^{\mu\nu}, \tag{12.78}$$

where

$$D_\mu = \partial_\mu - ig\frac{\vec{\tau}}{2}\cdot\vec{A}_\mu - i\frac{f}{2}B_\mu, \tag{12.79}$$

$$G_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{12.80}$$

Further analysis, in fact, just repeats the previous analysis, so that we drop the details. It is convenient to introduce, instead of fields $A_{\mu1}, A_{\mu2}, A_{\mu3}$ and B_μ, the linear combinations

$$W_\mu = \frac{1}{\sqrt{2}}(A_{\mu1} + iA_{\mu2}), \tag{12.81}$$

$$\begin{aligned} Z_\mu &= A_{\mu3}\cos\alpha - B_\mu\sin\alpha, \quad A_{\mu3} = \ \ Z_\mu\cos\alpha + A_\mu\sin\alpha, \\ A_\mu &= A_{\mu3}\sin\alpha + B_\mu\cos\alpha, \quad B_\mu \ = -Z_\mu\sin\alpha + A_\mu\cos\alpha, \end{aligned} \tag{12.82}$$

where

$$\cos\alpha = \frac{g}{\tilde{g}}, \quad \sin\alpha = \frac{f}{\tilde{g}}, \quad \tilde{g} = \sqrt{g^2+f^2}, \quad tg\alpha = \frac{f}{g}. \tag{12.83}$$

Then our Lagrangian (12.78) is rewritten in terms of these new fields as

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\rho)^2 - V(\rho) + \frac{\rho^2}{4}\left[g^2W^{*\mu}W_\mu + \frac{\tilde{g}^2}{2}Z^\mu Z_\mu\right] - \frac{1}{4}\vec{F}_{\mu\nu}\vec{F}^{\mu\nu} - \frac{1}{4}G_{\mu\nu}G^{\mu\nu}. \tag{12.84}$$

Here it is important that field A_μ from (12.82) does not enter the square bracket term, so that after the appearance of the nonzero vacuum average for field ρ (spontaneous symmetry-breaking) this field remains massless and can be identified with the usual

electromagnetic field. At the same time, fields W_μ and Z_μ acquire the mass

$$m_W = \frac{g\eta}{2}, \qquad m_Z = \frac{\tilde{g}\eta}{2} = \frac{m_W}{\cos\alpha}. \tag{12.85}$$

This is immediately seen if we rewrite the Lagrangian (12.84) up to the quadratic terms in $\rho' = \rho - \eta$, W_μ and Z_μ:

$$\begin{aligned}\mathcal{L} = {} & \frac{1}{2}(\partial_\mu\rho')^2 - \frac{1}{2}m^2\rho'^2 - \frac{1}{2}(\partial_\mu W_\nu - \partial_\nu W_\mu)(\partial_\mu W_\nu^* - \partial_\nu W_\mu^*) \\ & + \frac{1}{4}g^2\eta^2 W_\mu^* W^\mu - \frac{1}{4}(\partial_\mu Z_\nu - \partial_\nu Z_\mu)^2 + \frac{1}{8}\tilde{g}^2\eta^2 Z_\mu^2 \\ & - \frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)^2 + \text{ const}\,,\end{aligned} \tag{12.86}$$

where $m^2 = 2|\mu^2|$.

Field W_μ (12.81) is complex, i. e., charged, while fields A_μ and Z_μ (12.82) are real (neutral). From the definition $\vec{F}_{\mu\nu}$ (12.58) and (12.81), (12.82) it follows that

$$\frac{1}{\sqrt{2}}(F^1 + iF^2)_{\mu\nu} = D_\mu W_\nu - D_\nu W_\mu\,, \tag{12.87}$$

where

$$D_\mu = \partial_\mu + igA_{\mu 3} = \partial_\mu + ig\sin\alpha A_\mu + ig\cos\alpha Z_\mu\,. \tag{12.88}$$

If we identify field A_μ with the Maxwell electromagnetic field, from (12.88) we get the following relation between Yang-Mills coupling constant g and the usual electric charge e:

$$e = g\sin\alpha\,. \tag{12.89}$$

An important property of this theory is its renormalizability. Renormalizability of QED is guaranteed by the masslessness of the photon and also by its neutrality. We can drop one of these properties, but the theory will still be renormalizable. For example, we can work with the renormalizable theory of fermions interacting with a massive vector neutral field. Thus, we do not worry much whether or not the photon has very small (undetected up to now) mass. The theory will not change much. Also renormalizable is the Yang–Mills theory of two charged and one neutral massless fields interacting with fermions. Due to the Higgs phenomenon the charged Yang–Mills fields may become massive: in the following we shall call them charged intermediate W-bosons, transmitting weak interactions, while Z-bosons are similar neutral particles. The electromagnetic field A_μ remains massless. The question arises whether or not our theory will remain renormalizable after the Higgs phenomenon (spontaneous symmetry-breaking). We can expect it to remain renormalizable, as the initial Lagrangian is definitely renormalizable, and all further results were obtained by clear transformations and change in notations. These expectations are actually confirmed by the rigorous proof which we shall not consider here.

12.4 The Weinberg–Salam model

The correct scheme for the unified description of weak and electromagnetic interactions was proposed (independently) by Weinberg and Salam.This model is well confirmed by experiments and forms the basis of the Standard Model. The main idea of the theory of electroweak interactions is that weak interactions are mediated by gauge bosons ($W^{\pm}, Z$), which are "initially" massless, while their masses (short range nature of weak forces!) are acquired as a result of the Higgs mechanism. The electromagnetic field obviously remains massless. This scheme for gauge fields was presented in the previous section. Now we have to include the leptons: the electron and the neutrino[7], which are also assumed to initially be massless. The Higgs mechanism (spontaneous symmetry-breaking) should generate the mass of an electron, leaving the neutrino massless[8].

The Dirac Lagrangian

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi \tag{12.90}$$

for $m = 0$ transforms to $i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi$. Let us introduce, as is usual for massless fermions,

$$\psi_L = \frac{1}{2}(1-\gamma^5)\psi\,, \quad \psi_R = \frac{1}{2}(1+\gamma^5)\psi\,, \quad \psi = \psi_L + \psi_R\,, \tag{12.91}$$

where $\gamma^5 = -i\gamma^0\gamma^1\gamma^2\gamma^3$. Then

$$i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi = i\bar{\psi}_R\gamma^{\mu}\partial_{\mu}\psi_R + i\bar{\psi}_L\gamma^{\mu}\partial_{\mu}\psi_L\,, \tag{12.92}$$

as γ^5 anticommutes with γ^{μ}. The electron (as well as the muon and the τ-lepton) have both L and R components, while the neutrino, according to the two-component neutrino model, i. e., ν_e (and also ν_{μ}, ν_{τ}) has only L-components. Then the initial Lagrangian of the leptons can be written as

$$\mathcal{L} = i\bar{e}_R\gamma^{\mu}\partial_{\mu}e_R + i\bar{e}_L\gamma^{\mu}\partial_{\mu}e_L + i\bar{\nu}_e\gamma^{\mu}\partial_{\mu}\nu_e + (e\to\mu) + (e\to\tau)\,, \tag{12.93}$$

where the fermion fields are denoted by the symbols of the appropriate particles. The contribution of the higher generations can be written in a similar form, but we skip it here for brevity.

The transformations of a gauge group should be applied to particles with the same space-time properties, i. e., the only possibility is to mix e_L è ν_e. We introduce the isospinor

$$\psi_L = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix} \tag{12.94}$$

and associate with this doublet the non-Abelian charge ("weak" isospin) $I_W = 1/2$, with its projections corresponding to its two components: neutrino ν_e corresponds to

[7] The other generations of leptons are described precisely in the same way.

[8] Here we neglect the possibility of very small neutrino mass.

$I_W^3 = +1/2$, while "electron" e_L to $I_W^3 = -1/2$. The remaining

$$\psi_R = e_R\,, \tag{12.95}$$

is considered to be an isosinglet: $I_W = 0$. Then we can write the Lagrangian as

$$\mathcal{L} = i\bar{\psi}_R\gamma^\mu\partial_\mu\psi_R + i\bar{\psi}_L\gamma^\mu\partial_\mu\psi_L\,, \tag{12.96}$$

which is invariant with respect to the $SU(2)$ group of isospin transformations:

$$\psi_L \to e^{-\frac{i}{2}\vec{\tau}\vec{\alpha}}\psi_L\,, \quad \psi_R \to \psi_R \tag{12.97}$$

or, in more detail,

$$\begin{pmatrix} \nu_e \\ e_L \\ e_R \end{pmatrix} \to \begin{pmatrix} e^{-i\frac{\vec{\tau}}{2}\vec{\alpha}} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \nu_e \\ e_L \\ e_R \end{pmatrix}. \tag{12.98}$$

The electric charge Q and the third component of the weak isospin I_W^3 for the left and right fields are connected by the obvious relations

$$L: \;\; Q = I_W^3 - \frac{1}{2}; \quad R: \;\; Q = I_W^3 - 1\,. \tag{12.99}$$

If we make this symmetry a local gauge symmetry, i. e., put $\vec{\alpha} = \vec{\alpha}(x)$, this will lead to the appearance of three massless Yang–Mills fields. However, the photon will not be there, as the right electron e_R, being an isosinglet, will not interact with these fields, while it is obviously interacting with the usual photons. To solve this problem we can use the fact that $SU(2)$ is not the maximal possible symmetry of our Lagrangian. We can additionally transform e_R by simple $U(1)$:

$$e_R \to e^{i\beta}e_R\,. \tag{12.100}$$

But this can be only the common transformation for all the fields of our model. Then ν_e and e_L should also acquire the common phase factor, but its phase is not necessarily the same as for R. Thus, we can write

$$\begin{pmatrix} \nu_e \\ e_L \\ e_R \end{pmatrix} \to \begin{pmatrix} e^{in\beta} & 0 & 0 \\ 0 & e^{in\beta} & 0 \\ 0 & 0 & e^{i\beta} \end{pmatrix} \begin{pmatrix} \nu_e \\ e_L \\ e_R \end{pmatrix}. \tag{12.101}$$

where n is some number to be determined later. This $U(1)$-symmetry leads to the existence of some conserving charge, with e_R having one value of this charge, while both ν_e and e_L have another value. This is obviously *not* the electric charge Q, as ν_e and e_L in reality have different values of Q. The gauge field corresponding to this $U(1)$ symmetry is not the usual electromagnetic field. Weinberg proposed to consider this

symmetry as the conservation of the "weak hypercharge" Y_W, defined by the relation[9]

$$Q = I_W^3 + \frac{Y_W}{2}. \tag{12.102}$$

Comparing this expression with (12.99), we see that for the left and right leptons it is necessary to introduce

$$L: \ Y_W = -1; \quad R: \ Y_W = -2. \tag{12.103}$$

Thus, in equation (12.101) we have to put $n = 1/2$, so that the coupling constant for interaction of the hypercharge gauge field for left fields is twice as small as the corresponding constant for right fields. As a result, the $U(1)$ group transformation is finally written as

$$\begin{pmatrix} \nu_e \\ e_L \\ e_R \end{pmatrix} \to \begin{pmatrix} e^{i\beta/2} & 0 & 0 \\ 0 & e^{i\beta/2} & 0 \\ 0 & 0 & e^{i\beta} \end{pmatrix} \begin{pmatrix} \nu_e \\ e_L \\ e_R \end{pmatrix}. \tag{12.104}$$

Thus, the Lagrangian (12.93), (12.96) is invariant to the direct product of groups $SU(2) \otimes U(1)$. The Yang–Mills theory with such a symmetry was already examined in the previous section. We have introduced four gauge fields: isotriplet $\vec{A}_\mu$ and isosinglet B_μ, for all of them $Y_W = 0$.

Lepton fields ψ_L and ψ_R interact with fields $\vec{A}_\mu$, B_μ, and the Higgs field ϕ. First of all, let us consider this last interaction. The corresponding term in Lagrangian is written in a form similar to equation (12.12), which was already discussed in connection with the mass generation mechanism for fermions:

$$\mathcal{L}_M = -\sqrt{2}a(\bar{\psi}_L \psi_R \phi + \bar{\psi}_R \psi_L \phi^*), \tag{12.105}$$

where a is the dimensionless coupling constant of this (renormalizable!) interaction. We write the Higgs field as an isospinor:

$$\phi = \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix}, \quad \phi^* = (\varphi^-, \varphi^{0*}), \tag{12.106}$$

with components corresponding to projections of weak isospin $I_W^3 = \pm 1/2$. From (12.102) we find the corresponding quantum numbers

$$I_W = 1/2, \quad Y_W = 1. \tag{12.107}$$

Both fields φ^+ and φ^0 are complex, so that we can write

$$\phi = \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}}(\varphi_3 + i\varphi_4) \\ \frac{1}{\sqrt{2}}(\varphi_1 + i\varphi_2) \end{pmatrix}, \tag{12.108}$$

where $\varphi_1, \ldots, \varphi_4$ are real fields.

[9] This expression is written in analogy with the Gell-Mann–Nishijima formula for the hypercharge of hadrons [40]

The covariant derivative describing the interaction of the Higgs field with the gauge fields has the form

$$D_\mu\phi = (\partial_\mu - \frac{i}{2}g\vec{\tau}\cdot\vec{A}_\mu - \frac{i}{2}fB_\mu)\phi\,. \tag{12.109}$$

Finally, the part of the Lagrangian containing field ϕ is equal to

$$\mathcal{L}_\phi = (D^\mu\phi)^*(D_\mu\phi) - \mu^2\phi^*\phi - \lambda(\phi^*\phi)^4 - \sqrt{2}a(\bar{\psi}_L\psi_R\phi + \bar{\psi}_R\psi_L\phi^*)\,. \tag{12.110}$$

The part of the Lagrangian corresponding to Higgs field interaction with the leptons is written in more detail as

$$-\sqrt{2}a(\bar{\nu}_e e_R\varphi^+ + \bar{e}_L e_R\varphi^0 + \bar{e}_R\nu_e\varphi^- + \bar{e}_R e_L\varphi^0)\,. \tag{12.111}$$

Further, we have

$$\phi^*\phi = (\varphi^+)^*\varphi^+ + (\varphi^0)^*\varphi^0 = \frac{1}{2}(\varphi_1^2 + \varphi_2^2 + \varphi_3^2 + \varphi_4^2)\,. \tag{12.112}$$

For $\mu^2 < 0$ the Higgs field Bose condenses and the minimum energy corresponds to

$$\langle 0|(\varphi^*\varphi)|0\rangle = \eta^2 = -\frac{\mu^2}{\lambda}\,. \tag{12.113}$$

Let us choose a vacuum to satisfy

$$\langle 0|\varphi_1|0\rangle = \eta\,, \quad \langle 0|\varphi_2|0\rangle = \langle 0|\varphi_3|0\rangle = \langle 0|\varphi_4|0\rangle = 0\,, \tag{12.114}$$

i. e.,

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \eta \end{pmatrix}. \tag{12.115}$$

Then, the lowest-order (over excitations) interaction is written as

$$\mathcal{L}_M = \sqrt{2}a(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)\phi = a(\bar{e}_L e_R + \bar{e}_R e_L)\eta = a\eta\bar{\psi}\psi\,, \tag{12.116}$$

so that the Higgs condensate η interacts only with electrons. Thus, we achieved what was desired: the electron acquired the mass

$$m_e = a\eta\,, \tag{12.117}$$

while the neutrino remained massless!

Consider now lepton interactions with gauge fields, which is described by the covariant derivatives

$$D_\mu\psi = \left(\partial_\mu - ig\vec{T}\cdot\vec{A}_\mu - if\frac{Y}{2}B_\mu\right)\psi\,, \tag{12.118}$$

where Y is the weak hypercharge of field ψ, and g and f are the corresponding coupling constants. For ψ_L we have $\vec{T} = \frac{1}{2}\vec{\tau}$, $Y = -1$, while for ψ_R we have $\vec{T} = 0$ and

$Y = -2$. Then the part of the Lagrangian corresponding to the interaction between the leptons and the gauge fields has the form

$$\mathcal{L} = i\bar{\psi}_L\gamma^\mu\left(\partial_\mu - ig\frac{\vec{\tau}}{2}\cdot\vec{A}_\mu + i\frac{f}{2}B_\mu\right)\psi_L + i\bar{\psi}_R\gamma^\mu(\partial_\mu + ifB_\mu)\psi_R\,. \quad (12.119)$$

The gauge fields entering these expressions, as we seen in (12.81), (12.82), (12.83), can be divided into three types: a field of charged heavy nosons W_μ, a field of neutral heavy bosons Z_μ, and an electromagnetic field A_μ. Let us separately write the parts of Lagrangian $\mathcal{L}_W$, $\mathcal{L}_Z$, $\mathcal{L}_A$, corresponding to the interaction with these fields. First we write

$$\mathcal{L}_W = \frac{g}{2}\bar{\psi}_L\gamma^\mu(\tau_1 A_{1\mu} + \tau_2 A_{2\mu})\psi_L = \frac{g}{\sqrt{2}}\gamma^\mu(\bar{\nu}_e W_\mu e_L + \bar{e}_L W_\mu^* \nu_e)\,. \quad (12.120)$$

This is the Lagrangian of the weak interaction of the leptons due to the exchange by $W^\pm$-bosons (the so-called charged currents). There are also interactions with fields $A_{\mu 3}$ and B_μ, from (12.119); these are written as

$$\frac{1}{2}\bar{\psi}_L\gamma^\mu(g\tau_3 A_{\mu 3} - fB_\mu)\psi_L - f\bar{\psi}_R\gamma^\mu B_\mu\psi_R \quad (12.121)$$

or, using (12.82), (12.83),

$$\frac{\tilde{g}}{2}\gamma^\mu[\bar{\nu}_e(\cos\alpha A_{\mu 3} - \sin\alpha B_\mu)\nu_e - \bar{e}_L(\cos\alpha A_{\mu 3} + \sin\alpha B_\mu)e_L - 2\sin\alpha\bar{e}_R B_\mu e_R]\,, \quad (12.122)$$

so that, expressing $A_{\mu 3}$ and B_μ via Z_μ and A_μ according to (12.82), we get

$$\mathcal{L}_Z = \frac{\tilde{g}}{2}\gamma^\mu(\bar{\nu}_e Z_\mu \nu_e - \cos 2\alpha\bar{e}_L Z_\mu e_L + 2\sin^2\alpha\bar{e}_R Z_\mu e_R), \quad (12.123)$$

the weak interaction due to the exchange by neutral Z-bosons (the so-called neutral currents), and

$$\mathcal{L}_A = -g\sin\alpha\gamma^\mu(\bar{e}_L A_\mu e_L + \bar{e}_R A_\mu e_R), \quad (12.124)$$

the usual electromagnetic interaction. Note that (12.124) once more confirms the relation for the electric charge (12.89). Thus, the model under discussion gives a unified description of weak and electromagnetic interactions, where the fields corresponding to $W^\pm$- and Z-bosons as well as to the electromagnetic field appear from the fundamental requirement of invariance with respect to the local gauge transformations of $SU(2)\otimes U(1)$.

During the first years since the construction of Weinberg–Salam model, weak interaction processes due to neutral currents (12.123) were not known, which was considered to be a shortcoming of the model. The experimental discovery of such processes in 1973 in CERN was actually the first serious confirmation of the theory of electroweak interactions. One of the simplest processes due to weak interactions is muon μ decay, described by the diagram shown in Figure 12.3. If the mass of the

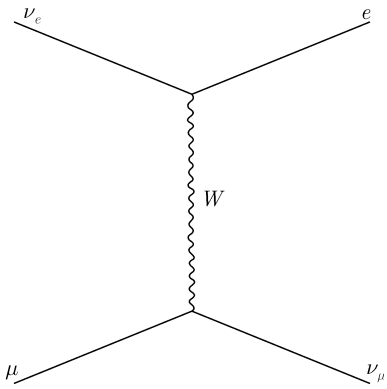


Figure 12.3

W-boson is a significantly larger muon mass, its propagator is simply proportional to $\frac{1}{m_W^2}$, and the appropriate transition amplitude is equivalent to the amplitude derived from the phenomenological (nonrenormalizable) 4-fermion interaction introduced by Fermi (10.96), (10.270):

$$\frac{g^2}{2m_W^2}(\bar{e}_L\gamma_\alpha\nu_e)(\bar{\nu}_\mu\gamma^\alpha\mu)\,. \tag{12.125}$$

Comparing with (10.270), we obtain the following expression for Fermi coupling:

$$\frac{G}{\sqrt{2}} = \frac{g^2}{8m_W^2}\,. \tag{12.126}$$

The value of G is well known experimentally (10.273). We see, that its smallness ("weakness" of the weak interactions) is actually due to the *large mass* of the intermediate boson in the denominator of (12.125), while the fundamental coupling constant is actually $g \sim e$! Using (12.89)) and the experimental values of e and G, we can, with the help of (12.85) and (12.126), find the following estimates for the masses of intermediate gauge bosons:

$$m_W = \frac{e}{2^{5/4}G^{1/2}\sin\alpha} = \frac{37\,\text{GeV}}{\sin\alpha}\,, \quad m_Z = \frac{m_W}{\cos\alpha} = \frac{74\,\text{GeV}}{\sin 2\alpha}\,, \tag{12.127}$$

so that $m_W > 37\,\text{GeV}$ and $m_Z > 74\,\text{GeV}$. Using (12.85) and (12.127) we can write

$$\eta = \frac{m_W}{g} = \frac{37\,\text{GeV}}{e} = 122\,\text{GeV}\,. \tag{12.128}$$

Then from (12.117) we have

$$a = \frac{m_e}{\eta} \approx 5\cdot 10^{-6}\,, \tag{12.129}$$

so that the coupling constant of leptons with the Higgs field is very small.

Experimental studies of the weak interactions due to neutral currents has already led in the early 1980s to the following estimate of "angle" α:

$$\sin\alpha \approx 0.47\,. \tag{12.130}$$

Then from (12.127) we have

$$m_W \approx 78.6\,\mathrm{GeV}\,, \quad m_Z \approx 89.3\,\mathrm{GeV}\,. \tag{12.131}$$

The triumph of the theory was the experimental discovery in 1983 of $W^{\pm}$- and Z-bosons in CERN with masses $m_W \approx 80\,\mathrm{GeV}$, $m_Z \approx 92\,\mathrm{GeV}$. Since then the theory has been confirmed in numerous experiments, and at present it is the commonly accepted scheme for describing electroweak interactions. The constants of the theory are known with high accuracy. The present-day experimental situation is well described in [67].

For many years the main unsolved problem was the absence of any direct experimental observation of Higgs bosons. One of the difficulties was due to the inability to make definite predictions for Higgs boson mass within the Weinberg–Salam model; only a rather wide interval of possible values were theoretically predicted. However, as we already mentioned in Chapter 1, in July 2012 it was announced that a particle with the properties expected for the Higgs boson was discovered in a number of LHC experiments, with mass around 125–126 GeV. If finally confirmed by further studies, this will definitely be the final triumph of the theory. The brief discussion of the present-day situation with the experimental discovery of the Higgs oson can be found in [55].

The Standard Model

The "Standard Model" of elementary particles is the combination of the electroweak theory due to Weinberg and Salam with quantum chromodynamics (QCD), which we discussed above babove. The full gauge symmetry is given by the direct product of color symmetry and the symmetries of weak isospin and weak hypercharge: $SU(3) \otimes SU(2) \otimes U(1)$. If we limit ourselves to the most important first generation of fermions, the fermion sector of the model is defined by

$$L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \quad e_R, \quad Q_L = \begin{pmatrix} u_\alpha \\ d_\alpha \end{pmatrix}_L, \quad u_{R\alpha}\,, \quad d_{R\alpha}\,, \tag{12.132}$$

where u and d denote the corresponding quarks (α is the color index). The covariant derivative which determine the fermion interactions with Yang–Mills fields has the form

$$D_\mu = \partial_\mu - i g_1 \frac{Y}{2} B_\mu - i g_2 \frac{\tau^i}{2} W_\mu^i - i g_3 \frac{\lambda^a}{2} G_\mu^a\,, \tag{12.133}$$

where λ^a are the generators of the color group $SU(3)$ (cf. Chapter 2) and G_μ^a are the vector fields of gluons. The Higgs sector of the theory was described above. Gluons

remain massless, but they are not observed as free particles, due to the phenomenon of "confinement", which we shall discuss later. This theory is sufficient, in principle, to describe the entire world (or universe) surrounding us. At present all the predictions of this theory are rather satisfactorily confirmed by existing experiments[10]. The attempts of the real unification of all known interactions within some single gauge group, which includes symmetries $SU(3) \otimes SU(2) \otimes U(1)$ of the Standard Model as a subgroup, are usually called "great unification theories" (GUT). We shall briefly discuss such attempts in the next chapter.

Phase transitions in quantum field theory at finite temperatures

Finally, let us briefly discuss one very interesting direction of research in modern quantum field theory. We have seen that the basis of the unified theory of electroweak interactions is the phenomenon of spontaneous symmetry-breaking and the Higgs mechanism. We already noted that this is a typical phase transition, like that taking place e. g., in superconductors. From the theory of condensed matter we know that any symmetry-breaking disappears at high enough temperature $T > T_c$, when the system goes to a symmetric phase. The same phenomenon takes place in the models of quantum field theory introduced above. This was clearly demonstrated for the first time by Kirzhnitz and Linde [41]. An appropriate theoretical analysis can be performed using the standard (Matsubara) formulation of quantum field theory at finite temperatures, which is widely used in statistical physics [1]. We have no room for a detailed discussion of these interesting problems here and therefore limit ourselves only to formulating some of the main conclusions.

The vacuum average of the Higgs field, which plays the role of the order parameter, becomes zero for $T > T_c$, where

$$T_c \approx \sqrt{\frac{3|\mu^2|}{\lambda}} \sim \eta(0) \sim 10^2 \div 10^3 \,\mathrm{GeV}\,. \tag{12.134}$$

For $T < T_c$ the order parameter behaves in the more or less usual way:

$$\eta^2(T) = \frac{|\mu^2|}{\lambda} - \Phi(T)\,, \tag{12.135}$$

where $\Phi(T)$ is some increasing function of temperature. As a result, we obtain the order parameter dependence shown in Figure 12.4(a). But we have seen above that masses appearing due to spontaneous symmetry-breaking are proportional to the vacuum average η at $T = 0$. Correspondingly, as temperature increases the masses of the gauge bosons, leptons, and other particles diminish, and at $T = T_c$ become *zero*, as shown in Figure 12.4(b). Already at this elementary level of discussion it becomes

[10] A rief review of experimental situation with the Standard Model can be found in the review paper [22], while the importance of this theory for "common day" life is well-described in [12].

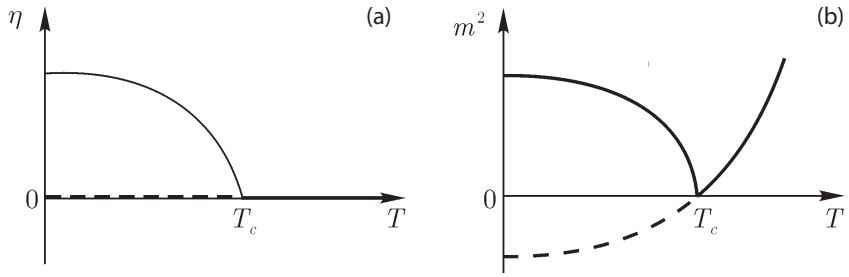


Figure 12.4

clear that the disappearance of the masses of the elementary particles creates strong long range forces which may play a decisive role in cosmology, because in the first moments after the "Big Bang" the temperature of the universe was extremely high. These conclusions were followed by explosive developments of new approaches in cosmology [16, 41]. Similar effects may be important for experiments with very high energy collisions of heavy nuclei when very high temperatures can also be generated.

At present, the analysis the effects of temperature in quantum field theory is an important part of elementary particles theory, which again stresses the unity of quantum field theory and modern statistical physics.

Chapter 13

Renormalization

13.1 Divergences in φ^4

The concept of renormalizability plays an absolutely fundamental role in modern quantum field theory. Only renormalizable theories are considered to be physical. In Chapter 8 we briefly discussed the renormalizability of QED. Now we will return to a more detailed discussion of the general situation for different field theory models.

Below we shall mainly consider the simplest scalar field $g\varphi^4$ theory, which was already discussed above in Chapter 10. There we already met with typical divergences like in equation (10.125). Now we shall present a more serious analysis of divergences. Using the rules of diagram technique we can once again write the first order ($\sim g$) correction to self-energy, corresponding to the diagram in Figure 13.1. The corresponding analytic expression is

$$\frac{1}{i}\Sigma = -ig\frac{1}{2}\int \frac{d^4q}{(2\pi)^4}\frac{1}{q^2-m^2}\,, \tag{13.1}$$

where we have taken into account the symmetry factor $1/2$. Here in the integrand we have the fourth power of q, while in the denominator we have q squared, so that our integral diverges quadratically at large q (i. e., at the upper limit, "ultraviolet divergence").

Another typical divergence arises in the order $\sim g^2$ from the diagram shown in Figure 13.2, where $p_1 + p_2 = q$ and $p_1 + p_2 + p_3 + p_4 = 0$. The corresponding analytic expression is

$$-g^2\int \frac{d^4p}{(2\pi)^4}\frac{1}{p^2-m^2}\frac{1}{(p-q)^2-m^2}\,. \tag{13.2}$$

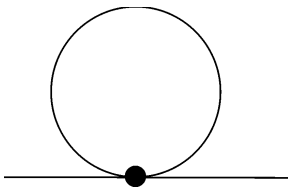


Figure 13.1

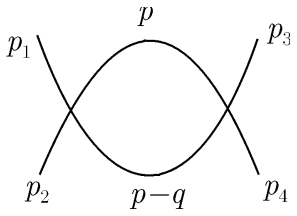


Figure 13.2

Here we have the fourth power of p both in the numerator and denominator of the integrand, which leads to logarithmic divergence[1].

Let us show how we can determine the divergence power of an arbitrary diagram. A similar analysis was already performed for QED in Chapter 8. But here we shall present more of the details. It is obvious that in an arbitrary diagram each propagator contributes $\sim p^2$ (for large p we can neglect m!) in the denominator of the integrand, while each vertex contributes $\sim p^4$ in the numerator, as well as the δ-function, expressing momentum conservation in this vertex. The number of independent integration momenta is equal to the number of closed loops in the diagram. For the diagrams, shown above, this number is 1 (one-loop diagrams). Consider a diagram of the order of $\sim g^n$, i. e., with n vertices. Suppose it has E external lines, I internal lines, and L loops. For generality we consider space-time with dimensionality d; in this case the vertices contribute p^d to the numerator. Let us define the *conditional degree of divergence* D of the given diagram as

$$D = dL - 2I\,. \tag{13.3}$$

For the diagrams shown above we have $D = 2$ and $D = 0$. Now we can express D via E and n, excluding I and L. In fact, we have in total I internal momenta. In each of the n vertices we have momentum conservation, and we also have the total momentum conservation law for the scattering process described by our diagram (external momenta are fixed). As a result, there are in total $n - 1$ relations between the integration momenta. Thus, there are only $I - n + 1$ independent integration momenta. But this number is equal to L:

$$L = I - n + 1\,. \tag{13.4}$$

In φ^4-theory each vertex is entered by four lines, so that we have in total $4n$ lines in the diagram, but part of the lines are internal and another part external. During these calculations the internal lines are counted twice, as each of them connects two vertices. Then we have

$$4n = E + 2I\,. \tag{13.5}$$

From (13.3), (13.4), (13.5) we immediately get

$$D = d - \left(\frac{d}{2} - 1\right) E + n(d-4)\,. \tag{13.6}$$

In particular, for $d = 4$ we have

$$D = 4 - E\,, \tag{13.7}$$

[1] In fact, we have already considered such diagrams during our discussion of the theory of critical phenomena in four-dimensional space, where the problem of divergence was solved by the introduction of a natural cutoff Λ, of the order of the inverse lattice constant.

which, by the way, gives the correct answers for the simplest diagrams discussed above. From (13.7) we can see that the degree of divergence diminishes with the growth of the number of external lines (and depends only on this number!)[2].

Let us return to the discussion of the general formula (13.6) and consider the last term in this expression. If the coefficient before n is positive, the situation is hopeless: the degree of divergence D grows with the growth of n, so that the full theory (summed over all n) will contain an infinite number of terms containing divergences (in each order of perturbation theory) with a higher degree of divergence (than in the previous order). This is equivalent to the nonrenormalizability of the theory. In φ^4-theory at $d = 4$, the degree of divergence depends only on E and does not depend on the order of perturbation theory, so that we have the finite number of the *types* of divergences, so we can hope that the corresponding infinite contributions can be hidden in a finite number of (infinite) renormalizations of the appropriate physical characteristics (renormalized theory). A finite number of the types of divergences is the *necessary* condition for renormalizability.

It is useful to analyze similar formulas for the theory with the general interaction φ^r. Equations (13.3) and (13.4) do not change, while equation (13.5) transforms into

$$rn = E + 2I\,, \tag{13.8}$$

so that equation (13.6) is rewritten as

$$D = d - \left(\frac{d}{2} - 1\right) E + n\left[\frac{r}{2}(d-2) - d\right]. \tag{13.9}$$

Now for $d = 4$ we have

$$D = 4 - E + n(r-4)\,. \tag{13.10}$$

For φ^6-theory we have $D = 4 - E + 2n$, and it is nonrenormalizable. On the other hand, for φ^3-theory we have $D = 4 - E - n$, and the degree of divergence diminishes with the increase of n, so that for fixed E there is only the finite number of divergent diagrams, and we are dealing with the so-called super-renormalizable theory[3]. Note that for $d = 2$ we have $D = 2 - 2n$ and independent of r.

Let us return to equation (13.7) and discuss the convergence or divergence of diagrams with $E > 4$. In φ^4-theory the number E is always even. Consider the examples of diagrams shown in Figure 13.3. Here $E = 6$, so that according to criterion (13.7) all of these diagrams seem to be convergent. This is correct for the diagram in Figure 13.3(a), but obviously wrong for diagrams (b) and (c), which contain "hidden" divergences from loops (considered above). It is because of such cases that we called D the conditional degree of divergence. It is important, however, that the inverse state-

[2] It may seem that all diagrams with the number of external lines greater than 4 are convergent. For example, for $E = 6$ we have $D = -2$. However, this is a wrong conclusion, as we shall see below.

[3] However, this theory is actually bad: there is a stable ground state!

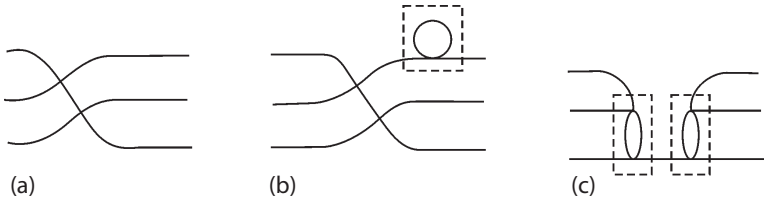


Figure 13.3

ment is always correct: the Feynman diagram converges if its degree of divergence D and the degrees of divergence of all its subdiagrams are negative (Weinberg theorem).

Two divergent diagrams shown above in Figures 13.1 and 13.2, are called *primitively divergent*. These are the only primitively divergent diagrams of φ^4-theory (types of divergences).

Dimensional analysis

Let us perform dimensional analysis in d-dimensional space. The action $S = \int d^d x \mathcal{L}$ is dimensionless. Then we easily find

$$[\mathcal{L}] = L^{-d}\,, \quad [\mathcal{L}] = \Lambda^d\,, \tag{13.11}$$

where L is some length, and Λ is the corresponding momentum. From the term $\sim \partial_\mu\varphi\partial^\mu\varphi$ in $\mathcal{L}$, taking into account $[\partial_\mu] = L^{-1}$, we have

$$[\varphi] = L^{1-\frac{d}{2}} = \Lambda^{\frac{d}{2}-1}\,. \tag{13.12}$$

Consider the interaction $g\varphi^r$. If we define the dimensionality of the coupling constant as $[g] = L^{-\delta} = \Lambda^\delta$, we obviously get $-\delta + r\left(1 - \frac{d}{2}\right) = -d$, so that

$$\delta = d + r - \frac{rd}{2}\,. \tag{13.13}$$

Thus, the dimensionality of the coupling constants in different theories is

$$\begin{aligned}
&g\varphi^4: \ \delta = 4-d\,, \quad [g] = \Lambda^{4-d}\,, \quad \delta \geq 0 \quad \text{for } d \leq 4\,,\\
&g\varphi^3: \ \delta = 3-\tfrac{d}{2}\,, \quad [g] = \Lambda^{3-\frac{d}{2}}\,, \quad \delta \geq 0 \quad \text{for } d \leq 6\,,\\
&g\varphi^6: \ \delta = 6-2d\,, \quad [g] = \Lambda^{6-2d}\,, \quad \delta \geq 0 \quad \text{for } d \leq 3\,.
\end{aligned} \tag{13.14}$$

Excluding the r-form equations (13.9) and (13.13), we obtain

$$D = d - \left(\frac{d}{2} - 1\right) E - n\delta\,. \tag{13.15}$$

In particular, for $d = 4$ we have $D = 4 - E - n\delta$. Now it is clear that the necessary condition for the renormalizability of the theory is $\delta \geq 0$. Previously, for simplicity we have spoken of the dimensionless coupling constants ($\delta = 0$) as the necessary condition for renormalizability. From (13.14) we can see when this condition is fulfilled for the simplest models of interaction. These results show that the dependence on spatial dimensionality is also very important.

In conclusion, we present the table of "canonical" dimensions for different Green's functions and vertices [56]:

Field function	*Dimensionality in units of* Λ	*Dimensionality at* $d = 4$
φ	$\frac{d}{2} - 1$	1
$G^{(n)}(x_1, \ldots, x_n)$	$n\left(\frac{d}{2} - 1\right)$	n
$G^{(n)}(p_1, \ldots, p_n)$	$-nd + n\left(\frac{d}{2} - 1\right) = -n\left(\frac{d}{2} + 1\right)$	$-3n$
$\bar{G}^{(n)}(p_1, \ldots, p_{n-1})$	$d - n\left(\frac{d}{2} + 1\right)$	$4 - 3n$
$\Gamma^{(2)}(x - y)$	$2 + d$	6
$\Gamma^{(n)}(x_1, \ldots x_n)$	$n\left(\frac{d}{2} + 1\right)$	$3n$
$\Gamma^{(n)}(p_1, \ldots, p_n)$	$-dn + n\left(\frac{d}{2} + 1\right) = n\left(1 - \frac{d}{2}\right)$	$-n$
$\bar{\Gamma}^{(n)}(p_1, \ldots, p_{n-1})$	$d + n\left(1 - \frac{d}{2}\right)$	$4 - n$

In addition to the Green's functions and vertices already known to us, we introduced here $\bar{G}^{(n)}$ and $\bar{\Gamma}^{(n)}$, defined as

$$
\begin{aligned}
G^{(n)}(p_1, \ldots, p_n) &= \bar{G}^{(n)}(p_1, \ldots, p_{n-1})\delta(p_1 + \cdots + p_n)\,, \\
\Gamma^{(n)}(p_1, \ldots, p_n) &= \bar{\Gamma}^{(n)}(p_1, \ldots, p_{n-1})\delta(p_1 + \cdots + p_n)\,,
\end{aligned}
\tag{13.16}
$$

where the δ-function of the total momentum conservation is explicitly shown (in units of Λ it has dimensionality $-d$).

13.2 Dimensional regularization of φ^4-theory

To analyze divergences of Feynman diagrams, first of all we have to formulate well defined rules to separate the divergent parts of integrals. This is achieved by one or another method of *regularization* of Feynman integrals. Above (e. g., during our discussion of divergences in QED in Chapter 8) we used the simplest regularization procedure, introducing the upper integration limit cutoff Λ. This method explicitly breaks relativistic invariance, as it is equivalent to the introduction of some "minimal length".

Now we shall discuss a more modern and elegant method of the so-called dimensional regularization (t'Hooft and Veltman). The main idea of this method is somehow similar to the analysis of critical phenomena in the space with $d = 4 - \varepsilon$ dimensions (Wilson) and considers divergent integrals in space-time with continuous $d < 4$, with further limiting procedure of $d \to 4$. We shall see that the singularities of single-loop diagrams considered above are simple poles over the variable $\varepsilon = d - 4$.

First we have to generalize the Lagrangian of 4-dimensional theory

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{m^2}{2}\varphi^2 - \frac{g}{4!}\varphi^4 \tag{13.17}$$

to d dimensions. As field φ has dimensionality $\frac{d}{2} - 1$, while the dimensionality of the Lagrangian $\mathcal{L}$ is d, the coupling constant g is dimensionless for $d = 4$; and to make it dimensionless in d dimensions we have to multiply it μ^{4-d}, where μ is an *arbitrary* parameter of dimensionality of mass (or momentum)[4] . Thus, in the following we shall consider the theory with the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{m^2}{2}\varphi^2 - \frac{1}{4!}g\mu^{4-d}\varphi^4 . \tag{13.18}$$

Let us calculate the simplest self-energy correction shown by the diagram in Figure 13.1. Similarly to (13.1) it is determined by the integral

$$\frac{1}{2}g\mu^{4-d}\int \frac{d^d p}{(2\pi)^d}\frac{1}{p^2 - m^2} . \tag{13.19}$$

This integral should be calculated at arbitrary d.

Integration in *d*-dimensions

We are working in d-dimensional "Minkowski space" with one time and $d - 1$ spatial dimensions ($d \le 4$). Consider an integral of the general form

$$I_d(q) = \int d^d p \frac{1}{(p^2 + 2pq - m^2)^\alpha} , \tag{13.20}$$

where $p = (p_0, \mathbf{r})$. Let us introduce polar coordinates $(p_0, r, \varphi, \theta_1, \theta_2, \ldots, \theta_{d-3})$, so that

$$\begin{aligned} d^d p &= dp_0 r^{d-2} dr d\varphi \sin\theta_1 d\theta_1 \sin^2\theta_2 d\theta_2 \cdots \sin^{d-3}\theta_{d-3} d\theta_{d-3} \\ &= dp_0 r^{d-2} dr d\varphi \prod_{k=1}^{d-3} \sin^k\theta_k d\theta_k , \\ &(-\infty < p_0 < \infty , \ 0 < r < \infty , \ 0 < \varphi < 2\pi , \ 0 < \theta_k < \pi) . \end{aligned} \tag{13.21}$$

Then

$$I_d(q) = 2\pi \int_{-\infty}^{\infty} dp_0 \int_0^{+\infty} dr\, r^{d-2} \int_0^\pi \frac{\prod_{k=1}^{d-3} \sin^k\theta_k d\theta_k}{(p_0^2 - r^2 + 2p_0 q_0 - 2r|\mathbf{q}|\cos\theta_{d-3} - m^2)^\alpha} . \tag{13.22}$$

[4] The arbitrariness of parameter μ is obvious, because at the end we have to perform the limit of $d \to 4$

Direct calculations [56] give

$$I_d(q) = i\pi^{d/2}\frac{\Gamma\left(\alpha-\frac{d}{2}\right)}{\Gamma(\alpha)}\frac{1}{[-q^2-m^2]^{\alpha-d/2}}\,. \tag{13.23}$$

Using (13.23) we obtain for (13.19)

$$-\frac{ig}{32\pi^2}m^2\left(-\frac{4\pi\mu^2}{m^2}\right)^{2-\frac{d}{2}}\Gamma\left(1-\frac{d}{2}\right). \tag{13.24}$$

The Γ-function has poles at zero and at negative integers. We see that the divergence of (13.24) is reflected in the simple pole for $d \to 4$. It can be shown that

$$\Gamma(-n+\varepsilon) = \frac{(-1)^n}{n!}\left[\frac{1}{\varepsilon}+\psi_1(n+1)+O(\varepsilon)\right], \tag{13.25}$$

where $\psi_1(z) = d\ln\Gamma(z)/dz = \Gamma'(z)/\Gamma(z)$ is the logarithmic derivative of the Γ-function, for which $\psi_1(n+1) = 1+\frac{1}{2}+\cdots+\frac{1}{n}-\gamma$, where $\gamma = -\psi_1(1) = 0.577$ is the Euler constant. Taking $\varepsilon = 4-d$ we obtain

$$\Gamma\left(1-\frac{d}{2}\right) = \Gamma\left(-1+\frac{\varepsilon}{2}\right) = -\frac{2}{\varepsilon}-1+\gamma+O(\varepsilon)\,. \tag{13.26}$$

As a result, using $a^\varepsilon = 1+\varepsilon\ln a+\cdots$ we obtain the following expansion of (13.24) near $d = 4$:

$$\begin{aligned}
&-\frac{igm^2}{32\pi^2}\left[-\frac{2}{\varepsilon}-1+\gamma+O(\varepsilon)\right]\left[1+\frac{\varepsilon}{2}\ln\left(-\frac{4\pi\mu^2}{m^2}\right)\right]\\
&\quad=\frac{igm^2}{16\pi^2\varepsilon}+\frac{igm^2}{32\pi^2}\left[1-\gamma+\ln\left(-\frac{4\pi\mu^2}{m^2}\right)\right]+O(\varepsilon)\\
&\quad=\frac{igm^2}{16\pi^2\varepsilon}+\text{finite expression.}
\end{aligned} \tag{13.27}$$

The finite contribution here is of no special importance, but we note that it depends on the arbitrary factor μ. The important point is that we succeeded in a correct separation of the divergent part. For $\varepsilon > 0$ this contribution is finite, and we can deal with it in a normal way.

Let us now calculate a 4-point function up to the terms of the order of $\sim g^2$. Similarly to (13.2) we obtain the contribution of the diagram in Figure 13.2 as

$$-\frac{1}{2}g^2(\mu^2)^{4-d}\int\frac{d^dp}{(2\pi)^d}\frac{1}{p^2-m^2}\frac{1}{(p-q)^2-m^2}\,. \tag{13.28}$$

Denominators in the integrand here can be joined with the help of Feynman's formula

$$\frac{1}{ab} = \int_0^1\frac{dz}{[az+b(1-z)]^2}\,. \tag{13.29}$$

This formula is derived from

$$\frac{1}{ab} = \frac{1}{b-a}\left(\frac{1}{a} - \frac{1}{b}\right) = \frac{1}{b-a}\int_a^b \frac{dx}{x^2}, \tag{13.30}$$

taking $x = az + b(1-z)$, with a and b complex, to exclude singularity at $a = b$. Now we have

$$\frac{1}{p^2-m^2}\frac{1}{(p-q)^2-m^2} = \int_0^1 \frac{dz}{[p^2-m^2-2pq(1-z)+q^2(1-z)]^2}. \tag{13.31}$$

Changing variables to $p' = p - q(1-z)$, we see that the denominator of the integrand is the square of $p'^2 - m^2 + q^2 z(1-z)$. We have $d^d p' = d^d p$, so that after the change of notations $p' \to p$, (13.28) takes the form

$$-\frac{1}{2}g^2(\mu^2)^{4-d}\int_0^1 dz \int \frac{d^d p}{(2\pi)^d}\frac{1}{[p^2-m^2+q^2z(1-z)]^2}. \tag{13.32}$$

Using (13.23) we now have

$$\frac{ig^2}{2}(\mu^2)^{4-d}\left(\frac{1}{4\pi}\right)^{d/2}\frac{\Gamma(2-d/2)}{\Gamma(2)}\int_0^1 dz[q^2z(1-z)-m^2]^{d/2-2} =$$
$$\frac{ig^2}{32\pi^2}(\mu^2)^{2-d/2}\Gamma\left(2-\frac{d}{2}\right)\int_0^1 dz\left[\frac{q^2z(1-z)-m^2}{4\pi\mu^2}\right]^{\frac{d}{2}-2}. \tag{13.33}$$

In the limit of $d \to 4$, from (13.25) we get

$$\Gamma\left(2-\frac{d}{2}\right) = \frac{2}{\varepsilon} - \gamma + O(\varepsilon), \tag{13.34}$$

so that after writing $a^\varepsilon \approx 1 + \varepsilon \ln a$ (13.33) take the form

$$\frac{ig^2\mu^\varepsilon}{32\pi^2}\left(\frac{2}{\varepsilon} - \gamma + O(\varepsilon)\right)\left\{1 - \frac{\varepsilon}{2}\int_0^1 dz \ln\left[\frac{q^2z(1-z)-m^2}{4\pi\mu^2}\right]\right\} =$$
$$\frac{ig^2\mu^\varepsilon}{16\pi^2\varepsilon} - \frac{ig^2\mu^\varepsilon}{32\pi^2}\left\{\gamma + \int_0^1 \ln\left[\frac{q^2z(1-z)-m^2}{4\pi\mu^2}\right]\right\}. \tag{13.35}$$

In this expression, the main (diverging) term depends on μ, while the finite part depends on the square of $(p_1+p_2)^2 = q^2 = s$ (Mandelstam variable). Let us define ta function

$$F(s,m,\mu) = \int_0^1 dz \ln\left[\frac{sz(1-z)-m^2}{4\pi\mu^2}\right]. \tag{13.36}$$

Then the final expression for the contribution of the diagram of Figure 13.2 is written as

$$-\frac{ig^2\mu^\varepsilon}{16\pi^2\varepsilon} + \frac{ig^2\mu^\varepsilon}{32\pi^2}[\gamma + F(s,m,\mu)] = -\frac{ig^2\mu^\varepsilon}{16\pi^2\varepsilon} + \text{finite expression}. \tag{13.37}$$

Thus, we have explicitly written the lowest order corrections to 2-point and 4-point functions in φ^4-theory. Let us now write the corresponding irreducible vertices $\Gamma^{(2)}(p)$

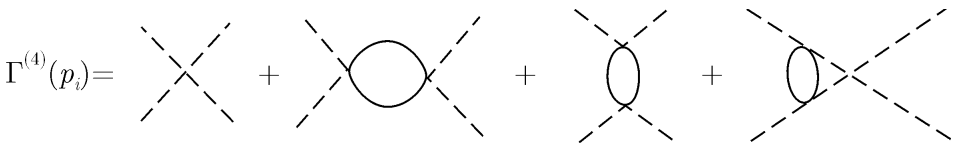


Figure 13.4

and $\Gamma^{(4)}(p_i)$. Equation (13.27) according to (13.1) reduces to $\frac{1}{i}\Sigma$, so that in the first order over g we have

$$\Sigma(p) = -\frac{g m^2}{16\pi^2 \varepsilon} + \text{finite expression.} \tag{13.38}$$

Accordingly, from the definition $\Gamma^{(2)}(p) = G^{-1}(p)G(p)G^{-1}(p) = p^2 - m^2 - \Sigma(p)$ we have

$$\Gamma^{(2)}(p) = G^{-1}(p) = p^2 - m^2 \left(1 - \frac{g}{16\pi^2 \varepsilon}\right). \tag{13.39}$$

Obviously, for $\varepsilon \to 4$ this expression diverges.

Further on, the 4-point vertex $\Gamma^{(4)}(p_1, \ldots, p_4)$ in momentum representation is written as

$$\Gamma^{(4)}(p_1, p_2, p_3, p_4) = G^{-1}(p_1)G^{-1}(p_2)G^{(4)}(p_1, p_2, p_3, p_4)G^{-1}(p_3)G^{-1}(p_4) \tag{13.40}$$

and is expressed by the sum of the diagrams shown in Figure 13.4, taking into account contributions of all the cross-channels, which are obtained from (13.37) and another two similar terms obtained from (13.37), changing the Mandelstam variable s to t and u (cf. Chapter 5):

$$s = (p_1 + p_2)^2, \quad t = (p_1 + p_3)^2, \quad u = (p_1 + p_4)^2. \tag{13.41}$$

The action of $G^{-1}(p_i)$ in (13.40) reduces to the "amputation" of the external lines. Finally we obtain

$$\begin{aligned} \Gamma^{(4)}(p_i) &= -ig\mu^{\varepsilon} - \frac{3ig^2\mu^{\varepsilon}}{16\pi^2\varepsilon} + \frac{ig^2\mu^{\varepsilon}}{32\pi^2}[3\gamma + F(s,m,\mu) + F(t,m,\mu) + F(u,m,\mu)] \\ &= -ig\mu^{\varepsilon}\left(1 + \frac{3g}{16\pi^2\varepsilon}\right) + \text{finite expression.} \end{aligned} \tag{13.42}$$

The main contribution here is also infinite for $\varepsilon \to 0$. To make vertices $\Gamma^{(2)}$ and $\Gamma^{(4)}$ physically sensible, we should make them finite! This is done by renormalization!

Loop expansion

Note that above we analyzed diagrams with the same number of loops equal to 1 (single-loop approximation). Actually, we can present arguments showing that loop-expansion is, in some sense, even more interesting than the usual perturbation expansion in powers of g. Expansion

in the number of loops L is equivalent to expansion in the powers of the Planck constant $\hbar$. In fact, restoring $\hbar$ in all expressions, we can write the generating functional as

$$Z[J(x)] = \int \mathcal{D}\varphi \exp\left\{\frac{i}{\hbar}\int dx[\mathcal{L}(x) + \hbar J(x)\varphi(x)]\right\}. \tag{13.43}$$

Introducing $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{int}$, we can write

$$Z[J] = \exp\left\{\frac{i}{\hbar}\mathcal{L}_{int}\left[\frac{1}{i}\frac{\delta}{\delta J}\right]\right\} Z_0[J], \tag{13.44}$$

where

$$Z_0[J] = \mathcal{N}\exp\left[-\frac{1}{2}i\hbar\int dx\int dy J(x)\Delta_F(x-y)J(y)\right]. \tag{13.45}$$

From (13.44) it follows that every vertex introduces the factor of $\hbar^{-1}$ into an arbitrary diagram of the n-th order of the usual perturbation theory, while from (13.45) it follows that each propagator contributes the factor of $\hbar$. Thus, the given diagram contains the factor of $\hbar^{I-n} = \hbar^{L-1}$ (where we have used the previously derived relation (13.4): $L = I - n + 1$, where I is the number of internal lines of this diagram). Then we conclude that the expansion over the number of loops is actually an expansion in the powers of $\hbar$, i. e., the expansion "around" the classical theory.

13.3 Renormalization of φ^4-theory

Our aim now is to make all physical quantities finite! In a single-loop approximation we can easily explicitly make renormalization. After the regularization, all the quantities we are dealing with are finite, and we can act in a direct way. From the definition of the physical mass of the particle it is clear that the inverse propagator *must* be of the form

$$G^{-1}(p) = \Gamma^{(2)}(p) = p^2 - m_1^2 \quad \text{or} \quad m_1^2 = -\Gamma^{(2)}(0) = -G^{-1}(0), \tag{13.46}$$

where the physical mass m_1 is finite. The initial ("bare") mass m entering the Lagrangian does not have any direct physical meaning and can even be infinite in the limit of $d \to 4$. This is a mass which characterizes a particle in the absence of interactions, which is unobservable; only m_1 is physically sensible, and it should be finite. From (13.39) and definition (13.46) we have

$$m_1^2 = m^2\left(1 - \frac{g}{16\pi^2\varepsilon}\right). \tag{13.47}$$

In the second term in the right-hand side we may, with the same accuracy $\sim g$, replace m by m_1, which gives

$$m_1^2 = m^2 - \frac{g}{16\pi\varepsilon}m_1^2, \tag{13.48}$$

so that we get

$$m^2 = m_1^2\left(1 + \frac{g}{16\pi^2\varepsilon}\right). \tag{13.49}$$

This is the value of the "bare" mass guaranteeing the finite value m_1 of the physical mass in single-loop approximation. We see that for $\varepsilon \to 0$ the value of m diverges, but m_1 is finite!

We can similarly analyze the vertex part $\Gamma^{(4)}$. We rewrite (13.42) as

$$i\Gamma^{(4)}(p_i) = g\mu^{\varepsilon} + \frac{g^2\mu^{\varepsilon}}{32\pi^2}\left[\frac{6}{\varepsilon} - 3\gamma - F(s,m,\mu) - F(t,m,\mu) - F(u,m,\mu)\right]. \tag{13.50}$$

Let us define the renormalized (finite!) coupling constant g_1 as

$$g_1 = i\Gamma^{(4)}(p_i = 0)\,, \tag{13.51}$$

i. e., as the vertex part for particles with zero momenta. Then, from (13.50) we obtain

$$g_1 = g\mu^{\varepsilon} + \frac{g^2\mu^{\varepsilon}}{32\pi^2}\left[\frac{6}{\varepsilon} - 3\gamma - 3F(0,m,\mu)\right]. \tag{13.52}$$

Considering g_1 to be fixed and finite, we immediately see that the "bare" coupling constant g should be infinite (for $\varepsilon \to 0$). In fact, rewriting equation (13.52) replacing g by $g_1\mu^{-\varepsilon}$ and m by m_1 everywhere (which always can be done with an accuracy up to terms $\sim g^2$), we can obtain an expression for g, expressed via g_1, similar to equation (13.49):

$$g = g_1\mu^{-\varepsilon} - \frac{3g_1^2\mu^{-2\varepsilon}}{32\pi^2}\left[\frac{2}{\varepsilon} - \gamma - F(0,m_1,\mu)\right]. \tag{13.53}$$

Then we can express $\Gamma^{(4)}$ (13.50) via g_1 as

$$i\Gamma^{(4)}(p_i) = g_1 - \frac{g_1^2\mu^{-\varepsilon}}{32\pi^2}[F(s,m_1,\mu) + F(t,m_1,\mu) + F(u,m_1,\mu) - 3F(0,m_1,\mu)]\,. \tag{13.54}$$

From here (13.51) follows directly, as for $p_1 = p_2 = p_3 = p_4 = 0$ we have $s = t = u = 0$. Thus, the physical (renormalized) coupling constant g_1 coincides with $i\Gamma^{(4)}$ with all external momenta equal to zero[5]. Now everything is finite! We completed renormalization in a single-loop approximation.

Now, how does this look in two-loop approximation? In this case we have to consider the diagrams shown in Figure 13.5. Appropriate analysis shows that in this case $G^{-1}(p) = \Gamma^{(2)}(p)$ acquires an additional divergence due to the diagram of Figure 13.5(b). This divergence is not cancelled by mass and coupling constant renormalization. It is hidden into an additional multiplicative factor, which is introduced by redefining 2-point function as

$$G_r^{-1} = \Gamma_r^{(2)} = Z_\varphi(g_1,m_1,\mu)\Gamma^{(2)}(p,m_1)\,. \tag{13.55}$$

[5] This is not the only way to define the renormalized coupling constant. Sometimes g_1 is defined via $i\Gamma^{(4)}$ at the so-called symmetric point $p_i^2 = m^2$, $p_i p_j = -m^2/3$ $(i \neq j)$, which corresponds to $s = t = u = 4m^2/3$.

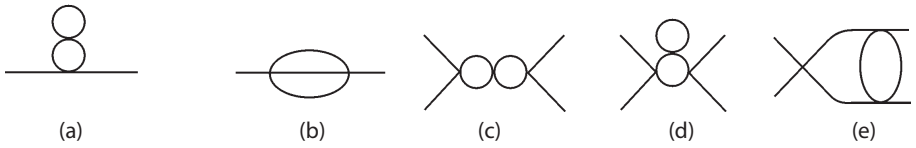


Figure 13.5

Here $\Gamma_r^{(2)}$ is finite, while the factor Z_φ is infinite. The factor of $Z_\varphi^{1/2}$ is called wave function renormalization. For Z_φ it is possible to write an expansion in the number of loops, which has the form

$$Z_\varphi = 1 + g_1 Z_1 + g_1^2 Z_2 + \cdots = 1 + g_1^2 Z_2 + \cdots , \tag{13.56}$$

because the single-loop contribution is absent. Wave function renormalization (renormalization of the field amplitude) cannot be just arbitrary. To define it we need to require that at some point, e. g., at $p^2 = 0$, we have

$$\frac{\partial}{\partial p^2} G_r^{-1}(p)\bigg|_{p^2=0} = \frac{\partial}{\partial p^2}\Gamma_r^{(2)}\bigg|_{p^2=0} = 1 . \tag{13.57}$$

The choice of $p^2 = 0$ is more or less arbitrary.

The divergence of Z_φ means that in a two-loop approximation the value of m_1 defined above is actually infinite (in the limit of $\varepsilon \to 0$). However, renormalized $G_r^{-1}(p) = \Gamma_r^{(2)}$ gives the finite value of renormalized mass m_r:

$$m_r^2 = Z_\varphi m_1^2 . \tag{13.58}$$

In other words, divergences of Z_φ and m_1^2 cancel each other. The value of renormalized coupling constant changes similarly. For $\Gamma_r^{(4)}$ we have the relation similar to (13.55):

$$\Gamma_r^{(4)} = Z_\varphi^2 \Gamma^{(4)}(p, m_1, \mu) \tag{13.59}$$

and the new renormalized coupling constant g_r, defined by the relation similar to (13.46), has the form

$$i\,\Gamma_r^{(4)}(p_i = 0) = g_r = Z_\varphi^2 g_1 . \tag{13.60}$$

The factor Z_φ is the function of $g\mu^\varepsilon$, so that writing this dependence explicitly we obtain the renormalized n-particle vertex part as

$$\Gamma_r^{(n)}(p_i, g_r, m_r, \mu) = Z_\varphi^{n/2}(g\mu^\varepsilon)\Gamma^{(n)}(p_i, g, m) \tag{13.61}$$

or

$$\Gamma^{(n)}(p_i, g, m) = Z_\varphi^{-n/2}(g\mu^\varepsilon)\Gamma_r^{(n)}(p_i, g_r, m_r, \mu) . \tag{13.62}$$

Thus, in two-loop approximation we also can make our theory finite. Is it so in any order? This is the problem of the *proof of renormalizability*. This proof is tedious

enough, but it can actually be done in all orders of perturbation theory (Dyson). A detailed presentation of this proof for different models of quantum field theory can be found in [9]. Note, that the proof of renormalizability in φ^4-theory is actually more difficult than the similar proof for QED, which is made more simple due to the gauge invariance.

Counter-terms

There is an alternative way to introduce renormalizability, which became popular after the publication of [9]. The point is that from the start we can consider parameters m and g in an initial Lagrangian as the physical mass and charge (coupling constant). The fact that this Lagrangian does not produce finite Green's functions now leads to the requirement that we introduce into the Lagrangian some additional terms which cancel the divergences. These terms are called *counter-terms*. Renormalized theory can be made finite by the introduction of the *finite* number of counter-terms. Let us briefly describe how this is done.

Consider again mass renormalization in a single-loop approximation, which is defined by equations (13.46)–(13.49). This may be described as follows. A single-loop correction to a free propagator is shown in Figure 13.6 and diverges for $\varepsilon \to 0$. Let us add to the initial Lagrangian $\mathcal{L}$ the term

$$\delta\mathcal{L}_1 = -\frac{gm^2}{32\pi^2\varepsilon}\varphi^2 \equiv -\frac{1}{2}\delta m^2\varphi^2\,. \tag{13.63}$$

This may be considered as an additional interaction, which we shall denote by the "cross" on the diagrams:

$$\times = -\frac{igm^2}{16\pi^2\varepsilon} = -i\delta m^2\,. \tag{13.64}$$

Then, up to terms of the order of $\sim g$, the total inverse propagator is represented by diagrams in Figure 13.7 and is equal to

$$\begin{aligned}\Gamma^{(2)}(p) &= iG(p)^{-1}\\ &= i\left[\frac{1}{i}(p^2-m^2) - \left(\frac{igm^2}{16\pi^2}\frac{1}{\varepsilon} + \text{finite part}\right) + \frac{igm^2}{16\pi^2\varepsilon}\right]\\ &= p^2 - m^2\,,\end{aligned} \tag{13.65}$$

where we have dropped the finite contribution (alternatively we can include it into m^2). Here m^2 is considered to be the finite physical mass, which in corresponding order of

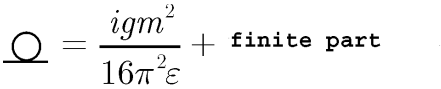


Figure 13.6

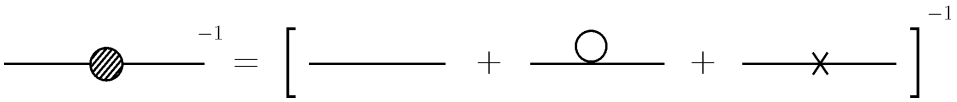


Figure 13.7

perturbation theory is equal to $-\Gamma^{(2)}(0)$. The Lagrangian is now $\mathcal{L}+\delta\mathcal{L}_1$, where $\delta\mathcal{L}_1$ is a diverging counter-term.

The meaning of the introduction of a mass term in the Lagrangian as an additional interaction is rather simple. Consider noninteracting theory:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\varphi)(\partial^\mu\varphi) - \frac{1}{2}m^2\varphi \tag{13.66}$$

and assume,that it describes massless field (first term in Lagrangian), with interaction determined by the second terms. Corresponding Feynman rules are shown in Figure 13.8. The full propagator is determined now by the diagrams shown in Figure 13.9. Then, the perturbation series reduces to the simple geometric progression

$$G(p) = \frac{i}{p^2} + \frac{i}{p^2}(-im^2)\frac{i}{p^2} + \frac{i}{p^2}(-im^2)\frac{i}{p^2}(-im^2)\frac{i}{p^2} + \cdots = \frac{i}{p^2-m^2}, \tag{13.67}$$

which gives the usual propagator of the massive field. This we actually used while considering the mass counter-term as perturbation.

In a similar way we can also deal with $\Gamma^{(4)}$. From (13.42) it is seen, that $\Gamma^{(4)}$, corresponding to diagrams $\sim g^2$, shown in Figure 13.10, diverges for $\varepsilon \to 0$. Then we can add to the Lagrangian the counter-term

$$\delta\mathcal{L}_2 = \frac{1}{4!}\frac{3g^2\mu^\varepsilon}{16\pi^2\varepsilon}\varphi^4 = \frac{Bg\mu^\varepsilon}{4!}\varphi^4, \tag{13.68}$$

corresponding to the additional interaction, shown in Figure 13.11. As a result $\Gamma^{(4)}$, as shown in Figure 13.12, becomes finite. The divergence of $\Gamma^{(2)}$ in a two-loop approximation, qualitatively described above and leading to the necessity to multiply $\Gamma^{(n)}$ by

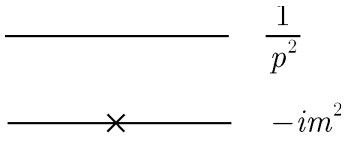


Figure 13.8

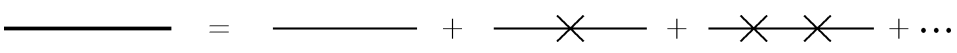


Figure 13.9

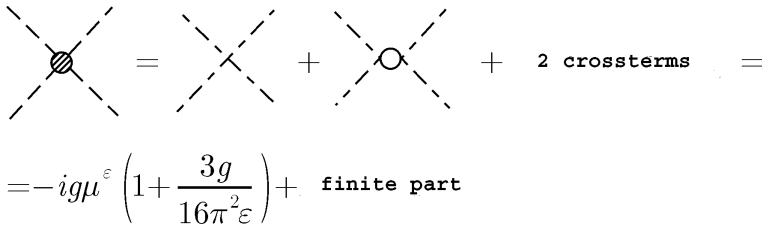


Figure 13.10

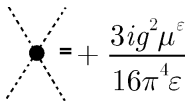


Figure 13.11

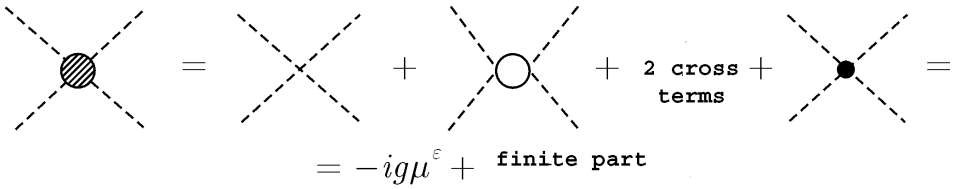


Figure 13.12

$Z_{\varphi}^{n/2}$, is equivalent to the addition to the Lagrangian of the counter-term

$$\delta\mathcal{L}_3 = \frac{A}{2}(\partial_\mu\varphi)^2\,, \tag{13.69}$$

where $1 + A = Z_\varphi$.

Thus, finite expressions for Green's functions and vertices can be obtained by adding to the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}m^2\varphi^2 - \frac{1}{4!}g\mu^{4-d}\varphi^4 \tag{13.70}$$

the counter-term $\mathcal{L}_{CT}$:

$$\mathcal{L}_{CT} = \frac{A}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}\delta m^2\varphi^2 - \frac{1}{4!}Bg\mu^{4-d}\varphi^4\,. \tag{13.71}$$

The total Lagrangian, which is usually called the "bare" Lagrangian $\mathcal{L}_B$, is equal to

$$\mathcal{L}_B = \mathcal{L} + \mathcal{L}_{CT} = \frac{1+A}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}(m^2+\delta m^2)\varphi^2 - \frac{1}{4!}(1+B)g\mu^{4-d}\varphi^4\,. \tag{13.72}$$

Thus, the addition of counter-terms is equivalent to the multiplication of φ, m, and g by some renormalization factors Z (multiplicative renormalization). If we define "bare" quantities

$$\begin{aligned} \varphi_B &= \sqrt{Z_\varphi}\varphi_r\,, \qquad & Z_\varphi &= 1 + A\,, \\ m_B &= Z_m m_r\,, & Z_m^2 &= \frac{m^2 + \delta m^2}{1 + A}\,, \\ g_B &= \mu^\varepsilon Z_g g_r\,, & Z_g &= \frac{1 + B}{(1 + A)^2}\,, \end{aligned} \tag{13.73}$$

the "bare" Lagrangian (13.72) is written as

$$\mathcal{L}_B = \frac{1}{2}\partial_\mu\varphi_B\partial^\mu\varphi_B - \frac{1}{2}m_B^2\varphi_B^2 - \frac{1}{4!}g_B\varphi_B^4\,. \tag{13.74}$$

Note that here there is no *explicit* dependence on μ. The values of A, B, and δm^2 are assumed to be chosen in such a way as to make Green's functions finite (for $\varepsilon \to 0$). In terms of the counter-terms approach, the theory is renormalizable if the counter-terms needed to cancel the divergences in every order of perturbation theory have the same form as the terms entering the initial Lagrangian. If this is so, the "bare" quantities can be defined with (infinite!) renormalization factors, as was done above, and the "bare" Lagrangian has the same form as the initial Lagrangian.

Lagrangian $\mathcal{L}_B$ leads to finite theory, in contrast to initial $\mathcal{L}$. This means that, "hiding" all divergences into φ_B, m_B and g_B, we can make theory finite: the divergences are absorbed by renormalization. All "bare" quantities are divergent for $\varepsilon \to 0$ [6], while renormalized quantities are finite for $\varepsilon \to 0$, but their values are more or less arbitrary and should be taken to be equal to the physical parameters of the theory.

Equation (13.62) is also obvious from the counter-terms approach. From equations (13.73) and (13.74) it is clear that, taking (13.74) as the initial Lagrangian, we have to replace $m \to m_B$, $g \to g_B$, $\varphi \to \varphi_B$ in all expressions for Green's functions. But now we can (and need to!) express the "bare" parameters via physical m_r, g_r, and φ_r according to expressions (13.73). Then we obtain

$$\Gamma_B^{(n)}(p_i, g_B, m_B) = Z_\varphi^{-n/2}\Gamma_r^{(n)}(p_i, g_r, m_r, \mu)\,, \tag{13.75}$$

which is equivalent to (13.62) (index B can now be dropped). The absence of explicit dependence of the left-hand side of this equation on μ is obvious from the form of the Lagrangian (13.74), where it is also absent.

[6] For finite ε there are no divergences at all!

13.4 The renormalization group

In Chapter 8 we already discussed briefly the renormalization group in QED. Renormalization groups plays the major role in quantum field theory [9,28,56] and statistical physics [3,42], as well as in some other fields of theoretical physics. Below we shall present a more detailed discussion. There are several (more or less equivalent) formulations of this method. For example, in Chapter 8 the renormalization group was related to transformations from one value of the cutoff parameter (for divergent integrals) to another, in the theory of critical phenomena [42] Wilson's formulation is quite popular, which is based on integrating out regions of momentum space, corresponding to large momenta, i. e., restricting the analysis of fluctuations to long enough wavelengths etc. Here we shall use the most common (though probably more formal) approach used in quantum field theory literature, which is based on dimensional regularization [56].

Within the framework of dimensional regularization we have introduced an arbitrary parameter μ with dimensionality of the mass. Dependence of renormalized irreducible vertices on μ is determined, according to equation (13.61), by a corresponding μ-dependence of renormalization factor Z_φ. In other words (cf. (13.62), (13.75)) the nonrenormalized ("bare") function $\Gamma^{(n)}$ does not depend on μ:

$$\Gamma^{(n)}(p_i, g, m) = Z_\varphi^{-n/2}(g\mu^\varepsilon)\Gamma_r^{(n)}(p_i, g_r, m_r, \mu) \tag{13.76}$$

and, in this sense, is invariant towards the group of transformations

$$\mu \to e^s\mu \quad \text{or} \quad \mu = e^s\mu_0 \quad \text{i. e.,} \quad s = \ln\frac{\mu}{\mu_0}\,. \tag{13.77}$$

These transformations represent the renormalization group. Introducing the dimensionless differential operator $\mu\frac{\partial}{\partial\mu}$, we get

$$\mu\frac{\partial}{\partial\mu}\Gamma^{(n)} = 0 \tag{13.78}$$

or, taking into account (13.76),

$$\mu\frac{\partial}{\partial\mu}[Z_\varphi^{-n/2}(g\mu^\varepsilon)\Gamma_r^{(n)}(p_i, g_r, m_r, \mu)] = 0\,, \tag{13.79}$$

where g_r and m_r depend on μ. Making a differentiation and multiplying the result by $Z_\varphi^{n/2}$, we obtain

$$\left[-n\mu\frac{\partial}{\partial\mu}\ln\sqrt{Z_\varphi} + \mu\frac{\partial}{\partial\mu} + \mu\frac{\partial g_r}{\partial\mu}\frac{\partial}{\partial g_r} + \mu\frac{\partial m_r}{\partial\mu}\frac{\partial}{\partial m_r}\right]\Gamma_r^{(n)} = 0\,. \tag{13.80}$$

In the following, for brevity we shall everywhere write g instead of g_r and m instead of m_r, assuming that we are dealing only with renormalized quantities. In general, only renormalized quantities enter equation (13.80), which are finite for $\varepsilon \to 0$.

Let us define the following functions:

$$
\begin{aligned}
m\gamma_m(g) &= \mu\frac{\partial m}{\partial \mu}\,,\\
\gamma(g) &= \mu\frac{\partial}{\partial \mu}\ln\sqrt{Z_\varphi}\,,\\
\beta(g) &= \mu\frac{\partial g}{\partial \mu}\,.
\end{aligned}
\tag{13.81}
$$

Then equation (13.80) takes the form

$$
\left[\mu\frac{\partial}{\partial \mu}+\beta(g)\frac{\partial}{\partial g}-n\gamma(g)+m\gamma_m(g)\frac{\partial}{\partial m}\right]\Gamma^{(n)}=0\,. \tag{13.82}
$$

This is the main differential equation of the renormalization group, usually called a Callan–Symanzik equation. It reflects invariance of the renormalized vertex $\Gamma^{(n)}$ to changes of the regularization parameter μ[7].

Let us write a similar equation, expressing the invariance of $\Gamma^{(n)}$ to changes of the momentum scale (mass). Consider the replacement $p_i \to tp_i, m \to tm, \mu \to t\mu$. The vertex $\Gamma^{(n)}$ has mass dimensionality D, determined according to the table presented above, by the following expression

$$
D = d + n\left(1-\frac{d}{2}\right) = 4-n+\varepsilon\left(\frac{n}{2}-1\right)\,, \tag{13.83}
$$

where $d = 4-\varepsilon$. Then we have

$$
\Gamma^{(n)}(tp_i, tm, t\mu) = t^D\Gamma^{(n)}(p_i, m, \mu)\,, \tag{13.84}
$$

which, after the simple variable changes $tm \to \tilde{m},\ m \to \tilde{m}/t,\ \tilde{m} \to m$, and $t\mu \to \tilde{\mu},\ \mu \to \tilde{\mu}/t,\ \tilde{\mu} \to \mu$, is rewritten as

$$
\Gamma^{(n)}(tp_i, m, \mu) = t^D\Gamma^{(n)}(p_i, m/t, \mu/t)\,. \tag{13.85}
$$

Thus, $\Gamma^{(n)}$ is actually the homogeneous function of its variables of the power D.

Homogeneous functions. The Euler theorem

Let us recall the basic facts about homogeneous functions. The function $u = f(x_1, x_2, \ldots, x_m)$ is called a homogeneous function of power p, if for any t we have

$$
u = f(tx_1, \ldots, tx_m) = t^p f(x_1, \ldots, x_m)\,. \tag{13.86}
$$

For homogeneous functions we have the Euler theorem

$$
x_1\frac{\partial u}{\partial x_1}+\cdots+x_m\frac{\partial u}{\partial x_m} = pu\,. \tag{13.87}
$$

[7] To avoid misunderstanding we note that here we are dealing with the vertex, defined in (13.16) and denoted previously by $\bar{\Gamma}^{(n)}$.

To prove this, consider $u = f(tx_1^0, \ldots, tx_m^0)$, where $(x_1^0, \ldots, x_m^0)$ is an arbitrary point, from the region of the definition of our function. Then we have

$$\left.\frac{du}{dt}\right|_{t=1} = \frac{\partial u}{\partial x_1}x_1^0 + \ldots + \frac{\partial u}{\partial x_m}x_m^0 . \tag{13.88}$$

On the other hand,

$$\begin{aligned} &\frac{du}{dt} = pt^{p-1} f(x_1^0, \ldots, x_m^0) , \qquad \text{so that} \\ &\left.\frac{du}{dt}\right|_{t=1} = pf(x_1^0, \ldots, x_m^0) = pu . \end{aligned} \tag{13.89}$$

Comparison of (13.88) with (13.89) gives (13.87).

From (13.85), using the Euler theorem, we have

$$\left(t\frac{\partial}{\partial t} + m\frac{\partial}{\partial m} + \mu\frac{\partial}{\partial \mu} - D\right)\Gamma^{(n)}(tp_i, g, m, \mu) = 0 . \tag{13.90}$$

Excluding $\mu\frac{\partial \Gamma^{(n)}}{\partial \mu}$ from (13.82) and (13.90), we obtain another form of the Callan–Symanzik equation:

$$\left[-t\frac{\partial}{\partial t} + \beta\frac{\partial}{\partial g} - n\gamma(g) + m(\gamma_m(g) - 1)\frac{\partial}{\partial m} + D\right]\Gamma^{(n)}(tp_i, g, m, \mu) = 0 , \tag{13.91}$$

which directly expresses the result of the scale change of momenta in $\Gamma^{(n)}$ by the factor of t. Note that for $\beta(g) = \gamma(g) = 0$, $\gamma_m(g) = 1$, this result reduces to the *canonical* dimension D, which is determined by the "naïve" dimensional analysis. The necessity of renormalization and nontrivial values of $\beta(g), \gamma(g), \gamma_m(g)$ is directly related to interactions which lead to *anomalous* dimensions.

Let us find the solution of equation (13.91). In fact, this equation reflects the fact that the change of t can be compensated by an appropriate change of m and g and of the common factor. Suppose the existence of functions $g(t)$, $m(t)$ and $f(t)$, such that

$$\Gamma^{(n)}(tp, m, g, \mu) = f(t)\Gamma^{(n)}(p, m(t), g(t), \mu) . \tag{13.92}$$

Differentiating by t, we obtain

$$\begin{aligned} &\frac{\partial}{\partial t}\Gamma^{(n)}(tp, m, g, \mu) = \\ &\qquad \frac{df(t)}{dt}\Gamma^{(n)}(p, m(t), g(t), \mu) + f(t)\left(\frac{\partial m}{\partial t}\frac{\partial \Gamma^{(n)}}{\partial m} + \frac{\partial g}{\partial t}\frac{\partial \Gamma^{(n)}}{\partial g}\right), \end{aligned} \tag{13.93}$$

or, taking into account (13.92),

$$t\frac{\partial}{\partial t}\Gamma^{(n)}(tp,m,g,\mu)$$
$$= \left(t\frac{df(t)}{dt} + f(t)t\frac{\partial m}{\partial t}\frac{\partial}{\partial m} + f(t)t\frac{\partial g}{\partial t}\frac{\partial}{\partial g}\right)\Gamma^{(n)}(p,m(t),g(t),\mu)$$
$$= \left(t\frac{df(t)}{dt} + tf(t)\frac{\partial m}{\partial t}\frac{\partial}{\partial m} + tf(t)\frac{\partial g}{\partial t}\frac{\partial}{\partial g}\right)\frac{1}{f(t)}\Gamma^{(n)}(tp,m,g,\mu)\,, \quad (13.94)$$

which, after the regrouping, reduces to

$$\left(-t\frac{\partial}{\partial t} + \frac{t}{f(t)}\frac{df(t)}{dt} + t\frac{\partial m}{\partial t}\frac{\partial}{\partial m} + t\frac{\partial g}{\partial t}\frac{\partial}{\partial g}\right)\Gamma^{(n)}(tp,m,g,\mu) = 0\,. \quad (13.95)$$

Compare now (13.91) with (13.95). Equating the coefficients at $\partial/\partial g$, we obtain the Gell-Mann–Low equation

$$t\frac{\partial g(t)}{\partial t} = \beta(g)\,. \quad (13.96)$$

The function $g(t)$ is called the "running" coupling constant, and the $\beta(g)$-function is called the Gell-Mann–Low function. This equation is of basic importance in the study of asymptotic properties in quantum field theory. The knowledge of $\beta(g)$ allows us to find $g(t)$. Of major interest, as we shall see, is the asymptotics of $g(t)$ at $t \to \infty$. The initial condition for equation (13.96) is $g(1) = g$.

Comparison of the coefficients before $\partial/\partial m$ in (13.91) and (13.95) gives

$$t\frac{\partial m}{\partial t} = m[\gamma_m(g) - 1]\,, \quad (13.97)$$

and comparison of the remaining terms gives

$$\frac{t}{f(t)}\frac{df(t)}{dt} = D - n\gamma(g)\,. \quad (13.98)$$

The last equation can be integrated to obtain

$$f(t) = t^D \exp\left[-\int_0^t dt\frac{n\gamma(g(t))}{t}\right]; \quad (13.99)$$

substituting this into (13.92) and taking $D = 4 - n + \varepsilon\left(\frac{n}{2} - 1\right)$, in the limit of $\varepsilon \to 0$, we get

$$\Gamma^{(n)}(tp,m,g,\mu) = t^{4-n}\exp\left[-n\int_0^t dt\frac{\gamma(g(t))}{t}\right]\Gamma^{(n)}(p,m(t),g(t),\mu)\,. \quad (13.100)$$

This is the solution of (13.91), expressed via the "running" coupling constant $g(t)$ and the "running" mass $m(t)$. The exponential determines the anomalous dimension. Thus, the physics at high momenta is determined by functions $g(t)$ and $m(t)$. Relations like (13.100) in some sense allow us to analyze the situation outside the region of applicability of perturbation theory.

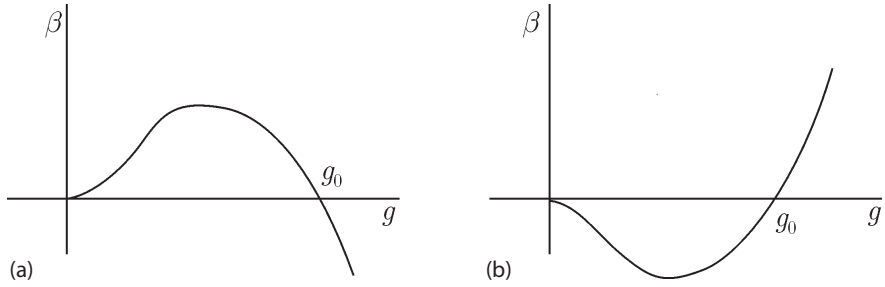


Figure 13.13

In the limit of very large momenta we can neglect particle masses. Thus, we can usually limit our analysis to studies of the Gell-Mann–Low equation (13.96). Consider the possible qualitative behavior which may appear. We shall be interested in the behavior of $g(t)$ for $t \to \infty$. The Gell-Mann–Low equation is written as (13.96), and the possible variants of the qualitative behavior of the $\beta(g)$-function are shown in Figure 13.13. We always have $\beta(g = 0) = 0$, which corresponds to free theory without interactions. Perturbation theory allows us to determine the behavior of $\beta(g)$ close to $g = 0$; it is always (as we shall see below) quadratic in g. In principle, where zeroes of $\beta(g)$ at finite g, may be, it is sufficient to consider only one, say at $g = g_0$, to understand the consequences of its existence. Consider first $\beta(g)$, shown in Figure 13.13(a). The zeroes of this function at $g = 0$ and $g = g_0$ correspond to the so-called *fixed points* of the Gell-Mann–Low equation. It is easy to see that, for $t \to \infty$ and initial values of g close to g_0, the value of $g(t)$, determined from (13.96), tends to g_0. In fact, for initial $g < g_0$ we have $\beta(g) > 0$, so that g grows with the growth of t and tends to g_0 (where further growth is just stopped). Similarly, for initial $g > g_0$ we have $\beta(g) < 0$, and g diminishes with the growth of t, i. e., also tends to g_0, moving in a negative direction. Thus $g(\infty) = g_0$, and we have an *ultraviolet* stable fixed point: the fixed value of the coupling constant (charge) at very large momenta. For small initial values of g in the limit of $t \to 0$ we always obtain $g = 0$, the *infrared* stable fixed point ("Moscow zero"). If the zeroes of the Gell-Mann–Low function at finite g are absent, equation (13.96) leads to the continuous growth of g for $t \to \infty$, and a fixed value of charge does not appear. If for the large values of argument $\beta(g) \sim g^\alpha$ and $\alpha > 1$, the theory becomes internally inconsistent: the inevitable divergence of g appears at the *finite* value of t (the Landau "ghost pole"). For $\alpha \leq 1$ we obtain the monotonous growth of g for $t \to \infty$; the theory is consistent, but for $t \to \infty$ we have a crossover to "strong coupling".

Consider now the $\beta(g)$-function, shown in Figure 13.13(b). Again we have two fixed points, but the sign of $\beta(g)$ is now opposite, so that $g = g_0$ is not the infrared stable fixed point at $(t \to 0)$, while $g = 0$ is the ultraviolet stable fixed point at $(t \to \infty)$. In the last case $g \to 0$ for $t \to \infty$ and effective interaction diminishes with

the growth of energy (momentum) becoming zero in the limit. This is called *asymptotic freedom*. With the absence of zero of $\beta(g)$ at finite values of g now we have problems at small momenta; interaction grows and can produce unphysical divergence. In any case, here we obtain the transition to "strong coupling" at large distances (confinement?).

These possibilities in fact represent all the variants of asymptotic behavior in any reasonable model of quantum field theory. As a concrete example we can consider the $g\varphi^4$ ($g > 0$)-theory. Let us take the result of the single-loop approximation (13.52) for the renormalized coupling constant. Dropping irrelevant finite corrections we can write

$$g_1 = g\mu^\varepsilon \left(1 + \frac{3g}{16\pi^2\varepsilon}\right). \tag{13.101}$$

Then we have

$$\mu\frac{\partial g_1}{\partial \mu} = \varepsilon g\mu^\varepsilon + \frac{3g^2}{16\pi^2}\mu^\varepsilon . \tag{13.102}$$

For finite ε everything is finite, and we can (with the same accuracy) rewrite (13.102) as

$$\mu\frac{\partial g_1}{\partial \mu} = \varepsilon g_1 + \frac{3g_1^2}{16\pi^2}\mu^{-\varepsilon}, \tag{13.103}$$

and then just drop the index 1, assuming that we are working with a renormalized coupling constant. Then from (13.103), for $\varepsilon \to 0$, we obtain the Gell-Mann–Low function as

$$\beta(g) = \mu\frac{\partial g}{\partial \mu} = \frac{3g^2}{16\pi^2}. \tag{13.104}$$

Introducing $s = \ln t = \ln\frac{\mu}{\mu_0}$, so that $\mu\frac{\partial}{\partial\mu} = \frac{\partial}{\partial s} = t\frac{\partial}{\partial t}$, we can rewrite equation (13.103) as

$$\frac{\partial g}{\partial s} = \frac{3g^2}{16\pi^2}. \tag{13.105}$$

Now, without any calculations we can see that "running" coupling constant of φ^4-theory grows with the growth of s, i. e., with the growth of momentum, so that this theory is not asymptotically free. The Gell-Mann–Low function is $\sim g^2$. The elementary integration of equation (13.105) with initial condition $g(s = 0) = g_0$ gives

$$g = \frac{g_0}{1 - \frac{3}{16\pi^2}g_0 s} = \frac{g_0}{1 - \frac{3}{16\pi^2}g_0 \ln t} = \frac{g_0}{1 - \frac{3}{16\pi^2}g_0 \ln\frac{\mu}{\mu_0}}. \tag{13.106}$$

With the growth of t (or μ) the coupling constant grows, and finally we meet the unphysical singularity ("ghost pole") at $1 = \frac{3}{16\pi^2}g_0 \ln(\frac{\mu}{\mu_0})$, which corresponds to $\mu = \mu_0 \exp(\frac{16\pi^2}{3g_0})$. The situation here is quite similar to that in QED, which we discussed in Chapter 8. The same behavior was discussed in Chapter 10 in relation to critical phenomena.

Of course, this behavior of the Gell-Mann–Low function is completely based on a single-loop approximation and is formally valid only for small enough values of

the coupling constant g. The problem of the $\beta(g)$ behavior for large values of g and the related question of the consistency of $g\varphi^4$-theory remains open. Many researchers obtain for $g \to \infty$ the asymptotic behavior of $\beta(g)$, which is practically the same as the result of a single-loop approximation, which is equivalent to the internal inconsistency of the theory in accordance with the initial Landau claim, discussed above in Chapter 8. Alternative suggestions will be discussed in Chapter 14.

Note that $g\varphi^4$-theory is "easily" made asymptotically free if we assume $g < 0$. Then we obviously have to change the sign before the logarithm in the denominator of equation (13.106) and the effective coupling constant will drop with the growth of t and μ. However, such a theory is unstable: there is no ground state (potential energy can be arbitrarily negative), and this model is usually not considered in quantum field theory. However, the specific variant of such models, which is reduced to the generalized Landau functional (10.160), with the number of field components $n = 0$ (!), describes the motion of an electron in the random potential field of impurities with point-like potential V, chaotically distributed in space with fixed average density ρ, if in (10.160) we put $g = -\rho V^2$ and $\tau = -E$, where E is the electron energy. This problem is basic for the theory of electrons in disordered systems and related to the still unsolved problem of electron localization in such systems (Anderson localization – the basic mechanism of metal-insulator transitions). These problems are deeply connected with the description of the infrared region of asymptotically free models in quantum field theory. We shall not discuss these problems in more detail here, referring the readers to existing reviews [57, 64].

13.5 Asymptotic freedom of the Yang–Mills theory

Now let us consider the asymptotic properties of gauge theories. The situation in QED was discussed in Chapter 8, where it was shown that this theory is not asymptotically free, which leads to a "zero-charge" problem and pathological behavior at large momenta (energy). Remarkably, in non-Abelian theories the situation is different, and in these theories we can obtain asymptotic freedom. The discovery of this phenomenon by Gross and Wilczek has opened the way to construction of quantum chromodynamics and guaranteed the possibility of reliable calculations of QCD effects at high energies using perturbation theory.

Here we limit ourselves to the main results for the case of $SU(3)$ gauge theory (QCD) and a qualitative interpretation of asymptotic freedom, referring the reader for details to existing textbooks [13, 53, 56]. The key to finding asymptotic behavior is the Gell-Mann–Low $\beta(g)$-function. In QED in Chapter 8 we used the simplest single-loop approximation for vacuum polarization. In single-loop approximation of QCD we have additional contributions, related to the non-Abelian nature of the theory (self-interaction). We have to take into account the contribution to charge renormalization from the simple loop graphs for gluon–gluon interaction, shown in Figure 13.14, from gluon–ghost interaction, shown in Figure 13.15, and the QED-like contribution from

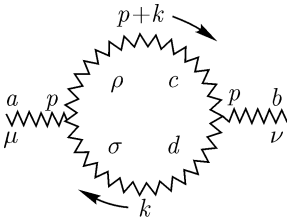


Figure 13.14

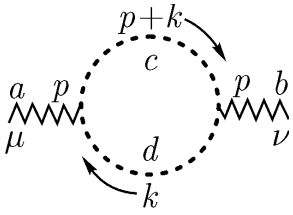


Figure 13.15

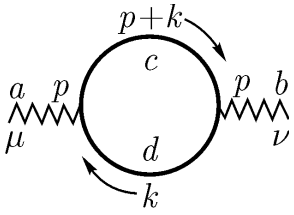


Figure 13.16

gluon–quark interaction, shown in Figure 13.16. After some tedious calculations [56] we obtain the renormalized QCD coupling constant in a single-loop approximation, similar to (13.101), in the form [8]

$$g_1 = g\mu^{\varepsilon/2}\left[1 + \frac{g^2}{4\pi\varepsilon}\left(-11 + \frac{2n_f}{3}\right)\right], \tag{13.107}$$

where n_f is the number of quark "flavors" (type of quarks). A similar single-loop correction for electrons in QED is $\frac{e^2}{4\pi}\left(-\frac{4}{3}\right)$. The sign of the fermion loop contribution in QCD is the same as in QED. However, the combined contribution from the diagrams of Figures 13.14 and 13.15 has the *opposite sign*! Correspondingly, for $n_f < 16$ the sign of the full polarization correction in (13.107) is *opposite* to that in QED ("anti-screening"). The physical reasons for such behavior will be explained below, while now, acting similarly to transformations leading from (13.101) to (13.104), in the limit of $\varepsilon \to 0$, we obtain

$$\beta(g) = \mu\frac{\partial g}{\partial \mu} = \frac{g^3}{12\pi}\left(-33 + 2n_f\right). \tag{13.108}$$

For $n_f \leq 16$ from (13.108) it follows that $\beta(g) < 0$ and the coupling constant g *diminishes* with the growth of the momentum (mass) scale, in accordance with the qualitative picture discussed above. We see that in this case the theory is asymptotically free. In nature we have $n_f = 6$.

Let us obtain the expression for the "running" coupling constant. Introducing once again $s = \ln t = \ln\frac{\mu}{\mu_0}$, $\mu\frac{\partial}{\partial\mu} = \frac{\partial}{\partial s}$, we obtain the Gell-Mann–Low equation

$$\frac{\partial g}{\partial s} = -\eta g^3, \quad \text{where} \quad \eta = \frac{33 - 2n_f}{12\pi}. \tag{13.109}$$

Let us rewrite it as

$$\frac{d}{ds}(g^{-2}) = 2\eta. \tag{13.110}$$

[8] Here we use the Gaussian system of units.

It is easy to see that the solution of this equation has the form

$$\frac{1}{g^2} = \frac{1}{g_0^2} + 2\eta s \tag{13.111}$$

or

$$g^2 = \frac{g_0^2}{1 + 2g_0^2\eta s} = \frac{g_0^2}{1 + 2g_0^2\eta \ln t}\,. \tag{13.112}$$

Introducing $t = Q/\mu$ and defining g_0 at $Q = \mu$, we obtain the result, already quoted in Chapter 8,

$$g^2(Q^2) = \frac{g^2(\mu)}{1 + \frac{g^2(\mu^2)}{12\pi}(33 - 2n_f)\ln\left(\frac{Q^2}{\mu^2}\right)}\,. \tag{13.113}$$

Only in the world with $n_f > 16$ the sign in the denominator of (13.113) will be the same as in QED. In real world the effective charge of QCD does not grow, but *drops* with the growth of Q^2 and becomes small at small distances. In contrast, for small enough Q^2 (at large distances between quarks) the effective coupling constant becomes large, which is reflected in *confinement* of quarks ("infrared jail"). For the value of Q^2, corresponding to the "ghost pole" in (13.113), we can introduce the notation Λ^2_{QCD}:

$$\Lambda^2_{QCD} = \mu^2 \exp\left[-\frac{12\pi}{(33 - 2n_f)g^2(\mu^2)}\right]\,. \tag{13.114}$$

so that (13.113) is rewritten as

$$g^2(Q^2) = \frac{12\pi}{(33 - 2n_f)\ln\left(\frac{Q^2}{\Lambda^2}\right)}\,. \tag{13.115}$$

For $Q^2 \gg \Lambda^2_{QCD}$ the effective coupling constant is small, and quark–gluon interaction (at small distances or large momenta) can be described by perturbation theory, similarly to electron-photon interactions in QED (at big distances of small momenta). For $Q^2 \sim \Lambda^2_{QCD}$ such a description becomes impossible, and quarks and gluons form strongly interacting clusters: hadrons. The experimental value of Λ_{QCD} is somewhere in the interval between 0.1 and 0.5 GeV. Then, for experiments being done at $Q^2 \sim (30\ \text{GeV})^2$ from (13.115) we obtain $g^2 \sim 0.1$, so that perturbation theory is applicable as in QED. In the limit of large Q^2 we can neglect all quark masses, but the theory still contains the mass scale μ^2, which appeared during the renormalization procedure.

Let us stress that the theoretical result (13.115) is well confirmed by experiments! In Figure 13.17 we show the experimental data for the effective coupling constant of QCD, as a function of the characteristic energy-momentum scale if with different scattering processes, studied at different experimental installations[9]. We see a rather convincing agreement between the theory and experiments.

[9] M. Schmelling. ArXiv: hep-ex/9701002.

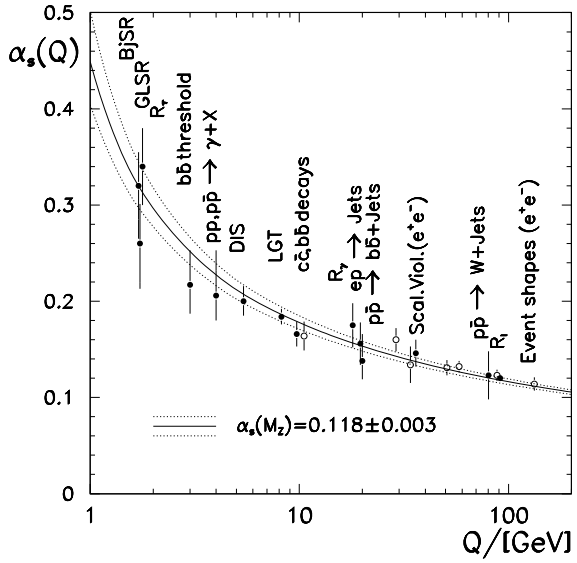


Figure 13.17

Antiscreening – the paramagnetism of Yang–Mills vacuum

We have seen that asymptotic freedom appears due to charge antiscreening in a Yang–Mills vacuum. This phenomenon has a rather simple explanation, based upon analogies with condensed matter theory[10].

Charge antiscreening means that a vacuum acts like dielectric a medium with dielectric permeability $\epsilon < 1$. The vacuum of quantum field theory differs from the usual polarizable medium in one important aspect: it is relativistically invariant. This means that its magnetic permeability μ is related to dielectric and both satisfy

$$\mu\epsilon = 1\,. \tag{13.116}$$

In fact, ϵ is the coefficient before the electric field term in action $\vec{E}\cdot\vec{D} \propto \epsilon F_{oi}F^{oi}$, while μ^{-1} is the coefficient before the magnetic field term $\vec{B}\cdot\vec{H} \propto \mu^{-1}F_{ij}F^{ij}$. The sum of these terms is relativistically invariant only if the condition $\epsilon = \mu^{-1}$ is satisfied. This relation allows us to connect electric properties of the medium with their magnetic properties, which may be of two types:

1. *Landau diamagnetism ($\mu < 1$).* Charged particles in the medium respond to the magnetic field creating the current, which itself induces the magnetic field with the direction opposite to the external field.

2. *Pauli paramagnetism ($\mu > 1$).* If particles possess magnetic moments, these are oriented along the field direction.

[10] Below we shall follow mainly F. Wilczek. Asymptotic Freedom. ArXiv: hep-th/9609099.

Then, the property of antiscreening of a Yang–Mills vacuum can be interpreted as $\mu > 1$, i. e., paramagnetism[11]. The thing is that non-Abelian gauge fields are Bose fields and, in contrast to Abelian photons, possess a gauge charge. Let us stress that the terminology of electromagnetism is used here, based on the analogy with $U(1)$ gauge theory (QED), while in reality we mean charges, corresponding to gauge $SU(3)$ (color) symmetry and color charges; electric and magnetic fields are understood as electric-like and magnetic-like components of a non-Abelian gauge field, corresponding to $SU(3)$ symmetry of QCD. When we are speaking about Yang–Mills fields in QCD (gluons) possessing charge and a magnetic moment, we mean that these fields possess a color charge and a color magnetic moment. Gluons are, of course. electrically neutral in the usual (electrodynamic) sense.

The well-known result of the theory of metals is that for an ideal gas of electrons Landau diamagnetism is overtaken by Pauli paramagnetism, so that the total response is paramagnetic [36]. We shall see that for non-Abelian gauge theories the situation is similar.

The standard classical Lagrangian of non-Abelian gauge theory has the form

$$\mathcal{L} = -\frac{1}{16\pi} G^a_{\alpha\beta} G^{a\alpha\beta} + \bar{\psi}(i\gamma^\nu D_\nu - m)\psi + \phi^\dagger(-D_\nu D^\nu - \mu^2)\phi + \text{ other contributions}, \tag{13.117}$$

where the field tensor is defined as $G^a_{\alpha\beta} \equiv \partial_\alpha A^a_\beta - \partial_\beta A^a_\alpha - g f^{abc} A^b_\alpha A^c_\beta$, and f^{abc} are the structural constants of the gauge group, covariant derivative $D_\nu = \partial_\nu + i g A^a_\nu \cdot T^a$, and T^a are the generators of the group (e. g., Pauli matrices $\frac{\tau}{2}$ for the fundamental representation of $SU(2)$, or Gell–Mann matrices $\frac{\lambda}{2}$ for fundamental representation of $SU(3)$). "Other contributions" are assumed to originate from Yukawa-type interactions and self-interactions of scalar fields. It is important that these contributions are independent of gauge fields. It is convenient to redefine $gA \to A$, so that the Yang–Mills constant g enters only the "free" part of the gauge field Lagrangian:

$$\mathcal{L} = -\frac{1}{16\pi g^2} G^a_{\alpha\beta} G^{a\alpha\beta} + \bar{\psi}(i\gamma^\nu D_\nu - m)\psi + \phi^\dagger(-D_\nu D^\nu - \mu^2)\phi + \text{ other contributions}, \tag{13.118}$$

where now $G^a_{\alpha\beta} \equiv \partial_\alpha A^a_\beta - \partial_\beta A^a_\alpha - f^{abc} A^b_\alpha A^c_\beta$, $D_\nu = \partial_\nu + i A^a_\nu \cdot T^a$ and g now enters only as a coefficient in the first term.

To calculate the magnetic susceptibility of a vacuum we need to know the change of its energy density due to a change of the external magnetic field. It may seem that everything is determined only by the first term in (13.118): $\frac{1}{8\pi g^2} B^2$. But this is only the classical contribution to energy; in quantum theory we need also to consider the charge of zero-point energy of all fields, entering (13.118), under the change of the external magnetic field. In fact, everything is similar to the theory of metals, where the vacuum corresponds to the filled Fermi sphere.

Before starting explicit calculations we shall write the correct answer and analyze its meaning and consequences. As we shall show, with the additional contribution of the zero-point

[11] The usual polarizable medium, in contrast, can simultaneously be screening ($\epsilon > 1$) and paramagnetic ($\mu > 1$). But still, there is some historical irony that the physical behavior leading to asymptotic freedom was, in fact, known to Landau, who made some fundamental contributions to the quantum theory of magnetism, but at the same time criticized the basics of quantum field theory, because of the pathological behavior of interactions at high momenta.

oscillations $\Delta\mathcal{E}$, the vacuum energy density in the external magnetic field B can be written as[12]

$$\mathcal{E} + \Delta\mathcal{E} = \frac{1}{8\pi g^2(\Lambda^2)} B^2 - \frac{1}{8\pi}\eta B^2 \ln\left(\frac{\Lambda^2}{B}\right) + \quad \text{finite contributions}\,, \tag{13.119}$$

where η was defined above in (13.109):

$$\eta = \frac{33 - 2n_f}{12\pi}\,, \tag{13.120}$$

and the neglected terms are finite in the limit of $g \to 0$ and $\Lambda \to \infty$. Here we introduced the usual cutoff Λ, i. e., dropped the contribution of all oscillations with wave vectors exceeding Λ. The origin of the notation $g^2(\Lambda^2)$ will soon become clear.

Consider the case when the cutoff Λ in (13.120) is changed to a smaller value Λ'. Then it is easy to see that all oscillation modes with wave vectors in the interval between Λ' and Λ give the following contribution to the change of vacuum energy:

$$\delta(\mathcal{E} + \Delta\mathcal{E}) = -\frac{1}{8\pi}\eta B^2 \ln\left(\frac{\Lambda^2}{\Lambda'^2}\right) = \left(\frac{1}{\mu} - 1\right)\frac{1}{8\pi g^2} B^2\,, \tag{13.121}$$

where in the second equality we introduced the contribution to vacuum magnetic susceptibility (permeability) from these modes, thus giving, in fact, its definition.

Now, for small g, we get

$$\mu - 1 = \eta g^2 \ln\left(\frac{\Lambda^2}{\Lambda'^2}\right)\,, \tag{13.122}$$

where we explicitly wrote the contribution to susceptibility from modes with energies (momenta) in the interval between Λ' and Λ. From equation (13.120) it is clear that here (as in the theory of metals) we have two contributions: the first is connected with the tendency of spins to orient along the field (paramagnetism), while the second is due to the orbital motion of charged particles (diamagnetism). For electron gas the paramagnetic response is three times greater, than the diamagnetic one [36]. The result (13.122) shows that in QCD the situation is similar and $\mu > 1$, which, as we have seen, corresponds to the antiscreening of the charge ($\epsilon < 1$). To determine the correct sign we have to take into account the fact that particles with spin 1 (gluons) have only two polarizations, and also that the fermion (quark) contribution to vacuum energy is negative (cf. Chapter 3), which leads to partial cancellation of the paramagnetic effect. In particular, in QED, where the Abelian electromagnetic field is not self-interacting, the entire effect is due to fermions, and we have the usual vacuum screening of the charge.

What are the consequences of equation (13.119) for physical observables? First of all we have to deal with the problem of arbitrary cutoff Λ. We define the effective coupling constant in such a way that the right-hand side of (13.119) becomes independent of Λ. To achieve this we require

$$\text{const} \equiv \frac{1}{g^2(\Lambda^2)} - \eta \ln\left(\frac{\Lambda^2}{B}\right), \tag{13.123}$$

which is equivalent to (13.111). It is better to write this condition in a differential form:

$$\frac{d}{d(\ln\Lambda^2)}\frac{1}{g^2(\Lambda^2)} = \eta\,, \tag{13.124}$$

[12] In our system of units $[B] = [L^{-2}] = [\Lambda^2]$, and we are using here the Gaussian system of units of electrodynamics.

which is the same as the Gell-Mann–Low equation (13.110). Now we see that the effective coupling constant drops with the growth of cutoff Λ, going to zero as the inverse logarithm of Λ for $\Lambda \to \infty$, when there are not too many quarks i. e., until $\eta > 0$. This is what we called asymptotic freedom.

Now let us proceed with the derivation of equation (13.119). A paramagnetic contribution to η from spin projections $\pm s$ is easily calculated as follows. Let the electric charge be 1, and gyromagnetic ratio g_m. As we are interested in the contribution of modes with very large momenta, the cutoff parameter Λ is much larger than the masses of all the particles, and we can consider all of them as massless (ignoring infrared divergences, which we may regularize, introducing a low momentum cutoff $\sim B$). Switching on the magnetic field leads to the energy shift of the relativistic particle [6]: $E^2 = k_1^2 + k_2^2 + k_3^2 \to E^2 \pm g_m B s$. Thus, the corresponding change of zero-point energy is

$$\Delta\mathcal{E} = \int \frac{d^3k}{(2\pi)^3}\frac{1}{2}(\sqrt{k^2 + g_m s B} + \sqrt{k^2 - g_m s B} - 2\sqrt{k^2})\,. \tag{13.125}$$

Expanding here up to terms quadratic in B and making angular integrations, we get

$$\Delta\mathcal{E} = -B^2(g_m s)^2\frac{1}{32\pi^2}\int_0^{\Lambda^2}\frac{dk^2}{k^2} = -B^2(g_m s)^2\frac{1}{32\pi^2}\ln\frac{\Lambda^2}{B}\,. \tag{13.126}$$

This gives the paramagnetic contribution to (13.119). The precise value of the numerical coefficient in (13.119) is related to group constants of $SU(3)$ and we shall not derive it here.

Calculation of diamagnetic contribution to η is more difficult°. Let us take the vector potential of a magnetic field in the Landau gauge: $A_y = Bx$. the Klein–Gordon equation for orbital motion of a relativistic particle in a magnetic field is

$$\left[E^2 + \frac{\partial^2}{\partial x^2} + \left(\frac{\partial}{\partial y} - iBx\right)^2 + \frac{\partial^2}{\partial z^2}\right]\phi = 0\,, \tag{13.127}$$

and its solutions are written as

$$\phi = e^{i(k_2 y + k_3 z)}\chi_n\left(x - \frac{k_2}{B}\right) \tag{13.128}$$

with corresponding eigenvalues $E_n^2 = k_3^2 + B(2n+1)$. Here χ_n is the usual oscillator wave function [35]. Energy levels are characterized by the integer n and momentum k_3, but are degenerate over k_2, as for the usual Landau levels in a magnetic field [35]. If we consider the states in the cube with side L, the coordinate of the center of the oscillator k_2/B should satisfy the inequality $0 \le k_2/B \le L$, which means that in the interval Δk_3 we have $\Delta k_2 \Delta k_3/(2\pi)^2 = \frac{B}{4\pi^2}\Delta k_3$ states with fixed n (for unit volume $L^3 = 1$). Then the corresponding contribution to the energy of zero-point modes is given by

$$\mathcal{E}_0 = \frac{B}{(2\pi)^2}\sum_{n=0}^{\frac{\Lambda^2}{2B}-\frac{1}{2}}\int_{-\infty}^{\infty} dk_3\theta\left[\Lambda^2 - k_3^2 - B(2n+1)\right]\sqrt{k_3^2 + B(2n+1)} \equiv \sum_{n=0}^{\frac{\Lambda^2}{2B}-\frac{1}{2}} f\left(n+\frac{1}{2}\right). \tag{13.129}$$

This is a rather complicated expression because of the sum over n. For us it is sufficient to take into account the first nontrivial contribution using the Euler–Maclaurin summation formula:

$$\sum_{n=0}^{p} g\left(n+\frac{1}{2}\right) = \int_0^{p+1} dn\, g(n) - \frac{1}{24}(g'(p+1) - g'(0)) + \ldots \tag{13.130}$$

as the next terms lead to contributions of higher orders in B/Λ^2. Applying (13.130) to (13.129), we see that the integral term is independent of B, while the significant contribution comes from the derivative at zero:

$$\frac{1}{24} f'(0) = \frac{1}{24} \frac{B}{4\pi^2} 2 \int_{\sqrt{B}}^{\Lambda} dk_3 \frac{B}{\sqrt{k_3^2}} = B^2 \frac{1}{96\pi^2} \ln \frac{\Lambda^2}{B} . \tag{13.131}$$

This gives the diamagnetic part of (13.119), which is smaller than paramagnetic term (13.126) for any reasonable values of g_m and s.

As we noted many times, the discovery of asymptotic freedom in non-Abelian gauge theories played a revolutionary role in modern quantum field theory, transforming QCD into a "respectable" theory and the foundation of the Standard Model. During the last thirty-five or so years this theory was tested in many experiments and was always confirmed. We shall not discuss this. Many aspects of QCD are discussed in [13]. A raher detailed presentation of the mathematical apparatus of QCD can be found in [62]. Among the unsolved problems we have to mention the problem of confinements, which is deeply related to the problem description of the strong coupling (nonperturbative) effects of QCD in the infrared region (large distances). We shall briefly discuss these problems in the next chapter.

In recent years there has been an intensive development of the theory of quark-gluon matter under extreme conditions of high temperatures and densities, important for problems of astrophysics and cosmology as well as for the study of heavy nuclei collisions in accelerator experiments. Here we meet some remarkable analogies with the physics of condensed matter. In particular, great attention is devoted to the study of the so-called *color superconductivity*, appearing in quark-gluon matter due to Cooper pairing of quarks, induced by the attraction due to gluon exchange. A detailed and clear presentation of these problems can be found in [54, 72][13].

13.6 "Running" coupling constants and the "grand unification"

In Chapter 12 we considered the $SU(2) \otimes U(1)$ symmetric unified theory of electroweak interactions, which is in remarkable agreement with experiments, similarly to the $SU(3)$ invariant QCD. But is it really a unified theory? In fact, $SU(2) \otimes U(1)$ represents the direct product of two disconnected groups of gauge transformations: the $SU(2)$ group of weak isospin with coupling constant g and the $U(1)$ group of weak hypercharge with coupling constant f. The ratio of these two coupling constants, introduced in equation (12.83) as

$$tg\alpha = \frac{f}{g} \tag{13.132}$$

[13] An elementary presentation of the successes of modern QCD is given in a mini-review [74].

is to be determined from experiments. However, if we consider groups $SU(2)$ and $U(1)$ as subgroups of some larger gauge group

$$G \supset SU(2) \otimes U(1)\,, \tag{13.133}$$

the constants g and f can be related to each other by group relations, which will determine the Weinberg angle α. Some of the transformations of the wider group G will connect previously disconnected subsets of groups $SU(2)$ and $U(1)$. It is natural to try to unify electroweak symmetries $SU(2)$ and $U(1)$ with the color gauge $SU(3)$ symmetry of QCD:

$$G \supset SU(3) \otimes SU(2) \otimes U(1)\,. \tag{13.134}$$

Then the gauge transformations of the group G will connect electroweak constants g and f with QCD coupling. As a result *all* known interactions will be described by a single gauge group with the single coupling constant g_G, while all observable constants of known interactions will be unambiguously defined by the group structure of G. This type of model is usually called *grand unified theories* (GUT). There are a number of such models under discussion for possible verification, and below we shall briefly discuss some aspects of this approach.

The foundation for such a description can be guessed from the real behavior of "running" coupling constants for known interactions. We shall denote these constants as $g_1(Q)$, $g_2(Q)$, and $g_3(Q)$, corresponding to gauge groups $U(1)$, $SU(2)$, and $SU(3)$. Let us introduce the following standard notations relating g_i $(i = 1, 2, 3)$ with the coupling constants used above:

$$\begin{aligned} SU(3):&\quad g^2(Q) = 4\pi g_3^2(Q)\,,\\ SU(2):&\quad g(Q) = g_2(Q)\,,\\ U(1):&\quad f(Q) = \tfrac{1}{C} g_1(Q)\,. \end{aligned} \tag{13.135}$$

Here we also introduced (not very important for us in the future) the coefficient C, which is usually defined by some group constants of G. In particular, the angle α from (13.132) becomes the function of Q:

$$tg\alpha(Q) = \frac{1}{C}\frac{g_1(Q)}{g_2(Q)}\,. \tag{13.136}$$

Figure 13.18 shows the behavior of "running" coupling constants of the Standard Model $\alpha_i = \frac{g_i^2}{4\pi}$ as functions of $\log_{10}(\mu/\mathrm{GeV})$, obtained from scattering experiments and (at very large momenta) from single-loop expressions, which were discussed above in some details for QED and QCD. We see that the QCD constant g_3 drops with the growth of momentum (asymptotic freedom), while the constants of electroweak theory g_1 and g_2 grow. However, we clearly observe the tendency for effective constants to become more or less equal in the region of $Q \sim 10^{15}$ GeV. It can be expected, that in the true theory of elementary particles at some large value of $Q \sim M_X$ (at small distances!) all three constants become just one universal constant

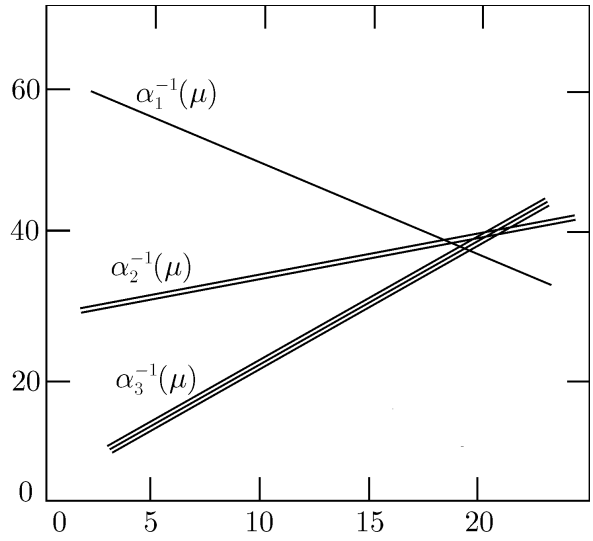


Figure 13.18

of "grand unification"[14]:

$$g_i(Q) = g_G(Q) \quad \text{for} \quad Q \geq M_X\,, \tag{13.137}$$

corresponding to gauge group G. For $Q < M_X$ constants $g_i(Q)$ separate and at large distances tend to the phenomenological constants g_i, describing the observable interactions, roughly corresponding to $Q \sim \mu \sim 10$ GeV. Such behavior of coupling constants is also obtained in some supersymmetric generalizations of the Standard Model[15]. An example of the "running" couplings behavior obtained in such models is shown in Figure 13.19. Such behavior of effective coupling constants is considered a strong argument for theories with supersymmetry. However, it should be noted that supersymmetry is, in any case, strongly broken in real nature. Also, up to now there is no experimental evidence for its existence. In particular, it is not known whether or not any "superpartners" of the known elementary particles exist.

Assuming the existence of the GUT group G and using the phenomenological values of coupling constants determined at $Q \sim \mu \sim m_W$, we can make a more accurate estimate of M_X. For the QCD constant, using (13.109)–(13.113), we can write

$$\frac{1}{g_3^2(\mu)} = \frac{1}{g_3^2(Q)} + 2b_3 \ln \frac{Q}{\mu}\,, \tag{13.138}$$

[14] In this region the Weinberg angle, in accordance with (13.136), is determined by group coefficient C.
[15] We recall that supersymmetry transforms fermions into bosons and back.

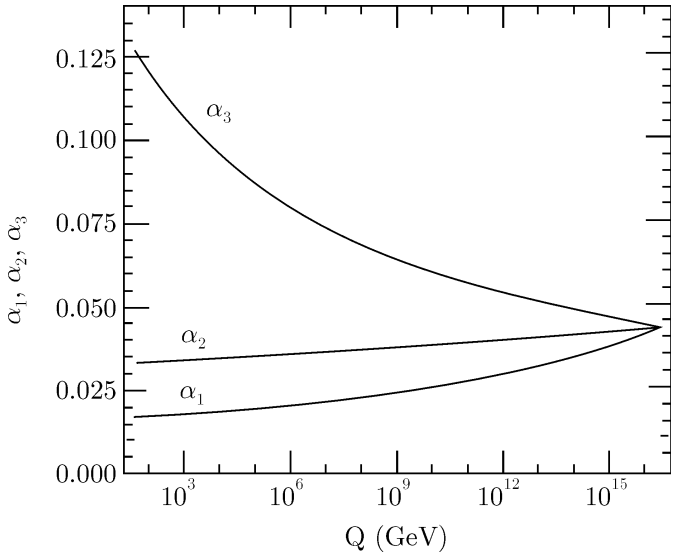


Figure 13.19

where we have introduced

$$b_3 = \frac{1}{(4\pi)^2}\left(\frac{2}{3}n_f - 11\right), \tag{13.139}$$

which differs from η introduced above, by its sign and constant factor. For $Q = M_X$ we have $g_3 = g_G$, so that from equation (13.138) we get

$$\frac{1}{g_i^2(\mu)} = \frac{1}{g_G^2} + 2b_i \ln\frac{M_X}{\mu}, \quad \text{where } i = 3. \tag{13.140}$$

The same relation can be applied to coupling constants g_1 and g_2 of gauge groups $SU(2)$ and $U(1)$, with

$$b_1 = \frac{1}{(4\pi)^2}\left(\frac{4}{3}n_g\right),$$

$$b_2 = \frac{1}{(4\pi)^2}\left(-\frac{22}{3}\right) + b_1, \tag{13.141}$$

$$b_3 = \frac{1}{(4\pi)^2}(-11) + b_1, \tag{13.142}$$

where n_g is the number of fermion flavors for the given model. In the general case of $SU(N)$ gauge group, we have

$$b_N = \frac{1}{(4\pi)^2}\left(-\frac{11}{4}N + \frac{4}{3}n_g\right), \tag{13.143}$$

where the first term is connected with loop contribution of gauge bosons, while the second one with fermion loops.

Excluding n_g and g_G from three equations like (13.142) and using (13.143) we can compose the following linear combination:

$$\frac{C^2}{g_1^2} + \frac{1}{g_2^2} - \frac{1+C^2}{g_3^2} = 2[C^2 b_1 + b_2 - (1+C^2)b_3] \ln \frac{M_X}{\mu}, \tag{13.144}$$

where $g_i^2 = g_i^2(\mu)$. The left hand side here is chosen in such a way that it can be expressed via e^2 and g_3^2. In fact, we have

$$\frac{C^2}{g_1^2} + \frac{1}{g_2^2} = \frac{1}{f^2} + \frac{1}{g^2} = \frac{1}{e^2}, \tag{13.145}$$

where we have used (13.135) and electroweak theory relation $e = g \sin\alpha = f \cos\alpha$. Substituting the coefficients b_i from (13.142) into (13.144), we obtain

$$\ln \frac{M_X}{\mu} = \frac{3(4\pi)^2}{22(1+3C^2)} \left[\frac{1}{e^2} - \frac{1+C^2}{g_3^2} \right]. \tag{13.146}$$

For $\mu \sim 10\,\text{GeV}$ we have $e^2 \sim 10^{-2}$ and $g_3^2 \sim 0.1$. Assuming[16] $C^2 = 5/3$ we have

$$M_X \sim 5 \cdot 10^{14}\,\text{GeV}. \tag{13.147}$$

This estimate is not very sensitive to the choice of μ and the precise value of C. Actually, the mass M_X is very large, but we can still neglect the gravitation effects[17].

A minimal group satisfying the condition of

$$G \supset SU(3) \otimes SU(2) \otimes U(1) \tag{13.148}$$

is $SU(5)$, leading to the simplest GUT model (Georgi–Glashow). What kinds of gauge bosons appear in this theory? In the general case of a $SU(N)$-symmetric gauge group we have $N^2 - 1$ gauge bosons. Then for $SU(5)$ we have

$$24 = (8,1)_{\text{Gluons}} + [(1,3) + (1,1)]_{W,Z,\gamma} + [(3,2) + (\bar{3},2)]_{X,Y}. \tag{13.149}$$

Thus, in this model superheavy bosons X and Y appeark. They have color and are

[16] This follows from (13.136) and $\sin^2\alpha \approx 0.2$. In the general case, from (13.136) we have $\sin^2\alpha = \frac{g_1^2(Q)}{g_1^2(Q)+C^2 g_2^2(Q)}$. If we take $C^2 = 5/3$, then for $Q = M_X$, i. e., for $g_1 = g_2$, we get $\sin^2\alpha = 3/8$. However, for $Q \approx \mu$ the value of $\sin^2\alpha$ is different because of $g_1 \neq g_2$..

[17] The account of gravitation becomes important for $\frac{GM^2}{r}|_{r=\frac{\hbar}{Mc}} \sim Mc^2$, which gives the Planck mass $M_P c^2 \sim (\frac{\hbar c^5}{G})^{1/2} \sim 1.2 \cdot 10^{19}\,\text{GeV}$.

intermediate bosons for interactions, transforming quarks into leptons:

$$(u,d)_L \to e_L^+ + (\bar{Y},\bar{X})\,, \tag{13.150}$$

which inevitably leads to proton decay[18]

Fermions in the $SU(5)$ model belong to fundamental representations $\bar{5}$ and 10. Explicitly, for left-handed states we have

$$\begin{aligned} \bar{5} &= (1,2) + (\bar{3},1) = (\nu_e, e^-)_L + \bar{d}_L\,, \\ 10 &= (1,1) + (\bar{3},1) + (3,2) = e_L^+ + u_L^+ + (u,d)_L\,. \end{aligned} \tag{13.151}$$

Theoretical estimates for the lifetime of proton give

$$\tau_p \sim \frac{M_X^4}{m_p^5}\,. \tag{13.152}$$

It is seen that its numerical value is not very sensitive to the precise value of M_X and is within the interval of 10^{30}–10^{32} years! The present day experimental limit is $\tau_p > 10^{32}$ years. This contradicts the simplest $SU(5)$ GUT model. However, in more complicated GUT models, the proton lifetime can be made much larger. Unfortunately, at present there is no clear experimental way to search for proton decay with a lifetime exceeding 10^{32} years. In this sense, and also because of the immense scale of the M_X masses, all GUT models represent a kind of theoretical "game". However, purely theoretical considerations stimulate further work in this direction [73].

[18] Proton decay is not so unexpected as it may seem. Conservation of electric charge is related to the existence of a massless photon, but apparently there are no particles, responsible for conservation of baryon charge (cf. Chapter 2). For $Q \sim M_{X,Y}$ the strong color interaction is mixed with electroweak interaction, and a clear distinction between color quarks and colorless leptons just vanishes.

Chapter 14

Nonperturbative approaches

14.1 The lattice field theory

Our previous presentation of the theory of interacting quantum fields was based on perturbation theory. In fact, it is the only universal method to deal with interactions. However, it is clear that there is are number of problems in quantum field theory which cannot be solved without the use of methods outside perturbation theory. In particular, we are meeting such problems in studies of the asymptotic properties of quantum field theory, where we have to use nonperturbative approaches in our attempts to find the correct behavior of the Gell-Mann–Low function. Among the physical problems of interest here we mention first of all the problem of quark confinement. It is obvious that there is no universal way to move outside the framework of perturbation theory. At the same time, a number of specific approaches were developed in the literature allowing us to analyze certain nonperturbative effects. This has led to some general concepts, which are currently important not only in quantum field theory, but also in other fields, such as condensed matter theory. In this chapter we shall discuss a number of such problems, concentrating mainly on these conceptual aspects.

An important part of modern quantum field theory is lattice gauge theory. It was proposed by Wilson, and so far is the only method allowing a more or less complete solution of the confinement problem. In this approach, instead of the usual space-time continuum, we introduce discrete space-time[1]. Now we do not have any problem with ultraviolet divergences, as we have a natural cutof: wavelengths in a discrete lattice cannot be smaller than the double lattice constant a, while the momentum projection can change from zero up to $\frac{\pi}{a}$ (i. e., within the first Brillouin zone of solid state theory). In this formulation, quantum field theory becomes similar to the statistical mechanics of lattice models, where we have well-developed methods which allow (sometimes) us to solve problems outside the limits of perturbation theory. In particular, in lattice models we can effectively use numerical approaches, such as Monte-Carlo simulations. Below we follow mainly [13]; a more detailed presentation of the lattice models in quantum field theory can be found in [30, 31].

Here we shall deal only with the Euclidean formulation of lattice quantum field theory, though there are methods allowing the explicit treatment of time dependence. We shall consider only simple the cubic lattice with the lattice constant a in four-

[1] The introduction of the lattice obviously breaks the relativistic invariance of the theory, but it is not very important for problems under discussion; our main interest will be QCD behavior at large distances, where we can just forget about the discrete lattice.

dimensional space. The lattice sites will be parametrized by 4-vector n. Then, four-dimensional integration is replaced by summation:

$$\int d^4x \cdots \to a^4 \sum_n \cdots . \tag{14.1}$$

Scalar fields

Consider the simplest case of scalar field $\phi(x)$. The action in continuous Euclidean space has the form

$$S(\phi) = \int d^4x \left[\frac{1}{2}(\partial_\mu \phi)^2 + V(\phi) \right] , \tag{14.2}$$

where

$$V(\phi) = \frac{1}{2} m^2 \phi^2 + \frac{\lambda}{4} \phi^4 . \tag{14.3}$$

To go to the lattice representation we note that scalar filed is now defined at every lattice site n:

$$\phi(x) = \phi_n . \tag{14.4}$$

The erivative of the field on the lattice is defined as

$$\partial_\mu \phi(x) \to \frac{1}{a}(\phi_{n+\hat{\mu}} - \phi_n) , \tag{14.5}$$

where $\hat{\mu}$ is the 4-vector with length a in direction μ.

Then for the lattice action we have

$$S(\phi) = \sum_n \left\{ \frac{a^2}{2} \sum_{\mu=1}^{4} (\phi_{n+\hat{\mu}} - \phi_n)^2 + a^4 \left(\frac{m^2}{2} \phi_n^2 + \frac{\lambda}{4} \phi_n^4 \right) \right\} . \tag{14.6}$$

It is useful to transform to momentum representation and define the excitation spectrum of free theory ($\lambda = 0$). Let us use Fourier transformation:

$$\phi_n = \int \frac{d^4k}{(2\pi)^4} e^{ik\cdot n} \phi(k) . \tag{14.7}$$

Integration in (14.7) is performed over the Brillouin zone of the inverse lattice, i. e.,

$$-\frac{\pi}{a} \le k_\mu \le \frac{\pi}{a} \quad \text{for every } \mu = 1, \ldots, 4 . \tag{14.8}$$

Here $k_\mu \equiv k \cdot \hat{\mu}$. After the substitution of (14.7) into (14.6) we can write terms, originating from "kinetic" energy, as

$$a^4 \sum_n \int \frac{d^4k}{(2\pi)^4} \int \frac{d^4k'}{(2\pi)^4} e^{i(k+k')\cdot n} (e^{iak_\mu} - 1)(e^{iak'_\mu} - 1) =$$

$$\int \frac{d^4k}{(2\pi)^4} (e^{iak_\mu} - 1)(e^{-iak_\mu} - 1) = 4 \int \frac{d^4k}{(2\pi)^4} \sin^2 \left(\frac{ak_\mu}{2} \right) , \tag{14.9}$$

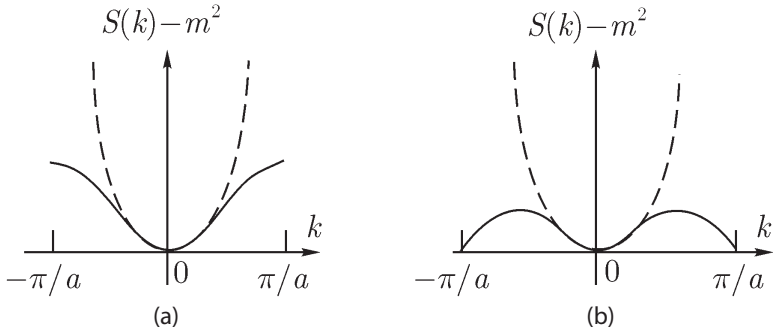


Figure 14.1

so that free action takes the form

$$S_0(\phi) = \frac{1}{2}\int \frac{d^4k}{(2\pi)^4}\left[\sum_\mu \frac{4}{a^2}\sin^2\left(\frac{ak_\mu}{2}\right) + m^2\right]\phi(-k)\phi(k)\,, \tag{14.10}$$

Thus, each mode gives the following contribution to action:

$$S(k) = \frac{1}{2}\left[\sum_\mu \frac{4}{a^2}\sin^2\left(\frac{ak_\mu}{2}\right) + m^2\right] \tag{14.11}$$

instead of standard $\frac{1}{2}\left(k^2+m^2\right)$. However, both expressions have the same continuous limit (the limit of small k), so that everything is consistent. The spectrum obtained is shown in Figure 14.1(a).

The theory with lattice action (14.6) can be quantized using functional integral formalism, when the vacuum average is defined as[2]

$$\langle 0|\phi_{n1}\phi_{n2}\cdots\phi_{nl}|0\rangle = \frac{1}{Z}\int \prod_n [d\phi_n](\phi_{n1}\phi_{n2}\cdots\phi_{nl})e^{-S[\phi]}\,, \tag{14.12}$$

where

$$Z = \int \prod_n [d\phi_n]e^{-S[\phi]}\,. \tag{14.13}$$

This is a typical statistical mechanics of the field (order parameter) ϕ_n on a lattice! The value of $S[\phi]$ corresponds to fluctuation-free energy. Equation (14.12) represents the correlation function of this order parameter at different lattice sites. It is useful to compare these expressions with equations (10.160), (10.162), and (10.164), used above in the theory of critical phenomena.

Let us change the variable (change the field scale):

$$\phi_n' = \sqrt{\lambda}\phi_n\,. \tag{14.14}$$

[2] In Euclidean theory there is no sense in introducing T-ordering!

Then the lattice action takes the form

$$S(\phi) = \frac{1}{\lambda} S'(\phi'), \tag{14.15}$$

where

$$S'(\phi') = \sum_n \left\{ \frac{a^2}{2} \sum_\mu (\phi'_{n+\mu} - \phi'_n)^2 + a^4 \left(\frac{m^2}{2} \phi'^2_n + \frac{1}{4} \phi'^4_n \right) \right\}, \tag{14.16}$$

so that the coupling constant λ becomes the common factor for the whole action. Then (14.12) and (14.13) are rewritten as

$$\langle 0 | \phi'_{n1} \phi'_{n2} \cdots \phi'_{nl} | 0 \rangle = \frac{1}{Z'} \int \prod_n [d\phi'_n] (\phi'_{n1} \phi'_{n2} \cdots \phi'_{nl}) \exp \left\{ -\frac{1}{\lambda} S'[\phi'] \right\}, \tag{14.17}$$

$$Z' = \int \prod_n [d\phi'_n] \exp \left\{ -\frac{S'[\phi']}{\lambda} \right\}. \tag{14.18}$$

If we change here

$$\frac{1}{\lambda} \to \beta = \frac{1}{T}, \tag{14.19}$$

where T is the temperature, the *strong coupling* expansion of quantum field theory, which is to be done over the *inverse* powers of coupling constant λ, becomes equivalent to the *high temperature* expansion of statistical mechanics. This opens wide prospects for studying such expansions, as the high-temperature expansions are widely used in lattice models of statistical mechanics (e. g., to study critical phenomena) and are fairly well developed [21, 63].

Fermion fields

Let us consider fermions. The same procedure as we used for scalar fields leads to Euclidean action of free fermions in the form

$$S_0(\psi) = \sum_n \left\{ \frac{a^3}{2} \sum_{\mu=1}^{4} \bar{\psi}_n \gamma_\mu (\psi_{n+\hat{\mu}} - \psi_{n-\hat{\mu}}) + m a^4 \bar{\psi}_n \psi_n \right\}, \tag{14.20}$$

where the γ-matrices of Euclidean theory satisfy anticommutation relations:

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}. \tag{14.21}$$

In momentum representation the action (14.20) is written as

$$S_0(\psi) = \int \frac{d^4k}{(2\pi)^4} \bar{\psi}(-k) \left\{ i \sum_\mu \gamma_\mu \frac{\sin a k_\mu}{a} + m \right\} \psi(k). \tag{14.22}$$

As compared with the continuous case we have the replacement $\gamma_\mu k_\mu \to \gamma_\mu \frac{1}{a} \sin a k_\mu$. In the same way as in the usual (Euclidean) Dirac's theory, the operator $\gamma_\mu k_\mu + m$

produces the spectrum $k^2 + m^2$; here we obtain the excitation spectrum

$$S(k) = \frac{\sin^2 ak_\mu}{a^2} + m^2 , \tag{14.23}$$

shown in Figure 14.1(b). We see that now we have two equivalent minima of the spectrum in the Brillouin zone. One is at $k = 0$ and leads to the correct continuous limit. The other mode, corresponding to the minimum at $k_\mu = \pm\frac{\pi}{a}$, corresponds (in the limit of $a \to 0$) to infinite momentum, but can be excited in the case of finite a. Correspondingly, we have to modify the theory in such a way that we exclude the contribution of the second minimum without changing the continuous limit. To achieve this, Wilson proposed adding to the lattice Lagrangian the term

$$\Delta\mathcal{L} = \frac{1}{2a}\bar{\psi}_n(\psi_{n+\hat{\mu}} + \psi_{n-\hat{\mu}} - 2\psi_n) , \tag{14.24}$$

so that in Euclidean space the action of free fermions takes the form

$$S_0(\psi) = \sum_n \left\{ \frac{a^3}{2} \sum_\mu \bar{\psi}_n[(1+\gamma_\mu)\psi_{n+\hat{\mu}} + (1-\gamma_\mu)\psi_{n-\hat{\mu}} - 2\psi_n] + ma^4\bar{\psi}_n\psi_n \right\}. \tag{14.25}$$

In momentum representation we have

$$S_0(\psi) = \int \frac{d^4k}{(2\pi)^4}\bar{\psi}(-k)\left\{ i\sum_\mu \gamma_\mu \frac{\sin ak_\mu}{a} + m - \sum_\mu \frac{\cos ak_\mu - 1}{a} \right\}\psi(k). \tag{14.26}$$

This action leads to the shift of the second minimum to finite energies, while the behavior at small k does not change. Then, in the continuous limit we remain only with the contribution from the "correct" minimum at $k = 0$.

Local gauge invariance

Let us now construct the lattice gauge theory. For concreteness we shall deal with $SU(3)$-symmetric QCD. Local (depending on the site!) gauge transformation is written as

$$\psi_n \to \Phi_n\psi_n , \quad \bar{\psi}_n \to \bar{\psi}_n\Phi_n^+, \tag{14.27}$$

where

$$\Phi_n = \exp\left\{ i\frac{\lambda^i}{2}\theta_n^i \right\} . \tag{14.28}$$

Here λ^i $(i = 1, 2, \ldots, 8)$ are Gell-Mann matrices (generators of $SU(3)$), cf. equations (2.186)).

Now we introduce the so-called *link variable*, defined on the lattice link, connecting the nearest neighbors sites:

$$U(n+\hat{\mu}, n) = \exp\left\{ iga\frac{\lambda^i}{2}A^i_{n\mu} \right\} , \tag{14.29}$$

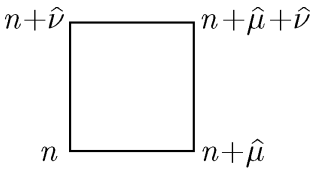


Figure 14.2

where $A^i_{n\mu}$ is the lattice field of the gluons, g is the Yang–Mills coupling constant. Gauge transformation for this matrix is defined as

$$U(n+\hat{\mu},n) \to \Phi_{n+\hat{\mu}} U(n+\hat{\mu},n)\Phi^+_n\,. \tag{14.30}$$

From (14.27) and (14.30) it follows that the combination $\bar{\psi}_n U(n,n+\hat{\mu})\psi_{n+\hat{\mu}}$ is the gauge invariant. Then in becomes clear how we should modify the action (14.25) to obtain the quark part of $SU(3)$-symmetric action of QCD:

$$S_{QCD} = S(q) + S(A)\,, \tag{14.31}$$

$$S(q) = \sum_n \left\{ \frac{a^3}{2} \sum_\mu \left[\bar{\psi}_n(1+\gamma_\mu)U(n,n+\hat{\mu})\psi_{n+\hat{\mu}} \right.\right. \\ \left.\left. +\bar{\psi}_n(1-\gamma_\mu)U(n,n-\hat{\mu})\psi_{n-\hat{\mu}} - 2\psi_n \right] + ma^4\bar{\psi}_n\psi_n \right\}. \tag{14.32}$$

In the continuous limit of $a \to 0$ the expansion of (14.32) in powers of a gives the usual expression for fermion action with the covariant derivatives of gauge theory.

How should we write the action for gauge (gluon) field itself? It is clear that it should be built of link variables. The simplest gauge invariant combination is defined on the elementary square of *plaquette* of the lattice, shown in Figure 14.2. Let us compose the matrix product of link variables, taken along the links of the plaquette p:

$$U_p = U(n,n+\hat{\mu})U(n+\hat{\mu},n+\hat{\mu}+\hat{\nu})U(n+\hat{\mu}+\hat{\nu},n+\hat{\nu})U(n+\hat{\nu},n). \tag{14.33}$$

This combination is obviously invariant with respect to transformations (14.30). Let us define the action of the gauge field as the following sum over all plaquettes on the lattice:

$$S(A) = -\frac{1}{8\pi g^2}\sum_p SpU_p\,. \tag{14.34}$$

Here the trace of the product is taken over the $SU(3)$ matrix indices. If we expand the exponents in (14.33), (14.34) in powers of a and neglect the terms of the order of $O(a^3)$, equation (14.34) can be rewritten as

$$S(A) = -\frac{1}{8\pi g^2}\sum_p Sp\{\exp(iga^2F_{n\mu\nu})\}\,, \tag{14.35}$$

where

$$F_{n\mu\nu} = \partial_\mu A_{n\nu} - \partial_\nu A_{n\mu} - ig[A_{n\mu}, A_{n\nu}], \tag{14.36}$$

where we have introduced the notation

$$\partial_\mu A_{n\nu} \equiv \frac{1}{a}(A_{n+\hat{\mu}\nu} - A_{n\nu}), \tag{14.37}$$

where $A_{n\mu} = A^i_\mu \lambda^i/2$ is the gluon field at lattice site n. This expression immediately gives the correct continuous limit:

$$S(A) = -\frac{1}{8\pi g^2}\sum_p \left\{1 - \frac{g^2a^4}{2}F^i_{\mu\nu}F^{i\mu\nu} + \cdots\right\} \to \text{const} + \frac{1}{16\pi}\int d^4x F^i_{\mu\nu}F^{i\mu\nu}, \tag{14.38}$$

where during the derivation we used $Sp\lambda^i = 0$ and $Sp(\lambda^i\lambda^j) = 2\delta^{ij}$.

The criterion for confinement. The Wilson loop

To define the confinement criterion for quarks in QCD, we shall find the energy of the system consisting of a quark at point $x = (t, \mathbf{0})$ and an antiquark at $x = (t, \mathbf{R})$. In the absence of confinement we obviously have

$$E(R) \to 2m \quad \text{for} \quad R \to \infty, \tag{14.39}$$

where m is the quark mass. Confinement corresponds to the infinitely growing (with interquark distance) interaction potential:

$$E(R) \to \infty \quad \text{for} \quad R \to \infty. \tag{14.40}$$

We shall denote the fermion quark field as $q(x)$ and introduce the gauge invariant $q\bar{q}$-operator:

$$\Gamma[x, x'; C] = \bar{q}(x')U(x', x; C)q(x), \tag{14.41}$$

where $U(x', x; C)$ is the ordered product of link variables along some path (trajectory) C, connecting points x and x' on the lattice[3]. Consider the gauge invariant correlator, describing the overlap of $q\bar{q}$ at the moment of (Euclidean!) time $t = 0$ and $q\bar{q}$ at the time moment $t = T$:

$$\Omega(T, R) = \langle 0|\Gamma^+[(0, \mathbf{0}), (0, \mathbf{R}); C]\Gamma[(T, \mathbf{0}), (T, \mathbf{R}); C]|0\rangle. \tag{14.42}$$

Inserting between operators the unity representation (completeness condition!) via the sum over the complete system of the energy eigenstates of our system, we obtain (compare the similar treatment in Chapter 9)

$$\Omega(T, R) = \sum_n |\langle 0|\Gamma^+[(0, \mathbf{0}), (0, \mathbf{R}); C]|n\rangle|^2 e^{-E_nT}. \tag{14.43}$$

[3] In a continuous limit: $U(x', x) = P\exp\{ig\int_x^{x'} dy^\mu \frac{\lambda^i}{2}A^i_\mu(y)\}$, where P is the ordering operator along path C.

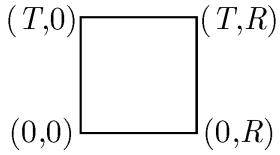


Figure 14.3

We see that for large T the main contribution here comes from the state with the smallest E_n. This minimal eigenvalue of energy obviously corresponds to the *potential* energy of the static $q\bar{q}$ system, with a quark and antiquark placed at distance R from each other:

$$\lim_{T\to\infty} \Omega(T,R) \sim e^{-E(R)T}. \tag{14.44}$$

In terms of quark fields we have

$$\Omega(T,R) = \langle 0|\bar{q}(0,\mathbf{R})U[(0,\mathbf{R}),(0,\mathbf{0});C]q(0,\mathbf{0})\bar{q}(T,\mathbf{0})U[(T,\mathbf{0}),(T,\mathbf{R});C]q(T,\mathbf{R})|0\rangle\,. \tag{14.45}$$

Considering quarks as very heavy (classical c-number) external sources and path C, represented by closed rectangle, shown in Figure 14.3, we may rewrite equation (14.45) as

$$\Omega(T,R) \sim e^{-2mT} W(C) \sim e^{-E(R)T}\,, \tag{14.46}$$

where

$$W(C) = \langle 0|SpU[x,x';C]|0\rangle \tag{14.47}$$

defines the so-called Wilson loop. The behavior of the correlator $W(C)$ determines the presence or absence of confinement. In fact, from (14.46) it is clear that

$$\lim_{T\to\infty} W(C) \sim \exp\{-T[E(R)-2m]\}\,. \tag{14.48}$$

As we shall see below, in the limit of the strong coupling ($g\to\infty$) of the lattice theory the Wilson loop satisfies the so-called *area law*, so that for large enough contour C we have

$$W(C) \sim \exp\{-KA(C)\}\,, \tag{14.49}$$

where K is some constant and $A(C)$ is the area encircled on the lattice by contour C (i. e., the minimal area of the surface, with its border defined by C). For the rectangular contour shown in Figure 14.3 we have

$$A(C) = TR\,. \tag{14.50}$$

Then from (14.48), (14.49), and (14.50) we obtain

$$T[E(R)-2m] \sim KTR \quad \text{or} \quad E(R)-2m \sim KR\,, \tag{14.51}$$

i. e., linearly growing with R interaction potential in the $q\bar{q}$ system, which obviously corresponds to the confinement. Coefficient K is called the string tension (the force of confinement). This term is connected with the picture of gluon fields between quarks being in a tube – "string" – to produce linearly growing potential. This string connects quarks, and its tension grows when quarks move from each other, thus preventing their separation at large distances.

Area law in strong coupling expansion

Let us present a schematic derivation of area law in the limit of strong coupling. Link variables, associated with gauge fields, can be used as the main dynamic degrees of freedom in lattice theories. This allows us to write (14.47) in the form of "functional" integral[4]

$$W(C) = \frac{1}{Z} \int \prod_{m,\mu} dU(n, n+\hat{\mu}) SpU(x, x; C) \exp\left\{ -\frac{1}{8\pi g^2} \sum_p SpU_p \right\}, \quad (14.52)$$

where

$$Z = \int \prod_{m,\mu} dU(n, n+\hat{\mu}) \exp\left\{ -\frac{1}{8\pi g^2} \sum_p SpU_p \right\}. \quad (14.53)$$

Note that, here, there is no need of additional gauge fixing terms, as the link variables change in the limited interval. Correspondingly, the volume of field configurations space generated by gauge transformations is actually finite.

Link variables, as was shown above, are the elements of the $SU(3)$ group. The unitary matrices of $SU(3)$ are parametrized by eight generalized Euler angles, so that the group integrals in (14.52), (14.53) can be explicitly written via these angles. We shall not do so, limiting ourselves to the quotation of the following orthogonality conditions [13]:

$$\begin{aligned}
&\int dU(n, n+\hat{\mu}) [U(n, n+\hat{\mu})]_{ij} = 0\,, \\
&\int dU(n, n+\hat{\mu}) [U(n, n+\hat{\mu})]_{ij} [U^+(n, n+\hat{\mu})]_{kl} = \frac{1}{3} \delta_{il} \delta_{jk}\,, \\
&\int dU(n, n+\hat{\mu}) [U(n, n+\hat{\mu})]_{ij} [U(n, n+\hat{\mu})]_{kl} = 0\,.
\end{aligned} \quad (14.54)$$

Equations (14.54) mean that during the computation of the integrals determining (14.52), the only nonzero contributions are from the lattice links, which are passed in opposite directions. Thus, if we consider two neighboring plaquettes of the same orientation, then after integration over the variables defined on their common link, these plaquettes are "joined" in one rectangle, as shown in Figure 14.4.

[4] On a lattice this is just the usual multiple integral!

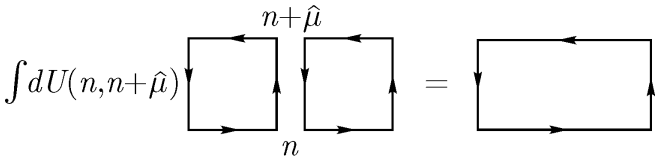


Figure 14.4

In the strong coupling limit the value of $\frac{1}{g^2}$ is considered to be a small parameter. Thus, the exponent in (14.52) can be expanded as

$$W(C) = \frac{1}{Z}\int \prod_{n,\mu} dU(n, n+\hat{\mu}) SpU(x,x;C)\Big[1 - \frac{1}{8\pi g^2}\sum_p SpU_p + \frac{1}{2!}\left(\frac{1}{8\pi g^2}\right)^2 \sum_p \sum_{p'} SpU_p SpU_{p'} + \cdots\Big]. \quad (14.55)$$

For simplicity we can consider a rectangular path C. According to equations (14.54), in this expansion a nonzero contribution comes only from those terms in the expansion in powers of $\frac{1}{g^2}$, for which the plaquettes completely fill the surface, encircled by the path C. Only in this case is each link in group integral passed twice in opposite directions (or is not passed at all), so that the corresponding integrals over the link variables give finite contributions. Thus, the nonzero contribution to $W(C)$ in the lowest order comes from the term of the order of $(\frac{1}{g^2})^{N_p}$, where N_p is the minimum number of plaquettes necessary to fill the surface encircled by C:

$$W(C) \sim \left(\frac{1}{g^2}\right)^{N_p}. \quad (14.56)$$

This corresponds to area law, as the surface area of C is given by

$$A(C) = a^2 N_p\,. \quad (14.57)$$

Then

$$W(C) \sim (g^2)^{-A(C)/a^2} = \exp\{-(TR\ln g^2)/a^2\}\,. \quad (14.58)$$

Comparing this expression with (14.49) and (14.51), we obtain linearly the growing potential

$$E(R) = KR\,, \quad \text{where } K = \frac{1}{a^2}\ln g^2\,, \quad (14.59)$$

which corresponds to $g^2(a) \sim e^{Ka^2}$.

We can also consider the weak coupling expansion for a Wilson loop, transforming to the continuous limit and taking the action in a Gaussian approximation. In this case the *perimeter* law is obtained, which corresponds to Coulomb potential $E(R) \sim \frac{1}{R}$.

Does all this mean, that we have proved the confinement? No! All our argumentation can actually be repeated also for the Abelian $SU(1)$ theory; we never used a non-Abelian nature of $SU(3)$. Strong coupling and weak coupling regimes can be separated by one or several phase transitions, taking place at different values of coupling constant g. There is no general proof of the absence of such transitions in QCD. This problem was thoroughly studied numerically, using Monte-Carlo simulations. These calculations has shown that in QCD there are no phase transitions at intermediate values of g, and there is continuous crossover from $g^2(a) \sim e^{Ka^2}$ dependence of equation (14.59) in the strong coupling region, to weak the coupling region with asymptotically free behavior $g^2(a) \sim \frac{1}{\ln a^{-1}}$, valid for $a \to 0$. The interaction potential of quarks, following from these calculations, is well approximated by the superposition of the Coulomb potential, dominating at small distances, and linearly growing potential, determining confinement at large distances: $V(R) = \frac{C}{R} + KR$. A typical result of such calculations is shown in Figure 14.5 [61], where we show the potential acting between two static quarks, calculated for the lattice with 32^4 sites, with the link $a = 0.055 \cdot 10^{-13}$ cm. the continuous line shows the fit with the superposition of the

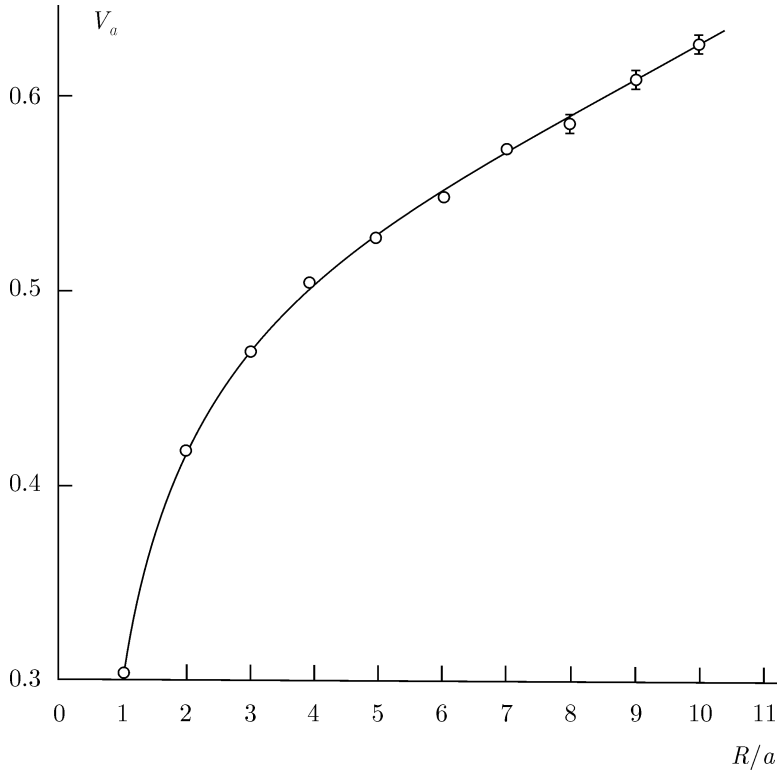


Figure 14.5

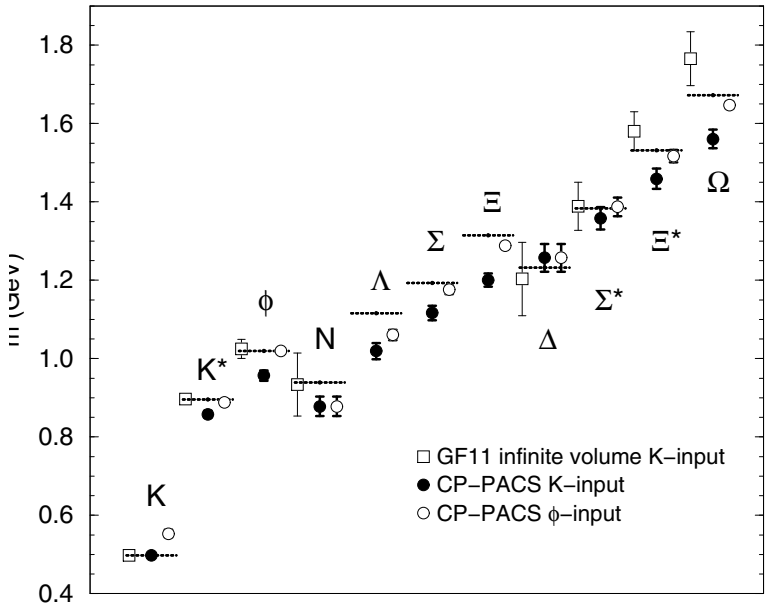


Figure 14.6

Coulomb and linear potentials. It is clearly seen that the linear growth of $V(R)$ takes place at distances $R > 0.25 \cdot 10^{-13}$ cm. At smaller distances we have the usual perturbation theory dynamics and asymptotic freedom. A typical value of string tension, following from these calculations, is $K \approx 0.2\,\text{GeV}^2 \approx 1.0{\cdot}10^{13}\,\text{GeVcm}^{-1} \approx 14$ tons! Thes effectively proves the confinement.

The details of Monte-Carlo calculations for the lattice field theories are well described in [11, 43]. The current situation with analytical models of confinement is reviewed in [61].

The study of lattice models has become one of the most important and actively developing directions in quantum field theory. As an illustration we show in Figure 14.6 [4] the early results of Monte-Carlo calculations of the masses of light hadrons, considered as bound states of quarks and gluons, which demonstrates a rather satisfactory agreement with experiments. Current results on hadron masses are well described in [11].

Confinement is not absolute, and at some very high temperature T_c (deconfinement temperature), or at some very high density there is a phase transition from the phase of hadron matter to quark–gluon plasma [72]. Physically this is rather clear. If Λ is some characteristic momentum scale characterizing the transition to asymptotically free behavior, then at $T \gg \Lambda$ the transferred momentum in scattering processes will almost always satisfy the inequality $Q^2 \gg \Lambda^2$. Correspondingly we can apply the usual perturbation theory. But in the perturbative approach to QCD, both quarks and gluons are

physical states of the theory. This means that at $T \gg \Lambda$ we have nearly an ideal gas of quarks and gluons (quark–gluon plasma). This phase transition is quite important in astrophysics for neutron stars and cosmology. Experimentally this transition can be observed in collisions of heavy nuclei, and there are already some indications for it in CERN experiments. The value of T_c was calculated by Monte-Carlo in lattice QCD. Typical values obtained show that T_c is somewhere in the interval of 0.15–0.20 MeV. A detailed review of phase transitions in QCD can be found in [45].

14.2 Effective potential and loop expansion

To study the theories with spontaneous symmetry-breaking the convenient concept there is the so-called effective potential. It allows a universal analysis of these theories and calculation of quantum corrections to the classical picture of spontaneous symmetry-breaking which was discussed above.

Let us once again discuss the simplest case of a scalar field:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \varphi)^2 - V(\varphi)\,, \quad V(\varphi) = \frac{m^2}{2}\varphi^2 + \frac{g}{4!}\varphi^4\,, \quad S[\varphi] = \int d^4x \mathcal{L}\,. \quad (14.60)$$

This Lagrangian is invariant with respect to $\varphi \to -\varphi$, but in the case of spontaneous symmetry-breaking this property is absent for the solutions of the equation

$$\left.\frac{dV}{d\varphi}\right|_{\varphi=\varphi_0} = 0\,, \quad (14.61)$$

where $\varphi_0 \neq 0$. This is already obvious from our previous analysis.

Quantum corrections, as we have seen above, appear from loop expansion containing divergences which require renormalization. The conditions for renormalization were formulated in terms of irreducible vertices $\Gamma^{(n)}$. The generating functional for $\Gamma^{(n)}(x_1, \ldots, x_n)$ is the *effective action* $\Gamma(\varphi)$, defined in (10.150). The meaning of this term will be clarified below.

The generating functional for connected diagrams $W[J]$ is defined according to (10.131) as

$$e^{iW[J]} = \langle 0|0\rangle_J\,. \quad (14.62)$$

Then the classical field φ_c (in the presence of source J) is determined from (10.151):

$$\varphi_c(x) = \frac{\delta W[J]}{\delta J(x)} = \frac{\langle 0|\varphi(x)|0\rangle_J}{\langle 0|0\rangle_J}\,. \quad (14.63)$$

The vacuum average $\langle\varphi\rangle$ is by definition

$$\langle\varphi\rangle = \lim_{J\to 0} \varphi_c\,. \quad (14.64)$$

According to (10.150) the effective action $\Gamma[\varphi_c]$ is given by

$$\Gamma(\varphi_c) = W[J] - \int dx J(x)\varphi_c(x) \tag{14.65}$$

and, in accordance with (10.151), it satisfies the equation

$$\frac{\delta\Gamma[\varphi_c]}{\delta\varphi_c(x)} = -J(x)\,. \tag{14.66}$$

For $J(x) \to 0$ the value of φ_c becomes constant, equal to $\langle\varphi\rangle$, so that the vacuum average of φ is the solution of the equation

$$\left.\frac{d\Gamma[\varphi_c]}{d\varphi_c}\right|_{\varphi_c=\langle\varphi\rangle} = 0\,. \tag{14.67}$$

The usual expansion of the functional $\Gamma[\varphi_c]$ in powers of φ_c is written as

$$\Gamma[\varphi_c] = \sum_{n=0}^{\infty} \frac{1}{n!} \int dx_1 \cdots dx_n \Gamma^{(n)}(x_1, \ldots, x_n)\varphi_c(x_1)\cdots\varphi_c(x_n) \tag{14.68}$$

or, equivalently, in momentum representation,

$$\Gamma[\varphi_c] = \sum_{n=0}^{\infty} \frac{1}{n!} \int dp_1 \cdots dp_n \delta(p_1 + \cdots + p_n)\Gamma^{(n)}(p_1, \ldots, p_n)\varphi_c(p_1)\cdots\varphi_c(p_n)\,. \tag{14.69}$$

Alternatively, we can expand $\Gamma[\varphi_c]$ over field φ_c and its derivatives:

$$\Gamma[\varphi_c] = \int dx \left\{-U(\varphi_c(x))\right\} + \frac{1}{2}(\partial_\mu\varphi_c)^2 Z(\varphi_c(x))\,. \tag{14.70}$$

In this case, the function (not a functional!) $U(\varphi_c)$ is called an *effective potential*. Below we shall see that in classical limit it just coincides with the potential $V(\varphi)$. In the case when $\varphi_c(x) = \langle\varphi\rangle = \varphi = \text{const}$, all terms of the last expansion except for the first one are zero, so that

$$\Gamma[\varphi] = -\Omega U(\varphi)\,, \tag{14.71}$$

where Ω is the total volume, filled by the field in space-time. Comparing (14.69) and (14.71), we have

$$U(\varphi) = -\sum_{n=0}^{\infty} \varphi^n \Gamma^{(n)}(p_i = 0)\,. \tag{14.72}$$

The normalization conditions for $\Gamma^{(2)}(p_i = 0)$ and $\Gamma^{(4)}(p_i = 0)$ can be reformulated in terms of potential U:

$$\left.\frac{d^2U(\varphi_c)}{d^2\varphi_c}\right|_{\varphi_c=\langle\varphi\rangle} = m^2\,, \tag{14.73}$$

$$\left.\frac{d^4U(\varphi_c)}{d^4\varphi_c}\right|_{\varphi_c=\langle\varphi\rangle} = g\,. \tag{14.74}$$

Besides, the condition (14.67) for the vacuum average takes the form

$$\left.\frac{dU(\varphi_c)}{d\varphi_c}\right|_{\varphi_c=\langle\varphi\rangle} = 0\,. \tag{14.75}$$

To study the properties of the theory with spontaneous symmetry-breaking, it is convenient to define the new field φ':

$$\varphi' = \varphi - \langle\varphi\rangle\,, \tag{14.76}$$

for which the vacuum average is simply zero.

Note that all divergences of the theory were hidden in counter-terms before we apply normalization conditions (14.73), (14.74), so that in the theory with spontaneous symmetry-breaking no new divergences appear (in addition to the theory without symmetry-breaking), so that the structure of the divergences in renormalized field theory is not changed by spontaneous symmetry-breaking.

Let us calculate the effective potential. We shall use (14.65) and start with calculations of the functional $W[J]$ by the stationary phase (or steepest descent method). Let us recall what steepest descent calculation is in the case of the usual integral of the form

$$I = \int_{-\infty}^{\infty} dx\, e^{-f(x)}\,. \tag{14.77}$$

Assume that the function $f(x)$ has a minimum at some point x_0. Then we have

$$f(x) = f(x_0) + \frac{1}{2}(x-x_0)^2 f''(x_0) + \cdots\,, \tag{14.78}$$

so that we can write

$$I \approx e^{-f(x_0)} \int_{-\infty}^{\infty} dx\, e^{-\frac{1}{2}(x-x_0)^2 f''(x_0),} \tag{14.79}$$

and the problem reduces to the calculation of the well-known Gaussian integral, which is an easy task.

Restoring the Planck constant in the definition of the functional $W[J]$, we have

$$e^{\frac{i}{\hbar}W[J]} = \int \mathcal{D}\varphi\, e^{\frac{i}{\hbar}S[\varphi,J]}\,, \tag{14.80}$$

where

$$S[\varphi, J] = \int d^4x[\mathcal{L}(\varphi) + \hbar\varphi(x)J(x)]\,. \tag{14.81}$$

From (14.60) and (14.61) it follows that

$$\left.\frac{\delta S[\varphi, J]}{\delta\varphi(x)}\right|_{\varphi_0} = \hbar J(x)\,. \tag{14.82}$$

For $J \to 0$ this reduces to the condition of extremal action. Let us expand the action in the vicinity of φ_0:

$$\begin{aligned} S[\varphi, J] &= S[\varphi_0, J] + \int dx[\varphi(x) - \varphi_0] \left.\frac{\delta S}{\delta\varphi(x)}\right|_{\varphi_0} \\ &\quad + \int dx \int dy\, \frac{1}{2}[\varphi(x) - \varphi_0] \left.\frac{\delta^2 S}{\delta\varphi(x)\delta\varphi(y)}\right|_{\varphi_0} [\varphi(y) - \varphi_0] + \cdots \\ &= S[\varphi_0, J] + \hbar \int dx[\varphi(x) - \varphi_0] J(x) \\ &\quad + \frac{1}{2} \int dx \int dy[\varphi(x) - \varphi_0] \left.\frac{\delta^2 S}{\delta\varphi(x)\delta\varphi(y)}\right|_{\varphi} [\varphi(y) - \varphi_0] + \cdots, \end{aligned} \tag{14.83}$$

where we have used (14.82). Performing functional differentiation we can understand that

$$\left.\frac{\delta^2 S}{\delta\varphi(x)\delta\varphi(y)}\right|_{\varphi_0} = -[\Box + V''(\varphi_0)]\delta(x - y). \tag{14.84}$$

If we take $\varphi' = \varphi - \varphi_0$, the expansion (14.83) takes the form

$$S[\varphi, J] = S[\varphi_0, J] + \hbar \int dx \varphi'(x) J(x) - \frac{1}{2} \int dx \varphi'(x)[\Box + V''(\varphi_0)]\varphi'(x) + \cdots . \tag{14.85}$$

Substitution of this expression into (14.80), to use the stationary phase approach, gives (here we write φ instead of φ')

$$\exp\left(\frac{i}{\hbar} W\right) = \exp\left\{\frac{i}{\hbar} S[\varphi_0, J]\right\} \int \mathcal{D}\varphi \exp\left\{-\frac{i}{\hbar}\frac{1}{2} \int dx \varphi[\Box + V''(\varphi_0)]\varphi\right\}, \tag{14.86}$$

where we have dropped the contribution of the second term in (14.85), as in the following we shall make the transition to the limit of $J \to 0$. To obtain the loop expansion (equivalent, as we have seen, to the expansion in $\hbar$), we replace $\varphi \to \hbar^{1/2}\varphi$, thus excluding $\hbar$ from the second exponent in (14.86). Let us transform the integral to Euclidean space so that i also vanishes. Now, calculating the functional integral we can use the usual Gaussian expression and obtain

$$\exp\left(\frac{i}{\hbar} W\right) = \exp\left\{\frac{i}{\hbar} S[\varphi_0, J]\right\} (\mathrm{Det}[\Box + V''(\varphi_0)])^{-1/2}. \tag{14.87}$$

Using the relation $\mathrm{Det}\, A = \exp Sp \ln A$, we get

$$W[J] = S[\varphi_0] + \hbar \int dx \varphi_0(x) J(x) + \frac{i\hbar}{2} Sp \ln[\Box + V''(\varphi_0)]. \tag{14.88}$$

This expression gives $W[J]$ with a single-loop correction, while the terms $O(\hbar^2)$ are dropped. Now we can substitute (14.88) into (14.65). But first we express $S[\varphi_c]$ via

$S[\varphi_0]$. Taking $\varphi_0 = \varphi_c - \varphi_1$, we have

$$S[\varphi_0] = S[\varphi_c - \varphi_1] \tag{14.89}$$

$$= S[\varphi_c] - \int dx \varphi_1(x) \left.\frac{\delta S}{\delta \varphi(x)}\right|_{\varphi_c} + \cdots$$

$$= S[\varphi_c] - \hbar \int dx \varphi_1(x) J(x) + \cdots . \tag{14.90}$$

Then the substitution of (14.88) and (14.90) into (14.65), in the limit of $J \to 0$, yields

$$\Gamma[\varphi_c] = S[\varphi_c] + \frac{i\hbar}{2} Sp \ln[\Box + V''(\varphi_0)] \tag{14.91}$$

which is the effective action with a single-loop quantum correction. Take now $\varphi_c(x) = \varphi = \text{const}$. Then $\Gamma[\varphi]$ is determined by (14.71), while from (14.60) it follows that $S[\varphi] = -\Omega V(\varphi)$. Correspondingly, from (14.91) we obtain the effective potential of the form

$$U(\varphi) = V(\varphi) - \frac{i\hbar}{2} \Omega^{-1} Sp \ln[\Box + V''(\varphi)] . \tag{14.92}$$

Now we see that for $\hbar \to 0$ (in the classical limit) the effective potential just coincides with the classical potential $V(\varphi)$, while the effective action (14.91) reduces to the classical action (14.60). A trace of an operator gives the sum (integral) over all eigenvalues, and we can (after the transformation to Euclidean momentum space) rewrite (14.92) as

$$U(\varphi) = V(\varphi) + \frac{\hbar}{2} \int \frac{d^4 k_E}{(2\pi)^4} \ln[k_E^2 + V''(\varphi)]$$

$$= V(\varphi) + \frac{\hbar}{2} \int \frac{d^4 k_E}{(2\pi)^4} \ln\left(k_E^2 + m^2 + \frac{1}{2} g\varphi^2\right) . \tag{14.93}$$

This type of expression can be used to study spontaneous symmetry-breaking taking into account the quantum corrections. In the classical limit we used above, for $m^2 > 0$ the vacuum is nondegenerate, while for $m^2 < 0$ there was spontaneous symmetry-breaking and the appearance of degeneracy of the vacuum (phase transition). What will happen taking into account the quantum corrections, in particular at $m^2 = 0$? From the single-loop expression for effective potential it follows that the nontrivial minimum $\varphi \neq 0$ already appears at $m^2 = 0$, so that we have spontaneous symmetry-breaking due to quantum corrections. Unfortunately, the single-loop approximation is insufficient for a complete understanding of this problem [13, 56]. A more detailed discussion of effective potential formalism, its relation to the renormalization group, and the other aspects of the theory are presented in [27].

The loop expansion, considered above, is in fact the expansion in powers of $\hbar$, not in powers of the coupling constant g. In this sense it is nonperturbative, but in fact this is not precisely

so! In classical theory, g is irrelevant by itself. It can be easily understood if we make the transformation to $\varphi' = g\varphi$. Then the Lagrangian of φ^4-theory can be rewritten as

$$\mathcal{L} = \frac{1}{g^2}\left(\frac{1}{2}\partial_\mu\varphi'\partial^\mu\varphi' - \frac{1}{2}m^2\varphi'^2 - \varphi'^4\right). \tag{14.94}$$

and g just drops from the classical equations of motion, becoming irrelevant. This is obviously not the case in quantum theory, which is essentially due to the appearance of $\hbar$. In quantum theory we are always dealing with the ratio:

$$\frac{1}{\hbar}\mathcal{L} = \frac{1}{g^2\hbar}\left(\frac{1}{2}\partial_\mu\varphi'\partial^\mu\varphi' + \cdots\right) \tag{14.95}$$

and the relevant parameter is $g^2\hbar$. Thus, the quasi-classical approximation (small $\hbar$) is, in fact, intimately connected with the weak coupling approximation (small g).

14.3 Instantons in quantum mechanics

Nontrivial nonperturbative effects can arise in quantum field theory even for small values of the coupling constant, and perturbation theory may become inadequate when naïvely it should be applicable[5]. A simple example is quantum tunneling, which we shall consider below.

Let us start with quantum mechanics. Consider a particle with unit mass, moving in one-dimensional potential:

$$H = \frac{1}{2}p^2 - V(x). \tag{14.96}$$

Below we shall derive the well-known results of quantum mechanics, but in a rather unusual way, which will be further generalized for the case of quantum field theory.

Euclidean path integrals

Consider the Euclidean (imaginary time!) version of the Feynman path integral:

$$\langle x_f|e^{-HT/\hbar}|x_i\rangle = \mathcal{N}\int [dx]e^{-S/\hbar}. \tag{14.97}$$

Here $|x_i\rangle$ and $|x_f\rangle$ are eigenstates of the coordinate operator of the particle, H is its Hamiltonian. Here the integration measure which we previously denoted as $\mathcal{D}x$, is written as $[dx]$; T is considered to be positive.

Let us introduce as usual the complete set of eigenstates of the Hamiltonian:

$$H|n\rangle = E_n|n\rangle \tag{14.98}$$

and write

$$\langle x_f|e^{-HT/\hbar}|x_i\rangle = \sum_n e^{-E_nT/\hbar}\langle x_f|n\rangle\langle n|x_i\rangle. \tag{14.99}$$

Then in the limit of $T \to \infty$ only the contribution of the ground state survives.

[5] Below we follow mainly lectures by Coleman [15]; these problems are discussed in more detail in the nice book by Rajaraman [52]

In the right-hand side of (14.97) is the Euclidean action:

$$S = \int_{-T/2}^{T/2} dt \left[\frac{1}{2} \left(\frac{dx}{dt} \right)^2 + V \right]. \tag{14.100}$$

Integration $[dx]$ is performed over all trajectories, with boundary conditions $x(-T/2) = x_i$, $x(T/2) = x_f$. In more detail, if $\bar{x}(t)$ is the given function and satisfies these conditions, the arbitrary function, satisfying the same conditions, can be written as

$$x(t) = \bar{x}(t) + \sum_n c_n x_n(t)\,, \tag{14.101}$$

where $x_n(t)$ is the complete set of orthonormalized functions, being zero at the boundaries:

$$\int_{-T/2}^{T/2} dt x_n(t) x_m(t) = \delta_{nm}\,, \quad x_n(\pm T/2) = 0\,. \tag{14.102}$$

Then the integration measure $[dx]$ can be defined as

$$[dx] = \prod_n (2\pi\hbar)^{-1/2} dc_n\,. \tag{14.103}$$

It is obvious, that acting in this way we take into account all the paths, and this definition differs from that of Feynman only by the normalization constant.

The right-hand side of equation (14.97) is easily calculated in quasi-classical approximation (small $\hbar$!). The main contribution to the action comes from the vicinity of the extremal trajectory, defined by

$$\frac{\delta S}{\delta \bar{x}} = -\frac{d^2\bar{x}}{dt^2} + V'(\bar{x}) = 0, \tag{14.104}$$

Euclidean Newton equations. Let us choose x_n as the eigenfunctions of the second variational derivative of action S at $\bar{x}$:

$$-\frac{d^2 x_n}{dt^2} + V''(\bar{x}) x_n = \lambda_n x_n\,. \tag{14.105}$$

Similarly to the analysis carried out at the beginning of Chapter 2, the first variation of the action, due to variation of trajectory $x \to \bar{x} + a$, in this case reduces to

$$\delta S = \int_{-T/2}^{T/2} dt\, a \left[-\frac{d^2 x}{dt^2} + V'(x) \right] = 0 \quad \text{for } x = \bar{x}\,, \tag{14.106}$$

which leads to Newton law (14.104). If we vary once again $x \to \bar{x} + a$, we get

$$\begin{aligned} \delta^2 S &= \int_{-T/2}^{T/2} dt\, a \left[-\frac{d^2(\bar{x}+a)}{dt^2} + V'(\bar{x}+a) \right] = \int_{-T/2}^{T/2} dt\, a[-\ddot{\bar{x}} - \ddot{a} + V'(\bar{x}) + V''(\bar{x})a] \\ &= \int_{-T/2}^{T/2} dt\, a[-\ddot{a} + V''(\bar{x})a]\,, \end{aligned} \tag{14.107}$$

where during the transformation to the last equality we used equations of motion (14.104). Now it is clear that the second variational derivative of action S is determined by the left-hand side of equation (14.105).

Then in the limit of small $\hbar$, after substituting (14.101) into (14.100), we can limit ourselves to quadratic deviations from the classical trajectory $\bar{x}$, so that the integrals over c_n become Gaussian, and we find

$$\begin{aligned}\langle x_f|e^{-HT/\hbar}|x_i\rangle &= \mathcal{N}e^{-S(\bar{x})/\hbar}\prod_n \lambda_n^{-1/2}[1+O(\hbar)] \\ &= \mathcal{N}e^{-S(\bar{x})/\hbar}\left[\mathrm{Det}(-\partial_t^2+V''(\bar{x}))\right]^{-1/2}[1+O(\hbar)]\,. \end{aligned} \tag{14.108}$$

If there are several stationary points of action, the corresponding contributions should be summed.

Note that the Euclidean equation of motion (14.104) is equivalent to the usual Newtonian equation for a particle with unit mass, moving in the inverted potential *minus* V! For such an equation we have the integral of motion:

$$E = \frac{1}{2}\left(\frac{d\bar{x}}{dt}\right)^2 - V(\bar{x})\,. \tag{14.109}$$

Consider the potential V shown in Figure 14.7(a). Let $x_i = x_f = 0$. The inverted potential is shown in Figure 14.7(b). It is obvious that the only solution of classical equations of motion in this potential, satisfying boundary conditions, is

$$\bar{x} = 0 \tag{14.110}$$

i. e., the particle stays at rest at the top. For this solution we have $S = 0$. Then from (14.108) we have

$$\langle 0|e^{-HT/\hbar}|0\rangle = \mathcal{N}[\mathrm{Det}(-\partial_t^2+\omega^2)]^{-1/2}[1+O(\hbar)]\,, \tag{14.111}$$

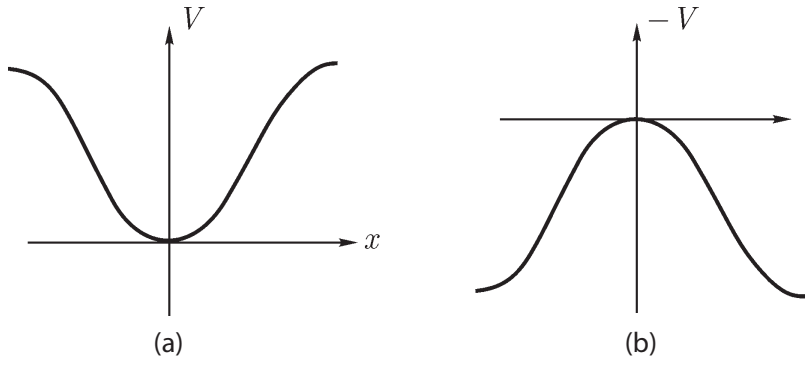


Figure 14.7

where zeroes denote transition from the origin to the origin of our coordinate system, while $\omega^2 = V''(0)$ is the square of the frequency of the small oscillations around the minimum of potential V. It can be shown [15] that for large T

$$\mathcal{N}[\mathrm{Det}(\partial_t^2 + \omega^2)]^{-1/2} = \left(\frac{\omega}{\pi\hbar}\right)^{1/2} e^{-\omega T/2}\,. \tag{14.112}$$

Then from (14.111) and the discussion around (14.99) we immediately see that the ground state energy in this problem is given by

$$E_0 = \frac{1}{2}\hbar\omega[1 + O(\hbar)]\,, \tag{14.113}$$

i. e., the zero-point energy of the oscillator near the minimum of V. The probability for a particle to be at the origin of the coordinate system when it is at its ground state is given by

$$|\langle x = 0|n = 0\rangle|^2 = \left(\frac{\omega}{\pi\hbar}\right)^{1/2}[1 + O(\hbar)]\,. \tag{14.114}$$

These are the well-known results of the quasi-classical approximation of quantum mechanics. Actually, from this correspondence we immediately see the validity of (14.112). The physics is also quite clear: in the limit of small $\hbar$ the particle is in the ground state of oscillator near th#e origin of coordinate system.

Double-well potential and instantons

Let us consider now a more interesting example: the double-well potential, shown in Figure 14.8(a). Here we have $V(x) = V(-x)$, and the potential minima are at points $x = \pm a$. We can also introduce $\omega^2 = V''(\pm a)$: the square of the frequency of the classical oscillations of a particle in the vicinity of the minima. Let us calculate the transition amplitudes:

$$\langle -a|e^{-HT/\hbar}|-a\rangle = \langle a|e^{-HT/\hbar}|a\rangle\,, \tag{14.115}$$

$$\langle a|e^{-HT/\hbar}|-a\rangle = \langle -a|e^{-HT/\hbar}|a\rangle\,, \tag{14.116}$$

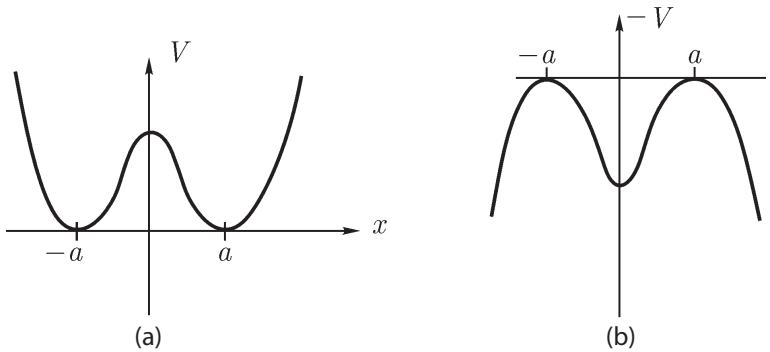


Figure 14.8

making a quasi-classical approximation for the path integral, similar to the case of a particle in single well. As a first step we shall again look for the solutions of the classical Euclidean equations of motion (14.104), satisfying the appropriate boundary conditions. There are two obvious solutions: one corresponding to the case of the particle remaining the entire time at the top of the left or right hill in Figure 14.8(b). However, there may be also another more interesting solution when the particle starts from one of the tops (e. g., the left one) at the moment $-T/2$ and rolls to the right top, reaching it at the moment $+T/2$ ($T \to \infty$). Here we are dealing with solutions of equations of motion, corresponding to energy $E = 0$ (because $E = 0$ in the initial states $x = \pm a$). Correspondingly (cf. (14.109)),

$$\frac{d\bar{x}}{dt} = \sqrt{2V}\,, \tag{14.117}$$

and the solution of this equation has the form

$$t = t_1 + \int_0^x dx' \frac{1}{\sqrt{2V}}\,, \tag{14.118}$$

where t_1 is the integration constant (time at which $x = 0$). This solution obviously has the form shown in Figure 14.9. This solution is called an *instanton*[6], centered at point t_1. A mirror reflection of this solution can be called an anti-instanton. It is important to stress that an instanton has a *finite* action:

$$S_0 = \int_{-T/2}^{T/2} dt \left[\frac{1}{2}\left(\frac{d\bar{x}}{dt}\right)^2 + V\right] = \int_{-T/2}^{T/2} dt \left(\frac{d\bar{x}}{dt}\right)^2 = \int_{-a}^{a} dx\sqrt{2V}\,, \tag{14.119}$$

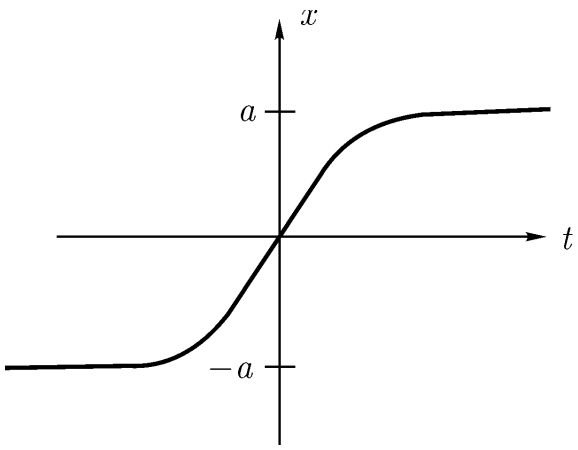


Figure 14.9

[6] The origin of this term is related to the obvious analogy with soliton, but stresses the fact that we are dealing here with solutions of Euclidean equations of motion.

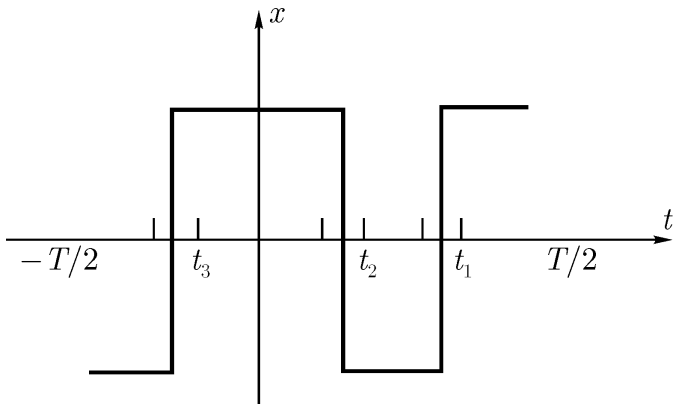


Figure 14.10

where we have used (14.117). For large t we have $\bar{x} \to a$, so that (14.117) can be approximated by

$$\frac{d\bar{x}}{dt} = \sqrt{\omega^2(\bar{x}-a)^2} = \omega(a-\bar{x}) \tag{14.120}$$

so that at large t,

$$(a-\bar{x}) \sim e^{-\omega t} \tag{14.121}$$

and the instanton is "localized in time" (at times $\sim \frac{1}{\omega}$), which clarifies its name.

It is clear that for large T the instanton and anti-instanton are not the only solutions of equations of motion; approximate solutions can be built as chains of well-separated instantons and anti-instantons. An example of such a configuration is shown in Figure 14.10, with n objects (instantons and anti-instantons), centered at points $t_1, \ldots, t_n$:

$$\frac{T}{2} > t_1 > t_2 > \cdots > t_n > -\frac{T}{2}\,. \tag{14.122}$$

Correspondingly, in the path integral we have to sum the contributions from all such configurations.

Now let us calculate! Obviously, n well-separated objects contribute additively to the actions giving $S = nS_0$, which in the path integral give the contribution of the order $\sim \exp(-nS_0)$. To find the determinant we need the more complicated procedure. Consider the time evolution operator e^{-HT} as the product of the operators of evolution between the points where instantons and anti-instantons are placed. In the absence of these, on all time axes we have $V'' = \omega^2$, and we get the same result as above for the case of single-well potential (14.112):

$$\left(\frac{\omega}{\pi\hbar}\right)^{1/2} e^{-\omega T/2}\,. \tag{14.123}$$

Intervals with instantons and anti-instantons lead to a correction, which we can write as

$$\left(\frac{\omega}{\pi\hbar}\right)^{1/2} e^{-\omega T/2} K^n \,, \tag{14.124}$$

where K can be determined from the requirement of a correct answer for the case of a single instanton. Later we shall give the appropriate explicit expression, while for the moment we note that in order to take into account all the possible contributions to the path integral we have to integrate over the arbitrary positions of all centers:

$$\int_{-T/2}^{T/2} dt_1 \int_{-T/2}^{t_1} dt_2 \cdots \int_{-T/2}^{t_{n-1}} dt_n = \frac{T^n}{n!} \,. \tag{14.125}$$

We should also take into account the fact, that we are not completely free in placing the instantons and anti-instantons. For example, if we start from $-a$, the first object we meet is to be an instanton, the next one an anti-instanton, etc. If we finally return to $-a$, n should be even, while if we end at $+a$, n should be odd. Thus we obtain

$$\langle -a|e^{-HT/\hbar}|-a\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{1/2} e^{-\omega T/2} \sum_{\text{Even } n} \frac{(Ke^{-S_0/\hbar}T)^n}{n!}[1+O(\hbar)] \,, \tag{14.126}$$

while $\langle a|e^{-HT/\hbar}|-a\rangle$ is given by the same expression, but with summation over odd n. The sums are elementary, and we obtain

$$\langle \pm a|e^{-HT/\hbar}|-a\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{1/2} e^{-\omega T/2} \frac{1}{2}\big[\exp(Ke^{-S_0/\hbar}T) \mp \exp(-Ke^{-S_0/\hbar}T)\big] \,. \tag{14.127}$$

Recalling (14.99),

$$\langle x_f|e^{-HT/\hbar}|x_i\rangle = \sum_n e^{-E_nT/\hbar}\langle x_f|n\rangle\langle n|x_i\rangle \,, \tag{14.128}$$

we understand that the two lowest energy levels correspond to energies

$$E_\pm = \frac{1}{2}\hbar\omega \pm \hbar K e^{-S_0/\hbar} \,. \tag{14.129}$$

If we denote the corresponding states as $|+\rangle$ and $|-\rangle$, we see that

$$|\langle +|\pm a\rangle|^2 = |\langle -|\pm a\rangle|^2 = \langle a|-\rangle\langle -|-a\rangle = -\langle a|+\rangle\langle +|-a\rangle = \frac{1}{2}\left(\frac{\omega}{\pi\hbar}\right)^{.1/2} \tag{14.130}$$

These are the well-known results of quantum mechanics [35]: we just obtained the splitting of the level in the double-well potential due to tunneling (degeneracy lifting of two levels in two potential wells). The size of this splitting is $\sim e^{-S_0/\hbar}$. The lowest state $|-\rangle$ is even combination of wave functions, corresponding to a particle, localized in each of the wells, while the first excited state $|+\rangle$ is described by antisymmetric combination of these functions.

Now let us calculate the factor K. First, let us study the properties of the solutions of equation (14.105):

$$-\frac{d^2x_n}{dt^2} + V''(\bar{x})x_n = \lambda_n x_n \,, \tag{14.131}$$

where $\bar{x}$ denotes a single instanton solution. Due to invariance with respect to the time-shift (the instanton center can be placed at an arbitrary point on the time axis), this equation has an eigenfunction with zero eigenvalue (the so-called zero translation mode). Explicitly this function is written as

$$x_1 = S_0^{-1/2}\frac{d\bar{x}}{dt} \,. \tag{14.132}$$

The normalization factor here appears from (14.119):

$$\int dt \left(\frac{d\bar{x}}{dt}\right)^2 = S_0 \,, \tag{14.133}$$

The existence of zero-mode can be confirmed as follows. The instanton $\bar{x}(t)$ satisfies the equation (14.104):

$$-\frac{d^2\bar{x}}{dt^2} + V'(\bar{x}) = 0 \,. \tag{14.134}$$

But $\bar{x}(t+T)$, with arbitrary T, is also the solution of this equation:

$$-\frac{d^2\bar{x}(t+T)}{dt^2} + V'(\bar{x}(t+T)) = 0 \,. \tag{14.135}$$

Differentiating this equation by T, we obtain

$$-\frac{d^2x_1}{dt^2} + V''(\bar{x})x_1 = 0 \,, \tag{14.136}$$

which proves our statement on (14.132).

The existence of zero-mode with $\lambda_1 = 0$ seems to lead to a problem. If we calculate the Gaussian integral around the extremal trajectory (instanton), as described in connection with (14.101), (14.103), and (14.108), integration over c_1 will lead to divergence! However, we have already done the appropriate integration, integrating over the centers of the instantons (anti-instantons) in (14.125). In fact, the change of $x(t)$ under the small shift of the instanton center t_1 is equal to

$$dx = \left(\frac{d\bar{x}}{dt}\right) dt_1 \,. \tag{14.137}$$

At the same time, the corresponding change due to variation of the coefficient c_1 in (14.101) is

$$dx = x_1 dc_1 \,. \tag{14.138}$$

Then, writing in (14.137) $\left(\frac{d\bar{x}}{dt}\right) dt_1 = \sqrt{S_0} x_1 dt_1$, and comparing this with (14.138), we get $dc_1 = \sqrt{S_0} dt_1$, or

$$(2\pi\hbar)^{-1/2} dc_1 = \left(\frac{S_0}{2\pi\hbar}\right)^{1/2} dt_1 \,, \tag{14.139}$$

where $\hbar$ is introduced to make normalization dimensionless. Thus, during the calculation of the determinant in the Gaussian integral in expressions like (14.108) we do not need to include the zero eigenvalue, but instead we have to include in K the factor $(\frac{S_0}{2\pi\hbar})^{1/2}$. Then the single-instanton contribution to the matrix element is given by

$$\langle a|e^{-HT/\hbar}|-a\rangle_{1inst} = \mathcal{N} T \left(\frac{S_0}{2\pi\hbar}\right)^{1/2} e^{-S_0/\hbar} [\mathrm{Det}'(-\partial_t^2 + V''(\bar{x}))]^{-1/2}, \tag{14.140}$$

where the prime over the determinant corresponds to dropping the zero eigenvalue. Comparing (14.140) with the single-instanton contribution to (14.126), we find

$$K = \left(\frac{S_0}{2\pi\hbar}\right)^{1/2} \left| \frac{\mathrm{Det}(-\partial_t^2 + \omega^2)}{\mathrm{Det}'(-\partial_t^2 + V''(\bar{x}))} \right|^{1/2} . \tag{14.141}$$

This completes our calculation!

Let us make some comments:

1. It can be shown that the results obtained are equivalent to the standard results of quantum mechanics [35].
2. We assumed that all $\lambda_n > 0$ (except $\lambda_1 = 0$). This is really so, as the lowest state x_1 (as is easily seen from its explicit form) does not have zeroes, as it should be for the solution of a one-dimensional Schroedinger equation. This is clear from the fact that our instanton is monotonously growing (ant-instanton: decreasing) function of t, so that its derivative $x_1 \sim \frac{d\bar{x}}{dt}$ has no zeroes.
3. The coefficient K is proportional to $\hbar^{-1/2}$, which is related to the contribution of the zero-mode. In fact, this is a general rule: each zero-mode (there may be several such modes!) produces the factor of $\hbar^{-1/2}$.

In a similar way we can analyze the problem of a particle moving in a periodic potential, as shown in Figure 14.11. The difference from the previous case is that now we have no restriction of alternating the placement of the instantons and anti-instantons, which is connected with the existence here of the multitude of equivalent potential minima. At the same time, the total number of instantons minus the total number of anti-instantons should now be equal to the change of $\bar{x}$ between the initial and final coordinates. Then, from (14.127) we obtain

$$\langle j_+|e^{-HT/\hbar}|j_-\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{1/2} e^{-\omega T/2} \sum_{n=0}^{\infty} \sum_{\bar{n}=0}^{\infty} \frac{1}{n!\bar{n}!} (K e^{-S_0/\hbar} T)^{n+\bar{n}} \delta_{n-\bar{n}-j_+ + j_-} \,, \tag{14.142}$$

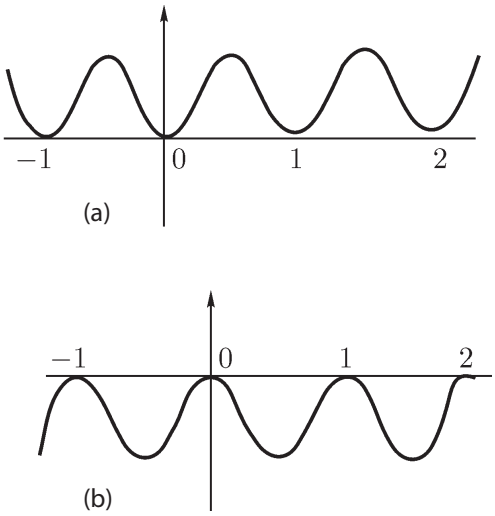


Figure 14.11

where n is the number of instantons, while $\bar{n}$ is the number of anti-instantons. Using now

$$\delta_{ab} = \int_0^{2\pi} \frac{d\theta}{2\pi} e^{i\theta(a-b)}, \tag{14.143}$$

twe rewrite (14.142) as

$$\langle j_+|e^{-HT/\hbar}|j_-\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{1/2} e^{-\omega T/2} \int_0^{2\pi} \frac{d\theta}{2\pi} \exp[2KT\cos\theta e^{-S_0/\hbar}] \exp[-i\theta(j_+ - j)]\,. \tag{14.144}$$

In this case we have the continuum of energy eigenvalues (band!), parametrized by an "angle" θ:

$$E(\theta) = \frac{1}{2}\hbar\omega - 2\hbar K e^{-S_0/\hbar}\cos\theta\,. \tag{14.145}$$

Matrix elements

$$\langle\theta|j\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{1/4} (2\pi)^{-1/2} e^{ij\theta} \tag{14.146}$$

represent, in fact, the appropriate Bloch wave.

Instantons and metastable states

Consider the potential shown in Figure 14.12(a). If we neglect tunneling, there will be a bound state at the origin. Reflected potential is shown in Figure 14.12(b). Classical equations of motion have the obvious solution corresponding to a particle starting from the top of the hill at $x = 0$, which is then reflected from the classical point of return σ and returns back to the top, as shown in Figure 14.13. Let us calculate the matrix element of transition from $x = 0$ to $x = 0$, summing over all the well-separated instantons of Figure 14.13. We can proceed as above (with obvious redefinition of S_0,

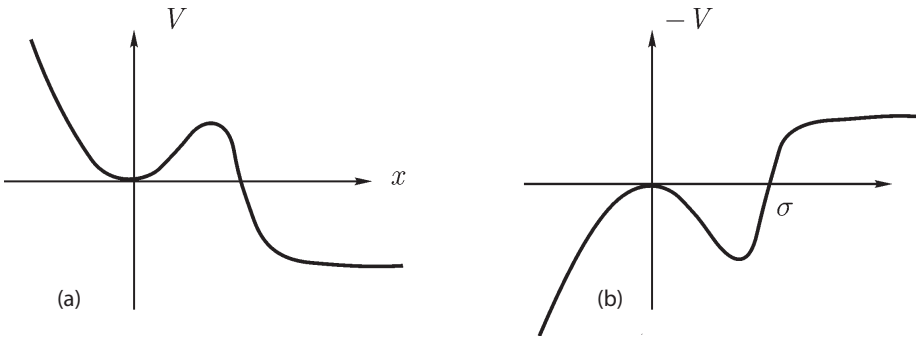


Figure 14.12

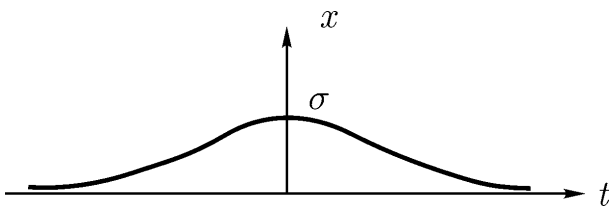


Figure 14.13

ω^2, etc.), but with no limitation on the number of instantons being even or odd. Then the summation produces the full exponent:

$$\langle 0|e^{-HT/\hbar}|0\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{1/2} e^{-\omega T/2} \exp[KTe^{-S_0/\hbar}] \tag{14.147}$$

and the ground state energy eigenvalue is

$$E_0 = \frac{1}{2}\hbar\omega - \hbar K e^{-S_0/\hbar} \tag{14.148}$$

But this is wrong! In fact, in this situation we have tunneling and the appearance of an unstable state! From the form of the instanton in Figure 14.13 it is clear that the eigenfunction $x_1 \sim \frac{d\bar{x}}{dt}$ has zero and cannot be the ground-state wave function. But its energy is zero, and now we understand that there is another state with $\lambda_0 < 0$ and an eigenfunction without zeroes. Then the factor K, containing the square root of the product of the eigenvalues, is *imaginary*. Thus, in fact we obtain

$$\text{Im}\, E_0 = \frac{\Gamma}{2} \sim \hbar|K|e^{-S_0/\hbar} \tag{14.149}$$

which corresponds to the finite-level *width*, corresponding to the metastable state.

14.4 Instantons and the unstable vacuum in field theory

Consider now Euclidean scalar field theory with the action

$$S = \int d^4x \left[\frac{1}{2}(\partial_\mu \phi)^2 + U(\phi)\right], \tag{14.150}$$

where potential $U(\phi)$ is shown in Figure 14.14. Here we have two nonequivalent minima at ϕ_+ and ϕ_-, and ϕ_- is an absolute minimum. Let us choose the origin of the energy scale so that $U(\phi_+) = 0$. In quantum theory, the minimum at $\phi = \phi_+$ plays the role of a "false" (metastable) vacuum. The description of the decay of such "false" vacuums is similar to the description of nucleation in statistical physics (e. g., during the boiling of a superheated liquid). In quantum field theory this problem is of importance for cosmology [41]. Who knows whether *our vacuum* is stable or metastable!

We have to calculate the value of $\frac{\Gamma}{V}$, the probability of metastable vacuum decay in units of time per unit volume. First we have to find the corresponding instanton $\bar{\phi}$ as a solution of the Euclidean equations of motion:

$$\partial_\mu \partial^\mu \bar{\phi} = U'(\bar{\phi}), \tag{14.151}$$

satisfying the boundary conditions

$$\lim_{x_4 \to \pm\infty} \bar{\phi}(\mathbf{x}, x_4) = \phi_+ . \tag{14.152}$$

It is easily seen that to guarantee the finiteness of the action at instanton we have to satisfy the condition

$$\lim_{|\mathbf{x}| \to \infty} \bar{\phi}(\mathbf{x}, x_4) = \phi_+ . \tag{14.153}$$

If an instanton is found, then in the leading approximation in $\hbar$ we have

$$\frac{\Gamma}{V} = K e^{-S_0}, \tag{14.154}$$

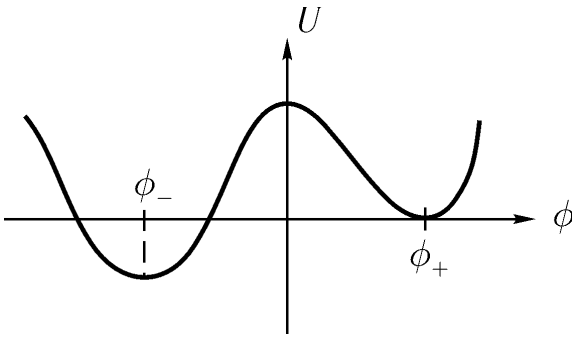


Figure 14.14

where $S_0 = S(\bar{\phi})$, while the preexponential factor K is determined by the appropriate determinant.

The trivial solution $\bar{\phi} = \phi_+$ is not interesting; for $\frac{\delta^2 S}{\delta \phi^2}$ does not have negative eigenvalues, so that it does not contribute to the vacuum decay. Equations (14.151)–(14.153) are invariant with respect to four-dimensional rotations ($O(4)$ group). We assume that an instanton is also $O(4)$ invariant[7], so that the corresponding $\bar{\phi}$ is the function of r only. Then equation (14.151) reduces to

$$\frac{d^2\bar{\phi}}{dr^2} + \frac{3}{r}\frac{d\bar{\phi}}{dr} = U'(\bar{\phi})\,, \tag{14.155}$$

and from (14.152) and (14.153) it follows that

$$\lim_{r\to\infty} \bar{\phi}(r) = \phi_+\,. \tag{14.156}$$

Obviously, we also have to require

$$\left.\frac{d\bar{\phi}}{dr}\right|_{r=0} = 0\,, \tag{14.157}$$

or $\bar{\phi}$ will be singular at the origin.

Equation (14.155) can be interpreted as an equation of motion (considering r as time!) of a particle, moving in potential *minus* U, shown in Figure 14.15, and under the action of a time-dependent friction force ($\sim \frac{1}{r}\times$ velocity). The particle can start from the state of rest (cf. (14.157)) at the moment $r = 0$ from the appropriate initial position and stops at $r \to \infty$ at point ϕ_+: such motion precisely corresponds to an instanton. Obviously, such a solution exists.

- The particle starting to the right of ϕ_0 will not reach ϕ_+; it will not have enough energy because of friction.

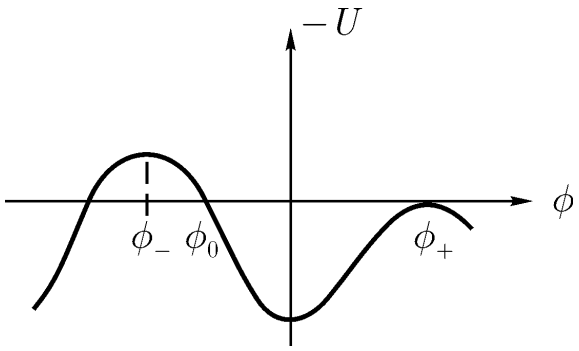


Figure 14.15

[7] This assumption can be rigorously justified: a spherically symmetric instanton has the minimal action.

- If we choose the initial point correctly to the left of ϕ_0 but to the right of ϕ_-, we can guarantee that for large r the particle will reach ϕ_+ and stop there.

In fact, for ϕ close to ϕ_- we can linearize the equation of motion and write it as

$$\left(\frac{d^2}{dr^2} + \frac{3}{r}\frac{d}{dr} - \mu^2\right)(\bar{\phi} - \phi_-) = 0\,, \tag{14.158}$$

where $\mu^2 = U''(\phi_-)$. This equation can be solved rather easily [15], and its solution is expressed via the modified Bessel function. Then we see that choosing $\bar{\phi}(0)$ close enough to ϕ_-, we can also guarantee that for large-enough r the particle will remain as close as possible to ϕ_-. But for large-enough r we can neglect friction, as it is $\sim 1/r$. But in absence of friction the particle will overshoot the top at ϕ_+. This means that in our problem there is always an intermediate point (between ϕ_- and ϕ_0), starting from which the particle will at $r \to \infty$ stop at ϕ_+.

Let $U_+(\phi)$ be some even function of ϕ:

$$U_+(\phi) = U_+(-\phi) \tag{14.159}$$

with minima at points $\pm a$:

$$U'_+(\pm a) = 0\,. \tag{14.160}$$

Let us define

$$\mu^2 = U''_+(\pm a)\,. \tag{14.161}$$

Let us add to U_+ a small term, breaking the symmetry between the minima:

$$U = U_+ + \varepsilon(\phi - a)/2a\,, \quad \varepsilon > 0\,. \tag{14.162}$$

In the first order in ε we have

$$\phi_\pm = \pm a\,. \tag{14.163}$$

The value of ε defines the energy difference between the "true" and "false" vacuum. Let us choose the initial position of the particle $\bar{\phi}(0)$ very close to ϕ_-. Then the particle remains close to ϕ_- up to some large moment of time $r = R$; then afterwards it rapidly passes through the valley and slowly approaches ϕ_+ for $r \to \infty$. Thus, our instanton looks like a large four-dimensional spherically symmetric "bubble" of radius R with a thin wall separating the "false" vacuum ϕ_+ (outside the bubble) from the "true" vacuum ϕ_- (inside the bubble). Correspondingly, our bubble (instanton) represents the nuclei of a new ("true") vacuum inside the metastable ("false") vacuum.

For $r \sim R$ we can neglect the friction as well as the ε-dependent term in U. Then the equation of motion has the form

$$\frac{d^2\bar{\phi}}{dr^2} = U'_+(\bar{\phi})\,, \tag{14.164}$$

which corresponds to the classical equation of motion of a particle in a double-well potential, which was analyzed in details above. This equation has as its solution the sim-

plest one-dimensional instanton of Figure 14.9, which we studied above (and which describes the transition from $-a$ to $+a$ at "moment" R with the growth of r). This is the approximate description of an instanton in our field problem.

Up to now we have not yet defined R. The action of the instanton is given by

$$S = 2\pi^2 \int_0^\infty dr\, r^3 \left[\frac{1}{2} \left(\frac{d\bar{\phi}}{dr} \right)^2 + U(\bar{\phi}) \right]. \tag{14.165}$$

Here we have three regions of integration: outside the bubble, close to its surface, and inside. Outside we can take $\bar{\phi} = \phi_+$ and $U = 0$, so that this contribution to the integral is just zero (which actually guarantees the finiteness of the instanton action). Inside the bubble we have $\bar{\phi} = \phi_-$, $U = -\varepsilon$, so that the corresponding contribution to the integral is

$$-\frac{1}{2}\pi^2 R^4 \varepsilon\,. \tag{14.166}$$

Close to the bubble surface, i. e., for $r \sim R$, we can neglect $\sim\varepsilon$ term in U, so that the integral reduces to

$$2\pi^2 R^3 \int dr \left[\frac{1}{2} \left(\frac{d\bar{\phi}}{dr} \right)^2 + U_+ \right] = 2\pi^2 R^3 S_1\,, \tag{14.167}$$

where

$$S_1 = \int_{-a}^{a} d\phi \sqrt{2U_+} \tag{14.168}$$

is the action of the one-dimensional instanton (14.119). Finally we get

$$S = -\frac{1}{2}\pi^2 R^4 \varepsilon + 2\pi^2 R^3 S_1\,. \tag{14.169}$$

Let us now define R from the requirement of extremal action:

$$\frac{dS}{dR} = -2\pi^2 R^3 \varepsilon + 6\pi^2 R^2 S_1 = 0\,, \tag{14.170}$$

which gives

$$R = \frac{3S_1}{\varepsilon}\,. \tag{14.171}$$

Then we have[8]

$$S_0 = \frac{27\pi^2 S_1^4}{2\varepsilon^3}\,. \tag{14.172}$$

The bubble radius (14.171) can be found from the elementary considerations used in nucleation theory of statistical mechanics: the energy gain within the bubble should compensate the energy loss, connected with the surface tension of the bubble:

$$\frac{4}{3}\pi R^3 \varepsilon = 4\pi R^2 \sigma\,, \quad \text{which gives } R = \frac{3\sigma}{\varepsilon}\,, \tag{14.173}$$

where σ is the surface energy of the bubble well. In our case $\sigma = S_1$.

[8] Our analysis is valid in the limit of small ε and in the limit when the bubble radius is much larger than the width of its wall: $R \gg \mu^{-1}$, which reduces to $3S_1\mu \gg \varepsilon$.

Finally, we obtain the probability of "false" vacuum decay as

$$\frac{\Gamma}{V} \sim \exp(-S_0)\,. \tag{14.174}$$

Determinants and renormalization

The preexponential factor in (14.174) should be defined in the same way as in the quantum mechanical problem discussed above. But there are some important differences and questions:

1. In quantum mechanics we had only one zero translational mode; here there are four.
2. It was very important that there was only one negative energy eigenvalue, leading to an imaginary contribution. Is this also the case in the present problem?
3. In quantum field theory we have ultraviolet divergences, and it is necessary to perform renormalization. What is the role of renormalization here?

Consider first the zero modes. Here we have four directions for instanton translation (the instanton can be place at an arbitrary point of four-dimensional Euclidean space); correspondingly we have four eigenfunctions of a differential operator, related to the second variational derivative of action, with zero eigenstates. These functions $\sim \partial_\mu \bar{\phi}$. The normalization condition reduces to

$$\int d^4x \partial_\mu \bar{\phi} \partial_\nu \bar{\phi} = \frac{1}{4}\delta_{\mu\nu} \int d^4x \partial_\mu \bar{\phi} \partial^\mu \bar{\phi} = \delta_{\mu\nu} S_0\,. \tag{14.175}$$

As a result, the preexponential in (14.174) contains four factors of $\left(\frac{S_0}{2\pi}\right)^{1/2}$ instead of one.

The proof of the last equality in (14.175) goes as follows. Consider $\phi_\lambda(x) = \bar{\phi}(x/\lambda)$. Then the action is

$$S(\phi_\lambda) = \frac{1}{2}\lambda^2 \int d^4x (\partial_\mu \bar{\phi})^2 + \lambda^4 \int d^4x U(\bar{\phi})\,. \tag{14.176}$$

As $\bar{\phi}$ is the solution of equations of motion, we should satisfy the condition of stationarity of the action (14.176) at $\lambda = 1$. This yields

$$\int d^4x (\partial_\mu \bar{\phi})^2 = -4 \int d^4x U(\bar{\phi}) \tag{14.177}$$

or

$$S_0 = \frac{1}{4} \int d^4x (\partial_\mu \bar{\phi})^2 > 0\,. \tag{14.178}$$

Finally, we obtain the preexponential factor as

$$K = \frac{S_0^2}{4\pi^2} \left| \frac{\mathrm{Det}'[-\partial_\mu \partial_\mu + U''(\bar{\phi})]}{\mathrm{Det}[-\partial_\mu \partial_\mu + U''(\phi_+)]} \right|^{-1/2}, \tag{14.179}$$

assuming there are no problems with negative eigenvalues and renormalization.

As to negative eigenvalues, this is really so. It is clear that $\frac{\delta^2 S}{\delta \phi^2}$ (at the instanton) has at least one negative eigenstate. It can be *rigorously* proved that there is only one negative eigenstate in this problem [15]. Thus, equation (14.179) gives the correct probability of vacuum decay.

We shall not discuss in detail the problem of the renormalization of (14.174). In principle, it is clear that in theories with renormalizable $U(\varphi)$, all expressions, including (14.179), can be rewritten via renormalized parameters, and everything should be finite. Some additional details can be found in [15].

The bubble, expanding in real Minkowski space-time, can be obtained as the analytical continuation of the instanton:

$$\phi(x_0, \mathbf{x}) = \bar{\phi}(r = \sqrt{|\mathbf{x}|^2 - x_0^2}) \,. \tag{14.180}$$

Thus, at small ε we have a thin wall at $r = R$, separating the two vacuums, and expansion of the bubble is determined by

$$|\mathbf{x}|^2 - x_0^2 \approx R^2 \,. \tag{14.181}$$

The value of R is determined, as we have seen, by the microscopic parameters of the theory and is itself microscopic. Then, equation (14.181) means that the expanding surface of the bubble moves practically with the speed of light ($v \sim 1$)! The wall transports the energy (per unit surface) $\frac{S_1}{\sqrt{1-v^2}}$. At the moment, when the bubble radius reaches $|\mathbf{x}|$, the wall energy becomes

$$E_{wall} = \frac{4\pi |\mathbf{x}|^2 S_1}{\sqrt{1 - v^2}} \,. \tag{14.182}$$

From (14.181) it is easy to find that

$$v = \frac{d\,|\mathbf{x}|}{dt} = \sqrt{1 - \frac{R^2}{|\mathbf{x}|^2}} \,. \tag{14.183}$$

Then the wall energy is

$$E_{wall} = \frac{4\pi |\mathbf{x}|^3 S_1}{R} = \frac{4\pi \varepsilon |\mathbf{x}|^3}{3} \,, \tag{14.184}$$

so that practically the whole energy released during the "false" vacuum decay goes to the wall acceleration. No particles are created; from both sides of the wall we have the corresponding vacuum states. In this sense, the "observer" will never know that the wall passed through him; he will also just "decay" in the corresponding microscopic time.

Examples of applications of this formalism to problems of relativistic cosmology can be found in [41].

The concept of instantons plays a major role in many problems of quantum field theory and statistical mechanics. As an example, we can again mention the $g\varphi^4$ with $g < 0$ and the number of filed components $n = 0$, which describes the motion of an electron in the random field of impurities in solid state theory. In this model with an unstable ground state there are t instanton solutions which determine the electron density of the states in the so-called "tail" region, appearing due to electron localization by random field fluctuations [57, 64].

Especially important are nontrivial instanton solutions in non-Abelian gauge theories, which are related to the topological properties of gauge transformations and the complicated structure of a Yang–Mills vacuum [15,51,52]. We shall not discuss these aspects of the theory here, as well as their importance for particle physics (QCD). A detailed presentation can be found in [52, 58].

14.5 The Lipatov asymptotics of a perturbation series

At the end of Chapter 8 we briefly discussed the asymptotic nature of a perturbation series in quantum field theory. Here we shall consider it in more detail, describing the elements of the elegant approach proposed by Lipatov. The idea of the Lipatov method is as follows. If we have some physical function $F(g)$ which is expanded in a perturbation series in powers of coupling constant g

$$F(g) = \sum_{N=0}^{\infty} F_N g^N, \tag{14.185}$$

the coefficients of this expansion F_N can be determined as

$$F_N = \int_C \frac{dg}{2\pi i} \frac{F(g)}{g^{N+1}}, \tag{14.186}$$

where integration contour C encircles the point $g = 0$ in the complex plane of the coupling constant. Rewriting the denominator here as $\exp\{-(N+1)\ln g\}$, for large N we can use the steepest-descent (stationary phase) approach to estimate (14.186).

We know that all problems solved by the diagram technique can be reformulated in terms of functional integrals like

$$Z(g) = \int D\varphi \, \exp\left(-S_0\{\varphi\} - gS_{int}\{\varphi\}\right), \tag{14.187}$$

and we can write the coefficients of perturbation expansion as

$$Z_N = \int_C \frac{dg}{2\pi i g} \int D\varphi \, \exp\left(-S_0\{\varphi\} - gS_{int}\{\varphi\} - N\ln g\right). \tag{14.188}$$

The Lipatov idea is to search for the steepest descent in (14.188) not simply over g, but over g and φ simultaneously:

$$\left.\frac{\delta S(\varphi)}{\delta \varphi}\right|_{\varphi_c} = 0\,, \tag{14.189}$$

$$\frac{S(\varphi_c)}{g_c} = 0\,. \tag{14.190}$$

The solution of these equations exists for all interesting models and is realized on a spatially localized *instanton* $\varphi_c(x)$. The steepest descent approach is applicable here for large N, independent of its applicability to the initial functional integral (14.187). This fact is of prime importance; in the general case an exact calculation of the functional integrals is impossible, but they are easily calculated by steepest descent.

This allows us to determine the general form of large N asymptotics of the perturbation theory coefficients for any physical characteristics (such as Green's functions, vertex parts, etc.) for different models of quantum field theory. The typical form of Lipatov asymptotics for the perturbation coefficients of an arbitrary function $F(g)$ has the form

$$F_N = c\,\Gamma(N+b)a^N\,, \tag{14.191}$$

where $\Gamma(x)$ is the Γ-function, and parameters $a,\ b,\ c$ depend on the specific problem under discussion. In a concrete model of field theory the constant a is universal, the parameter b depends on the physical function $F(g)$ under study, and c contains dependence on external momenta (or coordinates). The appearance of $\Gamma(N+b) \sim N!$ in (14.191) simply reflects the factorial growth of the *number* of diagrams with the order N of perturbation theory. Obviously, such asymptotic behavior of perturbation theory coefficients corresponds to the divergent series!

The knowledge of Lipatov asymptotics in combination with the exact results for a few lowest orders of perturbation theory, obtained by direct diagrammatic calculations, gives information on the perturbation series as a whole. Approximating the complete series by the sum of lowest order contributions with asymptotics of higher orders, and applying the mathematical methods of the summation of the divergent series, we can obtain approximate solutions of an arbitrary physical problem.

The most common method to deal with a divergent (asymptotic) series of perturbation theory is to use so-called Borel transformation. We can divide and multiply each term of the series by $N!$ and use the integral representation of the Γ-function, so that after the interchange of summation and integration, we can write

$$F(g) = \sum_{N=0}^{\infty} F_N g^N = \sum_{N=0}^{\infty} \frac{F_N}{N!} \int_0^{\infty} dx\, x^N \mathrm{e}^{-x} g^N = \int_0^{\infty} dx\, \mathrm{e}^{-x} \sum_{N=0}^{\infty} \frac{F_N}{N!} (gx)^N. \tag{14.192}$$

The power series in the right-hand side is in most cases converging (it actually has factorially improved convergence!) and defines Borel transform $B(z)$ of the function

$F(g)$, which can now be determined from the following integral transformation:

$$F(g) = \int_0^\infty dx\, \mathrm{e}^{-x} B(gx)\,, \quad B(z) = \sum_{N=0}^{\infty} \frac{F_N}{N!} z^N\,, \tag{14.193}$$

Borel transformation gives the natural method of summation of a factorially divergent perturbation series of quantum field theory[9].

14.6 The end of the "zero-charge" story?

In Chapter 13 we stressed the importance of the asymptotic behavior of the Gell-Mann–Low function $\beta(g)$ at large values of the coupling constant g for the internal consistency of quantum field theory. However, until recently, only perturbation theory estimates of $\beta(g)$ were available, and no definite conclusions on its behavior at large g could be drawn. Below we shall present some nonperturbative arguments due to Suslov, allowing us to find this asymptotic behavior in analytic form [66].

For simplicity we shall consider the $O(N)$ symmetric Euclidean φ^4 theory in d-dimensional space with the action[10]

$$S\{\varphi\} = \int d^d x \left\{ \tfrac{1}{2} \sum_{\alpha=1}^{N} (\nabla\varphi_\alpha)^2 + \tfrac{1}{2} m_0^2 \sum_{\alpha=1}^{N} \varphi_\alpha^2 + \tfrac{1}{8} u \Big(\sum_{\alpha=1}^{N} \varphi_\alpha^2 \Big)^2 \right\}, \tag{14.194}$$

where $u = g_0 \Lambda^\epsilon$ and $\epsilon = 4 - d$. Actually, this is the direct analogue of equation (10.160) used in the theory of critical phenomena. Here we are using lattice regularization of ultraviolet divergences, introducing the cut-off $\Lambda \sim a^{-1}$, where a is the lattice constant. Following the usual renormalization group formalism, we consider the "amputated" vertex $\Gamma^{(n)}$ with n external lines of field φ. The multiplicative renormalizability of the theory means that we may write the direct analogue of equation (13.76) as

$$\Gamma^{(n)}(p_i; g_0, m_0, \Lambda) = Z^{-n/2} \Gamma_R^{(n)}(p_i; g, m)\,, \tag{14.195}$$

so that divergence at $\Lambda \to \infty$ disappears after extraction of the proper Z-factors and their transfer to the renormalized charge and mass, which are denoted here as g and m. We shall accept the renormalization conditions at zero momentum:

$$\begin{aligned} \Gamma_R^{(2)}(p; g, m)\Big|_{p\to 0} &= m^2 + p^2 + O(p^4)\,, \\ \Gamma_R^{(4)}(p_i; g, m)\Big|_{p_i = 0} &= g m^\epsilon\,, \end{aligned} \tag{14.196}$$

[9] A detailed discussion of methods to deal with divergent series of perturbation theory can be found in the review paper [65]

[10] Generalization to QED is more or less straightforward.

which are typical for applications in the phase transitions theory. From equations (14.196) and (14.195) we can obtain expressions for renormalized g, m, Z in terms of the "bare" quantities:

$$\begin{aligned} Z(g_0, m_0, \Lambda) &= \left(\frac{\partial}{\partial p^2} \, \Gamma^{(2)}(p; g_0, m_0, \Lambda) \Big|_{p=0} \right)^{-1} , \\ m^2 &= Z(g_0, m_0, \Lambda) \, \Gamma^{(2)}(p; g_0, m_0, \Lambda) \Big|_{p=0} , \\ g &= m^{-\epsilon} Z^2(g_0, m_0, \Lambda) \, \Gamma^{(4)}(p_i; g_0, m_0, \Lambda) \Big|_{p_i=0} . \end{aligned} \tag{14.197}$$

Applying the differential operator $d/d \ln m$ to (14.195) for fixed g_0 and Λ gives the direct equivalent of the Callan–Symanzik equation (13.82), which for large momenta $|p_i|/m \gg 1$ has the form

$$\left[\frac{\partial}{\partial \ln m} + \beta(g) \frac{\partial}{\partial g} - n\gamma(g) \right] \Gamma_R^{(n)}(p_i; g, m) \approx 0 , \tag{14.198}$$

where functions $\beta(g)$ and $\gamma(g)$ are defined as

$$\beta(g) = \left. \frac{dg}{d \ln m} \right|_{g_0, \Lambda = \text{const}} , \quad \gamma(g) = \left. \frac{d \ln \sqrt{Z}}{d \ln m} \right|_{g_0, \Lambda = \text{const}} , \tag{14.199}$$

and according to the general theorems depend only on g.

Now we shall show how the renormalization group functions are expressed via functional integrals. The functional integrals of φ^4-theory are determined as

$$Z^{(M)}_{\alpha_1 \ldots \alpha_M}(x_1, \ldots, x_M) = \int D\varphi \, \varphi_{\alpha_1}(x_1) \varphi_{\alpha_2}(x_2) \cdots \varphi_{\alpha_M}(x_M) \exp\left(-S\{\varphi\}\right) . \tag{14.200}$$

Fourier transform of equation (14.200) can be written as

$$\begin{aligned} & Z^{(M)}_{\alpha_1 \ldots \alpha_M}(p_1, \ldots, p_M) \mathcal{N} \delta_{p_1 + \cdots + p_M} \\ & \quad = \sum_{x_1, \ldots, x_M} Z^{(M)}_{\alpha_1 \ldots \alpha_M}(x_1, \ldots, x_M) \mathrm{e}^{i p_1 x_1 + \ldots + i p_M x_M} \\ & \quad = K_M(p_i) I_{\alpha_1 \ldots \alpha_M} \delta_{p_1 + \cdots + p_M} \end{aligned} \tag{14.201}$$

where $\mathcal{N}$ is the number of sites on the lattice, which is implied in the definition of the (regularized) functional integral, and symmetry factors $I_{\alpha_1 \cdots \alpha_M}$ are similar to those discussed in Chapter 10 in relation to critical phenomena. Now we have

$$Z^{(0)} = K_0 , \quad Z^{(2)}_{\alpha\beta}(p, -p) = K_2(p) \delta_{\alpha\beta} , \quad Z^{(4)}_{\alpha\beta\gamma\delta}\{p_i\} = K_4\{p_i\} I_{\alpha\beta\gamma\delta} \tag{14.202}$$

where $I_{\alpha\beta\gamma\delta}$ is given by an expression similar to that in equation (10.168). Now we can introduce the vertex part $\Gamma^{(4)}$ by the usual relation for two-particle (4-point) Green's

function:

$$G^{(4)}_{\alpha\beta\gamma\delta}(p_1,\dots,p_4) = G^{(2)}_{\alpha\beta}(p_1)G^{(2)}_{\gamma\delta}(p_3)\,\mathcal{N}\delta_{p_1+p_2} + G^{(2)}_{\alpha\gamma}(p_1)G^{(2)}_{\beta\delta}(p_2)\,\mathcal{N}\delta_{p_1+p_3} + G^{(2)}_{\alpha\delta}(p_1)G^{(2)}_{\beta\gamma}(p_3)\,\mathcal{N}\delta_{p_1+p_4} - G^{(2)}_{\alpha\alpha'}(p_1)G^{(2)}_{\beta\beta'}(p_2)G^{(2)}_{\gamma\gamma'}(p_3)G^{(2)}_{\delta\delta'}(p_4)\Gamma^{(4)}_{\alpha'\beta'\gamma'\delta'}(p_1,\dots,p_4) \tag{14.203}$$

where $G^{(2)}_{\alpha\beta}(p_i)$ are single-particle (2-point) Green' functions. Extracting factors $I_{\alpha_1\dots\alpha_M}$ we have

$$G^{(2)}_{\alpha\beta}(p) = G_2(p)\delta_{\alpha\beta}\,, \quad G^{(4)}_{\alpha\beta\gamma\delta}\{p_i\} = G_4\{p_i\}I_{\alpha\beta\gamma\delta}\,, \quad \Gamma^{(4)}_{\alpha\beta\gamma\delta}\{p_i\} = \Gamma_4\{p_i\}I_{\alpha\beta\gamma\delta} \tag{14.204}$$

Now we can write

$$G_4 = \frac{K_4}{K_0}\,, \quad \Gamma_4 = -\frac{G_4}{G_2^4} = -\frac{K_4K_0^3}{K_2^4}\,, \tag{14.205}$$

and

$$G_2 = \frac{K_2(p)}{K_0}\,, \quad \Gamma_2(p) = \frac{1}{G_2(p)} = \frac{K_0}{K_2(p)} \approx \frac{K_0}{K_2} + \frac{K_0\tilde{K}_2}{K_2^2}\,p^2\,, \tag{14.206}$$

where for small p we have written

$$K_2(p) = K_2 - \tilde{K}_2 p^2 + \cdots\,. \tag{14.207}$$

Expressions for the Z-factors, renormalized mass, and charge follow from (14.197):

$$Z = \left[\frac{\partial}{\partial p^2}\Gamma_2(p)\right]^{-1}_{p=0} = \frac{K_2^2}{K_0\tilde{K}_2}\,, \tag{14.208}$$

$$m^2 = Z\Gamma_2(0) = \frac{K_2}{\tilde{K}_2}\,, \tag{14.209}$$

$$g = m^{-\epsilon}Z^2\Gamma_4 = -\left(\frac{K_2}{\tilde{K}_2}\right)^{d/2}\frac{K_4K_0}{K_2^2}\,, \tag{14.210}$$

and

$$\frac{dm^2}{dm_0^2} = \left(\frac{K_2}{\tilde{K}_2}\right)' = \frac{K_2'\tilde{K}_2 - K_2\tilde{K}_2'}{\tilde{K}_2^2}\,, \tag{14.211}$$

where the prime denotes the derivatives over m_0^2. Parameters g_0 and Λ are considered to be fixed, while m^2 is a function of m_0^2 only and derivative dm_0^2/dm^2 is defined by

the expression inverse to (14.211). Using the definitions (14.199) we have

$$\beta(g) = \left(\frac{K_2}{\tilde{K}_2}\right)^{d/2} \left\{ -d\,\frac{K_4 K_0}{K_2^2} + 2\frac{(K_4' K_0 + K_4 K_0')K_2 - 2K_4 K_0 K_2'}{K_2^2}\,\frac{\tilde{K}_2}{K_2\tilde{K}_2' - K_2'\tilde{K}_2} \right\} \tag{14.212}$$

$$\gamma(g) = -\frac{K_2\tilde{K}_2}{K_2\tilde{K}_2' - K_2'\tilde{K}_2} \left[2\frac{K_2'}{K_2} - \frac{K_0'}{K_0} - \frac{\tilde{K}_2'}{\tilde{K}_2} \right] \tag{14.213}$$

These equations determine $\beta(g)$ and $\gamma(g)$ in parametric form: for fixed g_0 and Λ, the right-hand side of these equations are functions of m_0^2 only, while dependence on the specific choice of g_0 and Λ is absent due to general theorems.

Any infinities in the right-hand sides of equations (14.212) and (14.213) can be induced only by the *zeroes of functional integrals*[11]. It is clear from equation (14.210) that the limit $g \to \infty$ can be achieved by two ways: tending to zero either K_2 or $\tilde{K}_2$. For $\tilde{K}_2 \to 0$ equations (14.210) and equations (14.212), (14.213) give

$$g = -\left(\frac{K_2}{\tilde{K}_2}\right)^{d/2} \frac{K_4 K_0}{K_2^2}\,, \quad \beta(g) = -d\left(\frac{K_2}{\tilde{K}_2}\right)^{d/2} \frac{K_4 K_0}{K_2^2}\,, \quad \gamma(g) \to 1\,, \tag{14.214}$$

and the parametric representation is resolved as

$$\beta(g) = dg\,, \quad \gamma(g) = 1\,, \quad (g \to \infty)\,. \tag{14.215}$$

For $K_2 \to 0$, the limit of $g \to \infty$ can be achieved only for $d < 4$:

$$\beta(g) = (d-4)g\,, \quad \gamma(g) \to 2 \quad (g \to \infty)\,. \tag{14.216}$$

The results (14.215) and (14.216) probably correspond to different branches of the function $\beta(g)$. It is easy to understand that the physical branch is the first one. Indeed, it is commonly accepted in phase transitions theory that the properties of φ^4-theory change smoothly as a function of space dimension, and the results for $d = 2, 3$ can be obtained by analytic continuation from $d = 4 - \epsilon$. All the available information indicates the positivity of $\beta(g)$ for $d = 4$, and consequently its asymptotics at $g \to \infty$ is also positive. The same property is expected for $d < 4$ by continuity. The result (14.215) does obey such a property, while the branch (14.216) does not exist for $d = 4$ at all.

According to our discussion in Chapter 13, the behavior of the Gell-Mann–Low function given by equation (14.215) corresponds to the continuous growth of the renormalized charge as we go to the region of strong coupling at small distances, and signifies the consistency of quantum field theory without "pathologies" like a Landau

[11] This is the most nontrivial moment of our discussion. Actually, it can be shown that zeroes of the functional integrals can be obtained by a rather subtle compensation of the contributions of the trivial vacuum and some instanton configuration with finite action.

“ghost pole” (or a “zero-charge” problem). However, it should be clearly understood, that during our discussion here we have skipped many subtle details, which are to be looked for in original papers, as well as the difficulties which are making this point-of-view less than commonly accepted.

Bibliography

[1] A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskii, *Quantum Field Theoretical Methods in Statistical Physics*, Pergamon Press, Oxford, 1965.

[2] A. I. Akhiezer and V. B. Berestetskii, *Quantum Electrodynamics*, Interscience Publishers, New York, 1965.

[3] D. Amit, *Field Theory, the Renormalization Group and Critical Phenomena*, McGraw-Hill, New York, 1978.

[4] S. Aoki et al., Quenched Light Hadron Spectrumm, *Phys. Rev. Lett.* **84** (2000), 238.

[5] V. B. Berestetskii, *Problems of Physics of Elementary Particles*, Nauka, Moscow, 1979 (in Russian).

[6] V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Quantum Electrodynamics*, Pergamon Press, Oxford, 1982.

[7] F. A. Berezin, *The Method of Second Quantization*, Academic Press, New York, 1966.

[8] J. Bjorken and S. Drell, *Relativistic Quantum Mechanics. Relativistic Quantum Fields* in two volumes, Mc Graw-Hill, New York, 1964, 1965.

[9] N. N. Bogolyubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields*, Interscience Publishers, New York, 1959.

[10] N. N. Bogolyubov and D. V. Shirkov, *Quantum Fields*, Addison-Wesley, New York, 1982.

[11] V. G. Bornyakov, M. I. Polikarpov, T. Suzuki, M. N. Chernodub, and G. Schierholz, *Usp. Fiz. Nauk* **174** (2004), 19.

[12] R. N. Cahn, *Rev. Mod. Phys.* /1996) **68**, 951.

[13] T.-P. Cheng and L.-F. Li, *Gauge Theory of Elementary Particle Physics*, Clarendon Press, Oxford, 1984.

[14] G. F. Chew, *The Analytic S-Matrix: A Basis for Nuclear Democracy*, W. A. Benjamin, New York, 1966.

[15] S. Coleman, The Uses of Instantons, in: A. Zichichi (ed.), *The Whys of Subnuclear Physics*, Plenum Press, New York, 1979.

[16] A. D. Dolgov, Y .B. Zel'dovich, and M. V. Sazhin, *Basics of Modern Cosmology*, Editions Frontieres, Gif sur Yvette Cedex, 1990.

[17] I. T. Dyatlov, V. V. Sudakov, and K. A. Ter-Martirosyan, Asymptotic theory of meson-meson scattering, *Zh. Eksp. Teor. Fiz. (JETP)* **32** (1957), 767.

[18] R. P. Feynman, Asymptotic theory of meson-meson scattering, *Quantum Electrodynamics*, W. A. Benjamin, New York, 1961.

[19] R. P. Feynman, *Statistical Mechanics*, W. A. Benjamin, New York, 1972.

[20] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, Mc Graw-Hill, New York, 1965.

[21] M. E. Fisher, *The Nature of Critical Points*, University of Colorado Press, Boulder, 1965.

[22] M. K. Gaillard, P. D. Grannis, and F. J. Sciulli, *Rev. Mod. Phys.* (1999) **71**, 96.

[23] K. Gottfried and V. F. Weisskopf, *Concepts of Particle Physics*, Clarendon Press, Oxford, 1984.

[24] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, John Wiley and Sons, New York, 1984.

[25] K. Huang, *Quarks, Leptons and Gauge Fields*, World Scientific, Singapore, 1982.

[26] K. Huang, *Fundamental Forces of Nature. The Story of Gauge Fields*, World Scientific, Singapore, 2007.

[27] J. Illiopoulos, C. Itzykson, and A. Martin, Functional methods and perturbation theory, *Rev. Mod. Phys.* **47** (1975), 165.

[28] C. Itzykson and J. Zuber, *Quantum Field Theory*, McGraw-Hill, New York, 1980.

[29] G. Kane, *Modern Elementary Particle Physics*, Addison-Wesley, New York, 1990.

[30] J. B. Kogut, An introduction to lattice gauge theory and spin systems, *Rev. Mod. Phys.* **51** (1979), 659.

[31] M. Kreutz, *Quarks, Gluons and Lattices*, Cambridge University Press, Cambridge, 1983.

[32] A. Lambrecht, The Casimir effect: a force from nothing, *Physics World* **15**(9) (2002), 29.

[33] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, Pergamon, Oxford, 1962.

[34] L. D. Landau and E. M. Lifshitz, *Mechanics*, Pergamon Press, Oxford, 1976.

[35] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, Pergamon Press, Oxford, 1977.

[36] L. D. Landau and E. M. Lifshitz, *Statistical Physics, Part I*, Pergamon Press, Oxford, 1980.

[37] L. D. Landau and E. M. Lifshitz, *Statistical Physics, Part II*, Pergamon Press, Oxford, 1980.

[38] A. I. Larkin and D. E. Khmelnitskii, Phase Transition in Uniaxial Ferroelectrics *Zh. Eksp. Teor. Phys. (JETP)* **56** (1969), 2087.

[39] T. D. Lee and C. N. Yang, Conservation of Heavy Particles and Generalized Gauge Transformations, *Phys. Rev.* **98** (1955), 1501.

[40] E. M. Lifshitz and L. P. Pitaevskii, *Relativistic Quantum Field Theory, Part 2*, Pergamon Press, Oxford, 1974.

[41] A. D. Linde, *Particle Physics and Inflationary Cosmology*, Harwood, Chur, 1990.

[42] S.-K. Ma, *Modern Theory of Critical Phenomena*, W. A. Benjamin, New York, 1976.

[43] Y. M. Makeenko, The Monte Carlo method in lattice gauge theories, *Usp. Fiz. Nauk* **143** (1984), 161.

[44] M. B. Mensky, Measurability of quantum fields and energy-time uncertainty relation,*Usp. Fiz. Nauk* **181** (2011), 543.

[45] H. Meyer-Ortmanns, Phase transitions in quantum chromodynamics, *Rev. Mod. Phys.* **68** (1996), 473.

[46] L. B. Okun, *Particle Physics. The Quest for the Substance of Substance*, Harwood, Chur, 1985.

[47] L. B. Okun, Current status of elementary particle physics, *Usp. Fiz. Nauk* **168** (1998), 625.

[48] A. Z. Patashinskii and V. L. Pokrovskii, *Fluctuation Theory of Phase Transitions*, Pergamon Press, Oxford, 1979.

[49] W. Pauli (ed.), *Niels Bohr and the Development of Physics*, Pergamon Press, London, 1955.

[50] D. H. Perkins, *Introduction to High Energy Physics*, Addison-Wesley, New York, 1987.

[51] A. M. Polyakov, *Gauge Fields and Strings*, Harwood, Chur, 1987.

[52] P. Rajaraman, *Solitons and Instantons*, North Holland, Amsterdam, 1982.

[53] P. Ramond, *Field Theory: A Modern Primer*, Benjamin/Cummings, London, 1982.

[54] K. Rajagopal and F. Wilczek, *The Condensed Matter Physics of QCD*, ArXiv: hep-ph/0011333.

[55] V. A. Rubakov, On Large Hadron Collider's discovery of a new particle with Higgs boson properties, *Usp. Fiz. Nauk* **182** (2012), 1017.

[56] L. Ryder, *Quantum Field Theory*, Cambridge University Press, Cambridge, 1985.

[57] M. V. Sadovskii, Electron localization in disordered systems: critical behavior and macroscopic manifestations, *Usp. Fiz. Nauk* **133** (1981), 223.

[58] T. Schäfer and E. V. Shuryak, Instantons in QCD, *Rev. Mod. Phys.* **70** (1998), 323.

[59] S. Schweber, *An Introduction to Relativistic Quantum Field Theory*, Row, Peterson and Co., New York, 1961.

[60] S. Schweber, H. Bethe, and F. de Hoffmann, *Mesons and Fields, vol. I*, Row, Peterson and Co., New York, 1955.

[61] Y. A. Simonov, The confinement, *Usp. Fiz. Nauk* **166** (1996), 337.

[62] A. V. Smilga, *Lectures on the Foundations of QCD*, ArXiv: hep-ph/9901412.

[63] H. Stanley, *Introduction to Phase Transitions and Critical Phenomena*, Clarendon Press, Oxford, 1971.

[64] I. M. Suslov, Divergent perturbation series, *Usp. Fiz. Nauk* **168** (1998), 503.

[65] I. M. Suslov, Development of (4-ϵ)-dimensional theory for the density of states of a disordered system near the Anderson transition, *Zh. Eksp. Teor. Fiz. (JETP)* **127** (2005), 1350.

[66] I. M. Suslov, Strong coupling asymptotics of the β-function in φ^4 theory and QCD, *Applied Numerical Mathematics* **60** 2010), 1418.

[67] S. V. Troitskii, Unsolved problems in particle physics, *Usp. Fiz. Nauk* **182** (2012), 77.

[68] R. Utiyama, Invariant Theoretical Interpretation of Interaction, *Phys. Rev.* **101** (1956), 1597.

[69] A. N. Vasil'ev, *The Field Theoretic Renormalization Group in Critical Behavior Theory and Stochastic Dynamics*, Chapman and Hall/CRC, London, 2004.

[70] V. S. Vladimirov, *Equations of Mathematical Physics*, Mir, Moscow, 1984.

[71] E. T. Whittaker and G. N. Watson. *A Course of Modern Analysis*. Cambridge University Press, Cambridge, 1927.

[72] F. Wilczek, *QCD in Extreme Conditions*, ArXiv: hep-ph/0003183.

[73] F. Wilczek, *The Future of Particle Physics as a Natural Science*, ArXiv: hep-ph/9702371.

[74] F. Wilczek, QCD Made Simple, *Phys. Today* **53**(8) (2000), 2.

[75] C. N. Yang and R. L. Mills, Conservation of Isotopic Spin and Isotopic Gauge Invariance, *Phys. Rev.* **96** (1954), 191.

[76] V. I. Zakharov, B. L. Ioffe, and L. B. Okun, New elementary particles, *Usp. Fiz. Nauk* **117** (1975), 227.

[77] Y. B. Zel'dovich, Classification of elementary particles and quarks "presented for pedestrians", *Usp. Fiz. Nauk* **86** (1965), 303.

Index