Introduction to Quantum Field Theory with Applications to Quantum Gravity

Iosif L. Buchbinder
Ilya L. Shapiro

OXFORD GRADUATE TEXTS

Introduction to Quantum Field Theory with Applications to Quantum Gravity

Introduction to Quantum Field Theory with Applications to Quantum Gravity

Iosif L. Buchbinder

Department of Theoretical Physics, Tomsk State Pedagogical University, Tomsk, 634061, Russia

Ilya L. Shapiro

Departamento de Física – Instituto Ciências Exatas, Universidade Federal de Juiz de Fora, Juiz de Fora, CEP 36036-330, MG, Brazil

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

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

First Edition published in 2021
Impression: 1

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

British Library Cataloguing in Publication
Data available

Library of Congress Control Number: 2020945719

ISBN 978–0–19–883831–9

DOI: 10.1093/oso/9780198838319.001.0001

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

Preface

For many decades, quantum field theory has played an important role in the successful description of the interactions of elementary particles. Besides, this area of theoretical physics has been always important due to the exchange of new ideas and methods with other branches of physics, such as statistical mechanics, condensed matter physics, gravitational physics, and cosmology. The last applications are becoming more important nowadays, especially because the amount of experimental and observational data demonstrates a fast growth and requires more detailed and reliable theoretical background. One of the most evident examples is the study of dark energy. Every few years, the estimates of its equation of state (EoS) become more precise and it can not be ruled out that, at some point, the EoS of the cosmological constant may be excluded from the list of phenomenologically acceptable possibilities. Does this necessarily mean that there is some special fluid (quintessence or alike) in the Universe? Or that the situation can be explained by the variable cosmological constant, e.g., some quantum effects? This is a phenomenologically relevant question, which should be answered at some point. On the other hand, this is a theoretical question, that can be answered only within a correctly formulated framework of quantum or semiclassical gravity.

In gravitational theory, general relativity represents a successful theory of relativistic gravitational phenomena, confirmed by various experiments in the laboratories and astronomical observations. Starting from the seventies and eighties, there has been a growing interest in the idea of the unification of all fundamental forces, including electroweak and strong interactions. Also, there is a general understanding that the final theory should also include gravitation. An important component of such unification is the demand for a quantum description of the gravitational field itself or, at least, a consistent formulation of the quantum theory of matter fields on the classical gravitational background, called semiclassical gravity.

The application of quantum field theory methods to gravitational physics, in both semiclassical and full quantum frameworks, requires a careful formulation of the fundamental base of quantum theory, with special attention to such important issues as renormalization, the quantum theory of gauge theories and especially effective action formalism. The existing literature on these subjects includes numerous review papers and also many books, e.g., [172, 56, 80, 150, 199, 240]. At the same time, the experience of the present authors, after giving many courses on the subject worldwide, shows that there is a real need to have a textbook with a more elementary introduction to the subject. This situation was one of the main motivations for writing this book which ended up being much longer than originally planned.

The textbook consists of two parts. Part I is based on the one-semester course given by I.B. in many places, including the Tomsk State Pedagogical University and the Federal University of Juiz de Fora. It includes a detailed introduction to the general methods of quantum field theory, which are relevant for quantum gravity, including its semiclassical part. Part II is mainly based on the one-semester course given regularly by I.Sh. in the Federal University of Juiz de Fora and on the numerous mini-courses in many countries. We did not pretend to do the impossible, that is, produce a comprehensive course of quantum field theory or quantum gravity. Instead, our purpose was to give a sufficiently detailed introduction to the fundamental, basic notions and methods, which would enable the interested reader to understand at least part of the current literature on the subject and, in some cases, start original research work.

It is a pleasure for us to acknowledge the collaborations on various subjects discussed in this book with M. Asorey, R. Balbinot, E.V. Gorbar, A. Fabbri, J.C. Fabris, J.-A. Helaël-Neto, P.M. Lavrov, T.P. Netto, S.D. Odintsov, F.O. Salles and A.A. Starobinsky. We would like also to thank many colleagues, especially A.O. Barvinsky, A.S. Belyaev, E.S. Fradkin, V.P. Frolov, S.J. Gates, E.A. Ivanov, D.I. Kazakov, S.M. Kuzenko, O. Lechnetfeld, H. Osborn, B.A. Ovrut, N.G. Pletnev, K. Stelle, A.A. Tseytlin, I.V. Tyutin, and G.A. Vilkovisky for fruitful discussions of the problems of quantum field theory.

We are grateful to Guilherme H.S. Camargo, Eduardo A. dos Reis, and especially to Wagno Cesar e Silva for communicating to us misprints and corrections; and also to Andreza R. Rodrigues and Yackelin Z. R. López for typing certain parts of the manuscript, and to Vadim Zyubanov for valuable technical assistance.

The main work on Part I of the book was done during the long-term visit of J.B. to the Federal University of Juiz de Fora (UFJF). The authors are grateful to UFJF and especially to the Physics Department for providing both kind hospitality and the conditions for productive work during this visit. Throughout the preparations of the manuscript, the work of the authors has been supported by a special project APQ-01205-16 from the Fundação de Amparo á Pesquisa de Minas Gerais (FAPEMIG). On the top of that, the scientific activity of I.Sh. was partially supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/Brazil). The authors are also grateful to the Russian Ministry of Science and High Education and Russian Foundation for Basic Research for their long-term support of the Center of Theoretical Physics at the Tomsk State Pedagogical University.

Contents

Part I

Introduction to Quantum Field Theory

1
Introduction

1.1 What is quantum field theory, and some preliminary notes

Quantum field theory (QFT) is part of the broader field of theoretical physics and is the study of quantum effects in continuous physical systems called fields. One can say that quantum field theory represents the unification of quantum mechanics and classical field theory. Since a natural and consistent description of fundamental interactions can be achieved in the framework of special relativity, it is also true to say that relativistic quantum field theory represents the unification of quantum mechanics and special relativity.

The main application of quantum field theory is the description of elementary particles and their interactions. However, QFT has also extensive applications in other areas of physics, including cosmology. Furthermore, quantum field theory plays an important role in the theoretical condensed matter physics, especially in the description of ensembles of a large number of interacting particles. The progress made in the theory of superconductivity, the theory of phase transitions and other areas of condensed matter physics is characterized by the consistent use of quantum field theory methods, and vice versa.

The first part of this book is devoted to the basic notions and fundamental elements of modern QFT formalism. In the second part, we present an introduction to the QFT in curved space and quantum gravity, which are less developed and essentially more complicated subjects.

The reader will note that the style of the two parts is different. In almost all of Part I and in most of Part II, we tried to give a detailed presentation, so that the reader could easily reproduce all calculations. However, following this approach for the whole topic of quantum gravity would enormously increase the size of the book and make it less readable. For this reason, in some places we avoided giving full technical details and, instead, just provided references of papers or preprints where the reader can find intermediate formulas. The same approach concerns the selection of the material. Since we intended to write an introductory textbook, in Part II we gave only the *need-to-know* information about quantum gravity. For this reason, many advanced subjects were not included. In addition, in some cases, only qualitative discussion and minimal references have been provided.

1.2 The notion of a quantized field

The field $\phi(x) = \phi(t, \mathbf{x})$ is defined as a function of time t and the space coordinates, that form a three-dimensional vector, $\mathbf{x}$. It is assumed that the values of the space

Introduction to Quantum Field Theory with Applications to Quantum Gravity. Iosif L. Buchbinder and Ilya L. Shapiro, Oxford University Press (2021).
DOI: 10.1093/oso/9780198838319.003.0001

coordinates correspond to a bounded or unbounded domain of the three-dimensional space. From the physical point of view, the field $\phi(t, \mathbf{x})$ can be treated as a dynamical object with an infinite number of degrees of freedom, marked by the three-dimensional vector index $\mathbf{x}$.

The notion of a field naturally arises in the framework of special relativity. Since there exists a maximal speed of propagation for any type of interaction, the physical bodies separated by space intervals can not affect each other instantly. Therefore, there should be a physical object responsible for transmitting perturbation from one body to another. Such an object is a field that fills the space between the bodies and carries perturbation from one body to another. The simplest example is an electromagnetic field that carries interaction between electrically charged bodies.

Taking into account quantum mechanical universality, it is natural to assume that fields should be quantized, like any other physical system. This means that quantum states are given by wave functions, while dynamical variables are given by operators acting on wave functions. Thus, in quantum theory, a field becomes an operator $\hat{\phi}(t, \mathbf{x})$, which is called a field operator.

As we have mentioned (and will discuss in more detail later on), a field is a system with an infinite number of degrees of freedom. However, it turns out that the state of the quantum field can be described in terms of either particles or fields. It turns out that the quantum field is a physical notion that is most suitable for the description of systems with an arbitrary number of particles.

It is well known that, in relativistic theory, there is a relation between momentum $\mathbf{p}$ and the energy $\varepsilon = \varepsilon(\mathbf{p})$ of a free particle,

$$\varepsilon^2 \,=\, m^2c^4 + c^2\mathbf{p}^2, \tag{1.1}$$

where c is the speed of light, and m is the mass of the particle. If the field can describe particles, it must take into account the relation (1.1) between energy and momentum. Let us try to clarify how the relation (1.1) can be implemented for the field. Let $\hat{\phi}(t, \mathbf{x})$ be the field operator, associated with the free particle. We can write the expansion as a Fourier integral,

$$\hat{\phi}(t, \mathbf{x}) \,=\, \int d^3p\, d\varepsilon\, e^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{x}-\varepsilon t)}\, \hat{\phi}(\varepsilon, \mathbf{p})\,, \tag{1.2}$$

where $\hbar$ is the Planck constant. According to the standard interpretation, the vector $\mathbf{p}$ is treated as the momentum of a particle, and the quantity ε as the energy of the particle. Then, since, for each Fourier mode, ε and $\mathbf{p}$ are related by Eq. (1.1), the quantity ε under the integral (1.2) is not an independent variable but is a function of $\mathbf{p}$. In order to satisfy this condition, one can write

$$\hat{\phi}(\varepsilon, \mathbf{p}) \,=\, \delta(\varepsilon^2 - \varepsilon^2(\mathbf{p}))\hat{\phi}_*(\mathbf{p}), \tag{1.3}$$

where $\hat{\phi}_*(\mathbf{p})$ depends only on $\mathbf{p}$. As a result, we arrive at the representation

$$\hat{\phi}(t, \mathbf{x}) \,=\, \int d^3p\, d\varepsilon\, e^{\frac{i}{\hbar}(\mathbf{p}\mathbf{x}-\varepsilon t)}\delta(\varepsilon^2 - \varepsilon^2(\mathbf{p}))\hat{\phi}_*(\mathbf{p}). \tag{1.4}$$

Consider the following expression showing a d'Alembert operator acting on the field (1.2):

$$
\begin{aligned}
\Box\hat{\phi}(t,\mathbf{x}) &= \Big(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta\Big)\hat{\phi}(t,\mathbf{x}) \\
&= \int d^3p\, d\varepsilon\, e^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{x}-\varepsilon t)}\,\frac{1}{\hbar^2}\Big(\mathbf{p}^2 - \frac{1}{c^2}\varepsilon^2\Big)\delta(\varepsilon^2-\varepsilon^2(\mathbf{p}))\hat{\phi}_*(\mathbf{p}) \\
&= \int d^3p\, e^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{x}-\varepsilon(\mathbf{p})t)}\,\frac{1}{\hbar^2}\Big[\mathbf{p}^2 - \frac{1}{c^2}\varepsilon^2(\mathbf{p})\Big]\,\delta(\varepsilon^2-\varepsilon^2(\mathbf{p}))\hat{\phi}_*(\mathbf{p}) \\
&= \int d^3p\, e^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{x}-\varepsilon(\mathbf{p})t)}\,\frac{1}{\hbar^2}\Big[\mathbf{p}^2 - \frac{1}{c^2}(c^2\mathbf{p}+m^2c^4)\Big]\delta(\varepsilon^2-\varepsilon^2(\mathbf{p}))\hat{\phi}_*(\mathbf{p}) \\
&= -\frac{m^2c^2}{\hbar^2}\,\hat{\phi}(t,\mathbf{x}).
\end{aligned}
$$

Thus, we find that the free field operator should satisfy

$$
\Big(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta + \frac{m^2c^2}{\hbar^2}\Big)\hat{\phi}(t,\mathbf{x}) \;=\; \Big(\Box + \frac{m^2c^2}{\hbar^2}\Big)\hat{\phi}(t,\mathbf{x}) = 0, \tag{1.5}
$$

the Klein–Gordon equation. Equation (1.5) is a direct consequence of the relativistic dispersion relation between the energy and the momentum of the particle.

If the field corresponds to a massless particle, the parameter m in Eq. (1.5) is zero. Therefore, the field operator of a free massless field satisfies the wave equation

$$
\Box\hat{\phi}(t,\mathbf{x}) \;=\; 0. \tag{1.6}
$$

Thus, any kind of a free relativistic quantum field is a spacetime-dependent operator satisfying the Klein–Gordon equation. In the case of interacting quantum fields, their dynamics is described by much more complicated equations which will be discussed in the following chapters.

1.3 Natural units, notations and conventions

It is evident that the units of measurements of physical quantities should correspond to the scales of phenomena where these units are used. For example, it is not reasonable to measure the masses of elementary particles in tons or grams, or the size of atomic nuclei in kilometers or centimeters.

When we consider the relativistic high-energy quantum phenomena in the fundamental quantum physics of elementary particles, it is natural to employ the units related to the fundamental constants of nature. This means that we have to choose the system of units where the speed of light is $c = 1$, and the Planck constant (which has the dimension of the action) is $\hbar = 1$. As a result, we obtain the natural system of units based only on the fundamental constants of nature. In these units, the action is dimensionless, the speed is dimensionless and the dimensions of energy and momentum coincide. As in quantum theory, there is a Planck formula, relating energy and frequency as $\varepsilon \sim \hbar\omega$, and $\omega \sim \frac{1}{t}$, where t is time, and the dimensions satisfy the relation

$$[\varepsilon] = [\mathbf{p}] = [m] = [l]^{-1} = [t]^{-1}. \tag{1.7}$$

Thus, we have only one remaining dimensional quantity, the unit of energy. Usually, the energy in high-energy physics is measured in electron-volts, such that the unit of energy is $1\ eV$, or $1\,GeV = 10^9\,eV$. The dimensions of length and time are identical. In what follows, we shall use this approach and assume the natural units of measurements described above, with $\hbar = c = 1$.

Other notations and conventions are as follows:

1) Minkowski space coordinates $x^\mu \equiv (x^0, \mathbf{x}) \equiv (t, \mathbf{x}) \equiv (x^0, x^i)$, where Greek letters represent the spacetime indices $\alpha, \dots, \mu = 0, 1, 2, 3$, while Latin letters are reserved for the space indices, $i, j, k, \dots = 1, 2, 3$.

2) Functions in Minkowski space are denoted as $\phi(x) \equiv \phi(x^0, x^i) \equiv \phi(x^0, \mathbf{x}) \equiv \phi(t, \mathbf{x})$.

3) The Minkowski metric is

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \equiv \text{diag}(1, -1, -1, -1), \tag{1.8}$$

and the same is true for the inverse metric, $\eta^{\mu\nu} = \text{diag}(1, -1, -1, -1)$. One can easily check the relations $\eta^{\mu\nu}\eta_{\nu\rho} = \delta^\mu{}_\rho$ and $\eta_{\mu\nu}\eta^{\nu\rho} = \delta_\mu{}^\rho$.

Furthermore, $\varepsilon^{\mu\nu\alpha\beta}$ is the four-dimensional, totally antisymmetric tensor. The sign convention is that $\varepsilon^{0123} = 1$ and hence $\varepsilon_{0123} = -1$.

4) Partial derivatives are denoted as

$$\frac{\partial}{\partial x^\mu} \equiv \partial_\mu, \qquad \frac{\partial^2}{\partial x^\mu \partial x^\nu} \equiv \partial_\mu \partial_\nu\,, \qquad \text{etc.} \tag{1.9}$$

5) Rising and lowering the indices looks like

$$A^\mu = \eta^{\mu\nu} A_\nu, \qquad A_\mu = \eta_{\mu\nu} A^\nu, \qquad \partial^\mu = \eta^{\mu\nu}\partial_\nu, \qquad \partial_\mu = \eta_{\mu\nu}\partial^\nu, \quad \text{etc.}$$

Let us note that these and some other rules will be changed in Part II, when we start to deal with curved spacetime.

6) The scalar product is as follows:

$$AB = A^\mu B_\mu = A^0 B_0 + A^i B_i = A_0 B_0 - A_i B_i.$$

In particular,

$$px = p_\mu x^\mu = p_0 x^0 + p_i x^i = p_0 x^0 - \mathbf{p}\cdot\mathbf{x},$$

where $p^\mu \equiv (p^0, \mathbf{p})$.

7) The integral over four-dimensional space is

$$\int d^4x = \int d^3x \int dx_0,$$

while the integral over three-dimensional space is $\int d^3x$.

8) Dirac's delta function in Minkowski space is

$$\delta^4(x - x') \equiv \delta(t - t')\delta(\mathbf{x} - \mathbf{x}') \equiv \delta(t - t')\delta(x^1 - x^{'1})\delta(x^2 - x^{'2})\delta(x^3 - x^{'3}). \quad (1.10)$$

In particular, this means $\int d^4x\, \delta^4(x - x')\phi(x') = \phi(x)$.

9) The d'Alambertian operator is

$$\Box = \partial_0^2 - \partial_1^2 - \partial_2^2 - \partial_3^2 = \eta^{\mu\nu}\partial_\mu\partial_\nu = \partial_0{}^2 - \Delta, \quad (1.11)$$

where the Laplace operator is

$$\Delta = \partial_0^2 + \partial_2^2 + \partial_3^2. \quad (1.12)$$

10) The convention is that repeated indices imply the summation in all cases, i.e.,

$$X_I Y_I = \sum_{I=1}^{N} X_I Y_I. \quad (1.13)$$

Comments

There are many books on quantum field theory that differ in their manner and level of presentation, targeting different audiences that range from beginners to more advanced readers. Let us present a short list of basic references, which is based on our preferences.

The standard textbooks covering the basic notions and methods are those by J.D. Bjorken and S.D. Drell [57], C. Itzykson and J.-B. Zuber [187], M.E. Peskin and D.V. Schroeder [250], M. Srednicki [304] and M.D. Schwartz [274].

A brief and self-contained introduction to modern quantum field theory can be found in the books by P. Ramond [256], M. Maggiore [215] and L. Alvarez-Gaume and M.A. Vazquez-Mozo [155].

Comprehensive monographs in modern quantum field theory, with extensive coverage but aimed for advanced readers are those by J. Zinn-Justin [356], S. Weinberg [345], B.S. DeWitt [106, 109] and W. Siegel [293].

There are also very useful lecture notes available online, e.g., those by H. Osborn [235]. For mathematical and axiomatical aspects and approaches to quantum field theory see, e.g., the book by N.N. Bogolubov, A.A. Logunov, A.I. Oksak and I. Todorov [60].

2
Relativistic symmetry

In this chapter, we briefly review special relativistic symmetry, which will be used in the rest of the book. In particular, we introduce basic notions of the Lorentz and Poincaré groups, which will be used in constructing classical and quantum fields.

In general, the principles of symmetry play a fundamental role in physics. One of the most universal symmetries of nature is the one that we can observe in the framework of special relativity.

2.1 Lorentz transformations

According to special relativity, a spacetime structure is determined by the following general principles:

1) Space and time are homogeneous.

2) Space is isotropic.

3) There exists a maximal speed of propagation of a physical signal. This maximal speed coincides with the speed of light. In all inertial reference frames the speed of light has the same value, c.

Let P_1 and P_2 be two infinitesimally separated events that are points in spacetime. In some inertial reference frame, the four-dimensional coordinates of these events are x^μ and $x^\mu + dx^\mu$. The interval between these two events is defined as

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu. \tag{2.1}$$

In another inertial reference frame, the same two events have the coordinates x'^μ and $x'^\mu + dx'^\mu$. The corresponding interval is

$$ds'^2 = \eta_{\mu\nu} dx'^\mu dx'^\nu. \tag{2.2}$$

The two intervals (2.1) and (2.2) are equal, that is, $ds'^2 = ds^2$, reflecting the independence of the speed of light on the choice of the inertial reference frame. Thus,

$$\eta_{\mu\nu} dx^\mu dx^\nu \;=\; \eta_{\alpha\beta}\, dx'^\alpha dx'^\beta. \tag{2.3}$$

Eq. (2.3) enables one to find the relation between the coordinates x'^α and x^μ. Let $x'^\alpha = f^\alpha(x)$, with some unknown function $f^\alpha(x)$. Substituting this relation into Eq. (2.3), one gets an equation for the function $f^\alpha(x)$ that can be solved in a general form. As a result,

$$x'^\alpha = \Lambda^\alpha{}_\mu x^\nu + a^\alpha, \tag{2.4}$$

Introduction to Quantum Field Theory with Applications to Quantum Gravity. Iosif L. Buchbinder and Ilya L. Shapiro, Oxford University Press (2021).
DOI: 10.1093/oso/9780198838319.003.0002

where $\Lambda \equiv (\Lambda^\alpha{}_\mu)$ is a matrix with constant elements, and a^α is a constant four-vector. Substituting Eq. (2.4) into Eq. (2.3), we get

$$\eta_{\alpha\beta}\Lambda^\alpha{}_\mu\Lambda^\beta{}_\nu = \eta_{\mu\nu}. \tag{2.5}$$

The coordinate transformation (2.4) with the matrix $\Lambda^\alpha{}_\mu$, satisfying Eq. (2.5), is called the non-homogeneous Lorentz transformation. One can say that the non-homogeneous Lorentz transformation is the most general coordinate transformation preserving the form of the interval (2.1). If in Eq. (2.4) the vector $a^\alpha = 0$, the corresponding coordinate transformation is called the homogeneous Lorentz transformation, or simply the Lorentz transformation. Such a transformation has the form

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{2.6}$$

with the matrix $\Lambda^\mu{}_\nu$ satisfying Eq. (2.5).

It is convenient to present the relation (2.5) in a matrix form. Let us introduce the matrices $\eta \equiv (\eta_{\alpha\beta})$ and $\Lambda \equiv (\Lambda^\alpha{}_\mu)$. Then Eq. (2.5) can be written as

$$\Lambda^T \eta \, \Lambda = \eta, \tag{2.7}$$

where Λ^T is the transposed matrix with the elements $(\Lambda^T)_\mu{}^\alpha = \Lambda^\alpha{}_\mu$. One can regard Eq. (2.7) as a basic relation. Any homogeneous Lorentz transformation is characterized by the matrix Λ satisfying the basic relation, and vice versa. Therefore, the set of all homogeneous Lorentz transformations is equivalent to the set of all matrices Λ, satisfying (2.7).

Let us consider some important particular examples of Lorentz transformations:

1. Matrix Λ has the form

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & R^i{}_j \end{pmatrix} \tag{2.8}$$

where the matrix $R = (R^i{}_j)$ transforms only space coordinates, $x'^i = R^i{}_j x^j$. Substituting eq.(2.8) into the basic relation (2.7), we obtain the orthogonality condition

$$R^T R = \mathbf{1_3}, \quad \text{or} \quad R^i{}_k \delta_{ij} R^j{}_l = \delta_{kl}, \tag{2.9}$$

where $\mathbf{1_3}$ is a three-dimensional unit matrix with elements δ_{ij}. Relation (2.9) defines the three-dimensional rotations

$$x'^0 = x^0, \qquad x'^i = R^i{}_j x^j. \tag{2.10}$$

If matrix R satisfies Eq. (2.9), then the transformation (2.10) is the Lorentz transformation. Thus, the three-dimensional rotations represent a particular case of Lorentz transformation.

2. Consider a matrix Λ with the form

$$\Lambda = \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 & \frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{\frac{v}{c}}{\sqrt{1-\frac{v^2}{c^2}}} & 0 & 0 & \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \end{pmatrix}, \tag{2.11}$$

where $(v/c)^2 < 1$. It is easy to show that this matrix satisfies the basic relation. Therefore, this matrix describes a Lorentz transformation,

$$x'^0 = \frac{x^0 + \frac{v}{c}x^1}{\sqrt{1-\frac{v^2}{c^2}}}, \qquad x'^1 = x^1, \qquad x'^2 = x^2, \qquad x'^3 = \frac{x^3 + \frac{v}{c}x^0}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{2.12}$$

This is the standard form of the Lorentz transformation for the case when one inertial frame moves with respect to another one in the x^3 direction. Indeed, one can construct a similar matrix describing relative motion in any other direction. Transformations of the type (2.12) are called boosts.

3. The matrix Λ corresponding to the time inversion, or T-transformation, is

$$\Lambda = \Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

This matrix corresponds to the Lorentz transformation

$$x'^0 = -x^0\,, \qquad x'^i = x^i. \tag{2.13}$$

4. Let the matrix Λ have the form

$$\Lambda = \Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

It is easy to check that the basic relation (2.7) is fulfilled in this case. This matrix corresponds to the following Lorentz transformation:

$$x'^0 = x^0 \qquad x'^i = -x^i, \tag{2.14}$$

which is called the space reflection or parity (P) transformation.

5. The matrix Λ with the form

$$\Lambda = \Lambda_{PT} = \Lambda_P\Lambda_T = \Lambda_T\Lambda_P = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

corresponds to the following Lorentz transformation:

$$x'^\mu = -x^\mu, \tag{2.15}$$

which is called the full reflection.

Eqs. (2.13), (2.14), (2.15) are called discrete Lorentz transformations.

We will mainly need only the subclass of all Lorentz transformations that can be obtained by small deformations of the identical transformation. Let the transformation matrix have the form $\Lambda = I$, where I is the unit 4×4 matrix with elements $\delta^\mu{}_\nu$. Matrix I satisfies the basic relation (2.7). This matrix realizes the identical Lorentz transformation

$$x'^\mu = x^\mu.$$

Stipulating small deformations of identical transformations means that we consider matrices Λ of the form

$$\Lambda = I + \omega, \tag{2.16}$$

where ω is a matrix with infinitesimal elements $\omega^\mu{}_\nu$. Requiring that the matrix Λ from (2.16) correspond to a Lorentz transformation, we arrive at the relation

$$(I + \omega)^T \eta (I + \omega) = \eta.$$

Taking into account only the first-order terms in ω, one gets

$$\omega^T \eta + \eta \omega = 0.$$

Recovering the indices, we obtain

$$(\omega^T)_\mu{}^\alpha \eta_{\alpha\nu} + \eta_{\mu\alpha}\omega^\alpha{}_\nu = 0 \quad \Longrightarrow \quad \omega^\alpha{}_\mu \eta_{\alpha\nu} + \eta_{\mu\alpha}\omega^\alpha{}_\nu = \omega_{\mu\nu} + \omega_{\nu\mu} = 0. \tag{2.17}$$

One can see that the matrix ω is real and antisymmetric, and hence it has six independent elements. The matrix Λ (2.16) corresponds to the coordinate transformation

$$x'^\mu = x^\mu + \omega^\mu{}_\nu x^\nu,$$

which is called the infinitesimal Lorentz transformation.

2.2 Basic notions of group theory

Group theory is a branch of mathematics devoted to the study of the symmetries. In this subsection, we consider the basic notions of group theory that will be used in the rest of the book. It is worth noting that this section is not intended to replace a textbook on group theory. In what follows, we consequently omit rigorous definitions and proofs of the theorems and concentrate only on the main notions of our interest.

A set G of the elements g_1, g_2, $g_3, \ldots$, equipped with a law of composition (or product of elements, or multiplication rule, or composition law), e.g., $g_1 g_2$, is called a group if for each pair of elements $g_1, g_2 \in G$, the composition law satisfies the following set of conditions:

1) Closure, i.e., $\forall g_1$, $g_2 \in G$: $g_1 g_2 \in G$.
2) Associativity, i.e., $\forall g_1$, g_2, $g_3 \in G$, for the product $g_1(g_2 g_3) = (g_1 g_2) g_3$.
3) Existence of unit element, i.e., $\exists e \in G$, such that $\forall g \in G : ge = eg = g$.
4) Existence of inverse element, i.e., $\forall g \in G$, $\exists g^{-1} \in G$ such that $gg^{-1} = g^{-1} g = e$.

Using these conditions, one can prove the uniqueness of the unit and inverse elements.

A group is called Abelian or commutative if, $\forall g_1,\, g_2 \in G$, the product satisfies $g_1 g_2 = g_2 g_1$. In the opposite case, the group is called non-Abelian or non-commutative, i.e., $\exists g_1, g_2 \in G$ such that $g_1 g_2 \neq g_2 g_1$.

A subset $H \subset G$ is said to be a subgroup of group G if H itself is the group under the same multiplication rule as group G. In particular, this means if $h_1, h_2 \in H$, then $h_1 h_2 \in H$. Also, $e \in H$, and if $h \in H$, then $h^{-1} \in H$.

A group consisting of a finite number of elements is called finite. In this case, it is possible to form a group table $g_i g_j$. A finite group is sometimes called a finite discrete group.

Let us consider a few examples:

1. Let G be a set of $n \times n$ real matrices M such that $\det M \neq 0$. It is evident that if $M_1,\, M_2 \in G$, then $\det M_1 M_2 = \det M_1 \det M_2 \neq 0$ and hence $M_1,\, M_2 \in G$. Thus, this set forms a group under the usual matrix multiplication. The unit element is the unit matrix E, and the element inverse to the matrix M is the inverse matrix M^{-1}. We know that the multiplication of matrices is associative. Thus, all group conditions are fulfilled. This group is called a general linear n-dimensional real group and is denoted as $GL(n|\mathbb{R})$. Consider a subset $H \subset GL(n|\mathbb{R})$ consisting of matrices N that satisfy the condition $\det N = 1$. It is evident that $\det (N_1 N_2) = \det N_1 \det N_2 = 1$. Hence

$$N_1,\, N_2 \,\in\, H \implies N_1 N_2 \in H.$$

Consider other properties of this group. It is evident that $E \in H$. On the top of this,

$$N \in G \quad\implies\quad \det\big(N^{-1}\big) \,=\, (\det N)^{-1} \,=\, 1.$$

The last means $N^{-1} \in H$. Hence H is a subgroup of the group $GL(n|\mathbb{R})$. Group H is called a special linear n-dimensional real group and is denoted as $SL(n|\mathbb{R})$. In a similar way, one can introduce general and special complex groups $GL(n|\mathbb{C})$ and $SL(n|\mathbb{C})$, respectively, where $\mathbb{C}$ is a set of complex numbers.

2. Let G be a set of complex $n \times n$ matrices U such that $U^+U = UU^+ = E$, where E is the unit $n \times n$ matrix. Here, as usual, $(U^+)_{ab} = (U^*)_{ba}$ or $U^\dagger = (U^*)^T$, where $*$ means the operation of complex conjugation, and T means transposition. Evidently, $E \in G$ and, for any $U_1,\, U_2 \in G$, the following relations take place:

$$\begin{aligned}
&(U_1U_2)^+(U_1U_2) = U_2^+(U_1^+U_1)U_2 = U_2^+U_2 = E,\\
&(U_1U_2)(U_1U_2)^+ = U_1(U_2U_2^+)U_1^+ = U_1U_1^+ = E.
\end{aligned} \tag{2.18}$$

In addition, if $U \in G$, then $(U^{-1})^+U^{-1} = (UU^+)^{-1} = U^{-1}(U^{-1})^+ = (U^+U)^{-1} = E$. Therefore, if $U \in G$, then $U^{-1} \in G$ too. As a result, the set of matrices under consideration form a group. This group is called the n-dimensional unitary group $U(n)$.

The condition $U^+U = E$ leads to $|\det U|^2 = 1$. Hence $\det U = e^{i\alpha}$, where $\alpha \in \mathbb{R}$. One can also consider a subset of matrices $U \in U(n)$, that satisfy the relation $\det U = 1$. This subset forms a special unitary group and is denoted $SU(n)$.

Since the multiplication of matrices is, in general, a non-commutative operation, the matrix groups $GL(n,\mathbb{R})$, $SL(n,\mathbb{R})$, $U(n)$ and $SU(n)$ are, in general, non-Abelian.

A group G is called the Lie group if each of its element is a differentiable function of the finite number of parameters, and the product of any two group elements is a differentiable function of parameters of each of the factors. That is, consider, $\forall g \in G$, and for $g_1 = g\big(\xi_1^{(1)},\dots,\xi_N^{(1)}\big)$ and $g_2 = g\big(\xi_1^{(2)},\dots,\xi_N^{(2)}\big)$, $g = g_1 g_2 = g(\xi_1,\dots,\xi_N)$. Then

$$\xi_I = f_I\big(\xi_1^{(1)},\dots,\xi_N^{(1)},\xi_1^{(2)},\dots,\xi_N^{(2)}\big), \tag{2.19}$$

where $I = 1,2,\dots,N$ are the differentiable functions of the parameters $\xi_1^{(1)},\dots,\xi_N^{(1)},\xi_1^{(2)},\dots,\xi_N^{(2)}$. The Lie group is called compact if the parameters $\xi_1,\dots,\xi_N$ vary within a compact domain. One can prove that the parameters $\xi_1,\dots,\xi_N$ can be chosen in such a way that $g(0,\dots,0) = e$, where e is the unit element of the group.

All matrix groups described in the examples above are the Lie groups, where the role of parameters is played by independent matrix elements.

The two groups G and G' are called homomorphic if there exists a map f of the group G into the group G' such that, for any two elements $g_1, g_2 \in G$, the following conditions take place: $f(g_1 g_2) = f(g_1)f(g_2)$, and if $f(g) = g'$, then $f(g^{-1}) = g'^{-1}$, where g'^{-1} is an inverse element in the group G'. Such a map is called homomorphism. One can prove that $f(e) = e'$, where e' is the unit element of the group G'. One-to-one homomorphism is called isomorphism, and the corresponding groups are called isomorphic. We will write, in this case, $G = G'$.

Let G be some group, and V be a real or complex linear space. Consider a map R such that, $\forall g \in G$, there exists an invertible operator $D_R(g)$ acting in the space V. Furthermore, let the operators $D_R(g)$ satisfy the following conditions:

1) $D_R(e) = I$, where I is a unit operator in the space V; and 2) $\forall g_1,\, g_2 \in G$, we have $D_R(g_1 g_2) = D_R(g_1)D_R(g_2)$.

The map R is called a representation of the group G in the linear space V. Operators $D_R(g)$ are called the operators of representation, and the space V is called the space of the representation. One can prove that, $\forall g \in G$, there is $D_R(g^{-1}) = D_R^{-1}(g)$, where $D_R^{-1}(g)$ is the inverse operator for $D_R(g)$. Thus, the set of operators $D_R(g)$ forms a group where a multiplication rule is the usual operator product.

We will mainly concern ourselves with matrix representations, where the operators $D_R(g)$ are the $n\times n$ matrices $D_R(g)^i{}_j$, $\quad i,j = 1,2,\dots,n$. Let v be a vector in a space of representation with the coordinates $v^1, v^2,\dots,v^n$, in some basis. The matrices $D_R(g)^i{}_j$ generate the coordinate transformation of the form

$$v'^i = D_R(g)^i{}_j v^j.$$

Let R be a representation of the group G in the linear space V, and $\tilde{V}$ be a subspace in V, i.e., $\tilde{V} \subset V$. We assume that, for any vector $\tilde{v} \in \tilde{V}$ and for any operator $D_R(g)$, the condition $D_R(g)\tilde{v} \in \tilde{V}$ takes place. Then, the subspace $\tilde{V}$ is called the invariant subspace of the representation R. Any representation always has two invariant subspaces, which are called trivial. These are the subspace $\tilde{V} = V$, and the subspace $\tilde{V} = \{0\}$, which consists of a single zero element. All other invariant subspaces, if they exist, are called non-trivial. A representation R is called reducible if it has non-trivial invariant subspaces, and irreducible if it does not. In other words, the representation R is called irreducible if it has only trivial invariant subspaces. A representation is called completely irreducible if all representation matrices $D_R(g)$ have the block-diagonal form. This means that, in a certain basis,

$$D_R(g) = \begin{pmatrix} \boxed{D_1(g)} & 0 & 0 \; ... & 0 \\ 0 & \boxed{D_2(g)} & 0 \; ... & 0 \\ & & \ddots \; ... & \\ 0 & 0 & & 0 \\ 0 & 0 & 0 \; ... & \boxed{D_k(g)} \end{pmatrix}.$$

This situation means that the representation space has k non-trivial invariant subspaces. In each of such subspaces, one can define an irreducible representation, $D_k(g)$.

A given Lie group can have different representations, where the matrices $D_R(g)$ may have different forms. However, some properties are independent of the representation. Some of these properties can be formulated, e.g., in terms of Lie algebra. Let $D_R(g)$ be the operators of representation, and $g = g(\xi)$. Then, the operators $D_R(g)$ will be the functions of N parameters $\xi^1, \xi^2, \dots, \xi^N$, i.e., $D_R(g) = D_R(\xi)$ and $D_R(\xi)|_{\xi^I=0} = D_R(e) = \mathbf{1}$, where $\mathbf{1}$ is a unit matrix in the given representation space. One can prove that, in an infinitesimal vicinity of the unit element, operators $D_R(\xi)$ can be presented in the form

$$D_R(\xi) = \mathbf{1} + i\xi^I T_{RI}, \quad \text{where} \quad T_{RI} = -i \left. \frac{\partial D_R(\xi)}{\partial \xi^I} \right|_{\xi=0}. \tag{2.20}$$

The operators T_{RI} are called the generators of the group G in the representation R. One can show that any operator $D_R(\xi)$ which is obtained by the continuous deformation from the unit element can be written as

$$D_R(\xi) = e^{i\xi^I T_{RI}}. \tag{2.21}$$

If the operator $D_R(\xi)$ is unitary, i.e., $D_R(\xi)D_R^+(\xi) = D_R^+(\xi)D_R(\xi) = \mathbf{1}$, then the generators T_{RI} are Hermitian, i.e., $T_{RI} = T_{RI}^+$. The generators of our interest satisfy the following relation in terms of commutators:

$$[T_{RI}, T_{RJ}] = i f_{IJ}{}^K T_{RK}, \tag{2.22}$$

where $f_{IJ}{}^K$ are the structure constants of the Lie group G. It is evident that $f_{IJ}{}^K = -f_{JI}{}^K$. In general, the form of the matrices T_{RI} depends on the representation. However, one can prove that the structure constants do not depend on the representation. Thus, these constants characterize the group G itself.

The group generators are closely related to the notion of Lie algebra. Let A be a real or complex linear space with the elements $a_1, a_2, \dots$. A linear space A is called Lie algebra, if for each two elements $a_1, a_2 \in A$, there exists a composition law (also called multiplication or the Lie product) $[a_1, a_2]$, such that

$$
\begin{aligned}
&1)\ [a_1, a_2] \in A,\\
&2)\ [a_1, a_2] = -[a_2, a_1],\\
&3)\ [c_1 a_1 + c_2 a_2, a_3] = c_1[a_1, a_3] + c_2[a_2, a_3] and\\
&4)\ [a_1, [a_2, a_3]] + [a_2, [a_3, a_1]] + [a_3, [a_1, a_2]] = 0.
\end{aligned} \tag{2.23}
$$

Here, c_1 and c_2 are arbitrary real or complex numbers, and $a_3 \in A$. The composition law $[a_1, a_2]$ is called the Lie bracket, or the commutator. Property 4 is called the Jacobi identity.

It is easy to check that the commutator of the generators (2.22) of the representation of the Lie group G satisfies all properties of the composition law for the Lie algebra. Therefore, the generators T_{RI} form the Lie algebra which is called the Lie algebra associated with a given Lie group G. To be more precise, they form a representation of the Lie algebra. It means that one can define the map $T(a)$ of the Lie algebra into a linear space of operators such that

$$
T : a \longrightarrow T(a) \quad \text{and} \quad [a_1, a_2] \longrightarrow [T(a_1), T(a_2)]. \tag{2.24}
$$

A Lie algebra is called commutative, or Abelian, if, for any two elements $a_1, a_2 \in A$, $[a_1, a_2] = 0$. In the opposite case, the Lie algebra is called non-commutative, or non-Abelian. One can prove that the Lie algebra associated with an Abelian Lie group is Abelian.

2.3 The Lorentz and Poincaré groups

Consider the group properties of Lorentz transformations. The Lorentz transformation has been defined in the form

$$
x'^\mu = \Lambda^\mu{}_\nu x^\nu,
$$

where the matrix Λ satisfies the basic relation (2.7). Let us show that Lorentz transformations form a group. Consider the set of all Lorentz transformations or, equivalently, the set of all matrices Λ satisfying $\Lambda^T \eta \Lambda = \eta$.

For the product of two matrices corresponding to the Lorentz transformations, Λ_1 and Λ_2, we have

$$
(\Lambda_1\Lambda_2)^T \eta (\Lambda_1\Lambda_2) = {\Lambda_2}^T {\Lambda_1}^T \eta \Lambda_1 \Lambda_2 = {\Lambda_2}^T \eta \Lambda_2 = \eta. \tag{2.25}
$$

Thus, the matrix product $\Lambda_1\Lambda_2$ satisfies the basic relation (2.7), and hence two consequent Lorentz transformations are equivalent to another Lorentz transformation,

$$
x''^\mu = {\Lambda_1}^\mu{}_\alpha {\Lambda_2}^\alpha{}_\nu x^\nu. \tag{2.26}
$$

Let I be the unit 4×4 matrix with the elements $\delta^\mu{}_\nu$. It is evident that $I^T \eta I = \eta$, *i.e.*, the matrix I corresponds to a Lorentz transformation.

The next step is to check the existence of an inverse element. The basic relation (2.7) can be recast in the form $\Lambda^T \eta = \eta \Lambda^{-1}$ or, equivalently, $\eta = (\Lambda^T)^{-1} \eta \Lambda^{-1}$, or $(\Lambda^{-1})^T \eta \Lambda^{-1} = \eta$. Thus, matrix Λ^{-1} also corresponds to a Lorentz transformation.

Lorentz transformations form a group where the multiplication law is a standard product of matrices Λ. Such a group is called the Lorentz group. Let us explore it in more detail. From the basic relation $\Lambda^T \eta \Lambda = \eta$ follows

$$\det \Lambda^T \det \eta \det \Lambda = \det \eta \qquad \text{and} \qquad \det \eta = -1 \neq 0.$$

As a result, $(\det \Lambda)^2 = 1$, and hence $\det \Lambda = \pm 1$.

Starting from the relation $\Lambda^\alpha{}_\mu \eta_{\alpha\beta} \Lambda^\beta{}_\nu = \eta_{\mu\nu}$, and setting $\mu = 0$ and $\nu = 0$, one gets $1 = \Lambda^0{}_0\, \eta_{00}\, \Lambda^0{}_0 + \Lambda^i{}_0\, \eta_{ij}\, \Lambda^j{}_0$. Since $\eta_{ij} = -\delta_{ij}$ and $\eta_{00} = 1$, one obtains

$$1 = (\Lambda^0{}_0)^2 - \Lambda^i{}_0 \Lambda^i{}_0 \qquad \Longrightarrow \qquad (\Lambda^0{}_0)^2 = 1 + \Sigma_{i=1}^3 (\Lambda^i{}_0)^2.$$

Therefore, $(\Lambda^0{}_0)^2 \geq 1$, and hence $|\Lambda^0{}_0| \geq 1$. As a result, we have the two relations

$$(\det \Lambda)^2 = 1, \qquad |\Lambda^0{}_0| \geq 0. \tag{2.27}$$

Thus, the following cases are possible:

$$\begin{array}{ll} \det \Lambda = 1, & \Lambda^0{}_0 > 0, \\ \det \Lambda = -1, & \Lambda^0{}_0 > 0, \\ \det \Lambda = 1, & \Lambda^0{}_0 < 0, \\ \det \Lambda = -1, & \Lambda^0{}_0 < 0. \end{array}$$

It means that the set of all Lorentz transformation is separated into four subsets:

$$\begin{array}{ll} L_+{}^\uparrow: & \text{the set of matrices}\, \Lambda\, \text{such that } \det \Lambda = 1,\ \Lambda^0{}_0 > 0, \\ L_-{}^\uparrow: & \text{the set of matrices}\, \Lambda\, \text{such that } \det \Lambda = -1,\ \Lambda^0{}_0 > 0, \\ L_+{}^\downarrow: & \text{the set of matrices}\, \Lambda\, \text{such that } \det \Lambda = 1,\ \Lambda^0{}_0 < 0, \\ L_-{}^\downarrow: & \text{the set of matrices}\, \Lambda\, \text{such that } \det \Lambda = -1,\ \Lambda^0{}_0 < 0. \end{array} \tag{2.28}$$

It is easy to see that $I \in L_+{}^\uparrow$. One can show that the subset $L_+{}^\uparrow$ forms a group which is called the proper Lorentz group. It is evident that the proper Lorentz group is a subgroup of the Lorentz group. All other subsets do not form groups.

Remark. The infinitesimal Lorentz transformations are generated by the matrix $\Lambda = I + \omega$. Since $I \in L_+{}^\uparrow$, the matrices $\Lambda = I + \omega \in L_+{}^\uparrow$. Since the matrix ω has six real independent elements, the proper Lorentz group is a six-parametric real Lie group.

Now let us consider a set of non-homogeneous Lorentz transformations

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu,$$

where the matrix Λ satisfies the basic relation, and a^μ is a constant four-vector. Applying two non-homogeneous Lorentz transformations, one after another, we get

$$x'^\alpha = \Lambda_1{}^\alpha{}_\nu x^\nu + a_1{}^\alpha, \qquad x''^\mu = \Lambda_2{}^\mu{}_\alpha x'^\alpha + a_2{}^\mu. \tag{2.29}$$

Substituting the first relation into the second one, we get

$$x''^{\mu} = \Lambda_2{}^{\mu}{}_{\alpha}(\Lambda_1{}^{\alpha}{}_{\nu}x^{\nu} + a_1{}^{\alpha}) + a_2{}^{\mu} \; = \; \Lambda_2{}^{\mu}{}_{\alpha}\Lambda_1{}^{\alpha}{}_{\nu}x^{\nu} + \Lambda_2{}^{\mu}{}_{\alpha}a_1{}^{\alpha} + a_2{}^{\mu}. \quad (2.30)$$

According to what we proved before, the matrix $\Lambda_2{}^{\mu}{}_{\alpha}\Lambda_1{}^{\alpha}{}_{\nu} = (\Lambda_2\Lambda_1)^{\mu}{}_{\alpha}$ satisfies the basic relation. Furthermore, the quantity $\Lambda_2{}^{\mu}{}_{\alpha}a_1{}^{\alpha} + a_2{}^{\mu}$ is a constant four-vector. Let us denote

$$\Lambda^{\mu}{}_{\nu} = \Lambda_2{}^{\mu}{}_{\alpha}\Lambda_1{}^{\alpha}{}_{\nu}\;, \qquad a^{\mu} = \Lambda_2{}^{\mu}{}_{\alpha}a_1{}^{\alpha} + a_2{}^{\mu}.$$

Then the relation (2.30) becomes

$$x''^{\mu} = \Lambda^{\mu}{}_{\nu}x^{\nu} + a^{\mu}. \quad (2.31)$$

Since the matrix $\Lambda^{\mu}{}_{\nu}$ satisfies the basic relation, Eq. (2.31) gives us again the non-homogeneous Lorentz transformation. We denote the non-homogeneous Lorentz transformation as (Λ, a), and define, as per (2.30), the multiplication rule on a set of all such transformations as follows:

$$(\Lambda_2, a_2)(\Lambda_1, a_1) = (\Lambda_2\Lambda_1, \Lambda_2 a_1 + a_2). \quad (2.32)$$

Using the transformation $(I, 0)$, we arrive at the relations

$$(\Lambda, a)(I, 0) = (\Lambda I, \Lambda \cdot 0 + a) = (\Lambda, a) \; = \; (I, 0)(\Lambda, a).$$

It is clear that the transformation $(I, 0)$ plays the role of the identity transformation. Let (Λ, a) be a non-homogeneous Lorentz transformation, and consider the transformation $(\Lambda^{-1}, -\Lambda^{-1}a)$. We have

$$(\Lambda, a)(\Lambda^{-1}, -\Lambda^{-1}a) = (\Lambda\Lambda^{-1}, \; -\Lambda\Lambda^{-1}a + a) = (I, 0) = (\Lambda^{-1}, -\Lambda^{-1}a)(\Lambda, a).$$

Hence, the transformation $(\Lambda^{-1}, -\Lambda^{-1}a)$ is the inverse of (Λ, a). As a result, the set of all non-homogeneous Lorentz transformations forms a group, with the multiplication rule given by Eq. (2.32). This is the Poincaré group.

An infinitesimal non-homogeneous Lorentz transformation has the form

$$x'^{\mu} = x^{\mu} + \omega^{\mu}{}_{\nu}x^{\nu} + a^{\mu}, \quad (2.33)$$

where $\omega^{\mu}{}_{\nu}$ is a matrix with infinitesimal elements, and a^{μ} is an infinitesimal constant four-vector. Since $\omega_{\mu\nu} = -\omega_{\nu\mu}$, the matrix $\omega_{\mu\nu}$ has six real independent elements. In the vicinity of the unit element $(I, 0)$, any element of the Poincaré group is determined by the real parameters $\omega_{\mu\nu}$ and a_{μ}. Thus, the Poincaré group is the ten-parametric Lie group.

2.4 Tensor representation

We defined a group representation as a map of the group into a group of matrices acting in a linear space. The Lorentz and Poincaré groups express the special relativity principles; therefore, the representations of these groups in the space of the fields define the types of the fields compatible with the principles of relativity. In this and the following sections, we consider the simplest representations of Lorentz and Poincaré groups. The general theory of the representations of these groups is well developed, but its detailed consideration is beyond the scope of this book.

Let us start with the linear space of the tensor fields. Consider a set of all coordinate systems related by the transformations

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu} + a^{\mu}, \tag{2.34}$$

with matrices Λ satisfying the basic relation (2.7). We will call all such coordinate systems admissible. Let us assume that, in some admissible coordinate system $\{x^{\mu}\}$, there is a set of 4^{m+n} functions $t^{\mu_1\cdots\mu_m}{}_{\nu_1\cdots\nu_n}(x)$, while, in another admissible coordinate system $\{x'^{\mu}\}$, there is a set of 4^{m+n} functions $t'^{\mu_1\cdots\mu_m}{}_{\nu_1\cdots\nu_n}(x')$. If these two sets are related to each other as

$$t'^{\mu_1\cdots\mu_m}{}_{\nu_1\cdots\nu_n}(x') = \frac{\partial x'^{\mu_1}}{\partial x^{\alpha_1}} \cdots \frac{\partial x'^{\mu_m}}{\partial x^{\alpha_m}} \frac{\partial x'^{\beta_1}}{\partial x^{\nu_1}} \cdots \frac{\partial x'^{\beta_n}}{\partial x^{\nu_n}} t^{\alpha_1\cdots\alpha_m}{}_{\beta_1\cdots\beta_n}(x), \tag{2.35}$$

then these functions form a tensor (or a tensor field) of the type (m, n). The numbers $t^{\mu_1\cdots\mu_n}{}_{\nu_1\cdots\nu_n}(x)$ are called the components of the tensor in the coordinate frame $\{x^{\mu}\}$.

Starting from the relation (2.34), we get

$$\frac{\partial x'^{\mu}}{\partial x^{\alpha}} = \Lambda^{\mu}{}_{\alpha}. \tag{2.36}$$

It is easy to see that

$$(\Lambda^{-1})^{\alpha}{}_{\mu}\Lambda^{\mu}{}_{\beta}\frac{\partial x^{\beta}}{\partial x'^{\nu}} = (\Lambda^{-1})^{\alpha}{}_{\mu}\delta^{\mu}{}_{\nu} \quad\Longrightarrow\quad \frac{\partial x^{\alpha}}{\partial x'^{\nu}} = (\Lambda^{-1})^{\alpha}{}_{\nu} = (\Lambda^{-1})^{T}{}_{\nu}{}^{\alpha}, \tag{2.37}$$

where T means a matrix transposition.

Then, the definition (2.35) can be rewritten as

$$\begin{aligned} t'^{\mu_t\cdots\mu_m}{}_{\nu_1\cdots\nu_n}(x') = {} & \Lambda^{\mu_1}{}_{\alpha_1} \dots \Lambda^{\mu_m}{}_{\alpha_m} \\ & \times (\Lambda^{-1})^{T}{}_{\nu_1}{}^{\beta_1} \dots (\Lambda^{-1})^{T}{}_{\nu_n}{}^{\beta_n} t^{\alpha_1\cdots\alpha_m}{}_{\beta_1\cdots\beta_n}(x). \end{aligned} \tag{2.38}$$

This relation means that, for any element of the Poincaré group (Λ, a), there is a tensor transformation (2.38). One can check that all conditions of the group representation are fulfilled for the transformations (2.38). Therefore, Eq. (2.38) defines a representation of the Poincaré group that is called the tensor representation.

Setting $a^{\mu} = 0$ in (2.38), we arrive at the tensor representation of the Lorentz group. There are special cases of tensor representations for either Poincaré or Lorentz groups. E.g., the tensor of rank zero, or type $(0, 0)$, is called a scalar. The relation (2.38) in this case is

$$\varphi'(x') = \varphi(x), \qquad x' = \Lambda x + a. \tag{2.39}$$

A tensor with components $t^{\mu}(x)$ is called a contravariant vector. The defining relation (2.38) in this case has the form

$$t'^{\mu}(\Lambda x + a) = \Lambda^{\mu}{}_{\nu} t^{\nu}(x). \tag{2.40}$$

A tensor with components t_μ is called a covariant vector. The defining relation (2.38) in this case looks like

$$t'_\mu(\Lambda x + a) = (\Lambda^{-1})^T{}_\mu{}^\nu t_\nu(x). \tag{2.41}$$

Remark. Using a metric, one can convert a covariant index into a contravariant, and vice versa. For example, for vectors, we have $t^\mu = \eta^{\mu\nu} t_\nu$ and $t_\mu = \eta_{\mu\nu} t^\nu$. Therefore, we can simply call the corresponding geometric object a vector, which may have covariant or contravariant components. Analogously, one can convert any upper tensor index into a lower tensor index, and vice versa.

Our main purpose in this section is to find an infinitesimal form of the Lorentz transformation of tensor components. Let us write

$$x'^\mu = x^\mu + \omega^\mu{}_\nu x^\nu + a^\mu, \tag{2.42}$$

where $\omega_{\mu\nu} = -\omega_{\nu\mu}$ is a matrix with infinitesimal elements, and a^μ is an infinitesimal vector. In what follows, we may omit the indices, when it is possible to do so without causing confusion.

Let us consider the transformation law for the scalar field,

$$\varphi'(x + \omega x + a) = \varphi(x) \quad \text{or} \quad \varphi'(x) + \partial_\mu \varphi(x)(\omega^\mu{}_\nu x^\nu + a^\mu) \;=\; \varphi(x). \tag{2.43}$$

Denote $\varphi'(x) - \varphi(x) = \delta\varphi(x)$, where $\delta\varphi(x)$ is a variation of a scalar field under the infinitesimal coordinate transformations. Let us introduce the operators P_μ and $J_{\alpha\beta}$ by the rule

$$P_\alpha = i\partial_\alpha\,, \qquad J_{\alpha\beta} = \eta_{\alpha\nu} x^\nu P_\beta - \eta_{\beta\nu} x^\nu P_\alpha. \tag{2.44}$$

P_μ and $J_{\alpha\beta}$ are called the generators of spacetime translations and the Lorentz rotations, respectively, in the scalar representation. Using these operators, the variation of the scalar field can be written as follows:

$$\delta\varphi(x) = \Big[ia^\alpha P_\alpha - \frac{i}{2}\,\omega^{\alpha\beta} J_{\alpha\beta}\Big]\varphi(x). \tag{2.45}$$

Let us consider the infinitesimal transformations of a vector field. We have

$$t'^\mu(x + \omega x + a) \;=\; (\delta^\mu{}_\nu + \omega^\mu{}_\nu) t^\nu(x),$$

which can be written as

$$t'^\mu(x) + \partial_\nu t^\mu(x)\big[\omega^\nu{}_\lambda x^\lambda + a^\nu\big] \;=\; t^\mu(x) + \omega^\mu{}_\nu t^\nu(x)$$

and, finally, as

$$\begin{aligned}
\delta t^\mu(x) \;&=\; -\partial_\nu t^\mu(x)\omega^\nu{}_\lambda x^\lambda - \partial_\nu t^\mu(x) a^\nu + \omega^\mu{}_\nu t^\nu(x) \;=\; ia^\alpha(i\partial_\alpha)\delta^\mu{}_\nu t^\nu(x) \\
&\quad + \frac{i}{2}\omega^{\alpha\beta}\Big[i(\delta^\mu{}_\beta \eta_{\alpha\nu} - \delta^\mu{}_\alpha \eta_{\beta\nu}) + i(\delta_\alpha{}^\gamma \eta_{\beta\lambda} x^\lambda - \delta_\beta{}^\gamma \eta_{\alpha\lambda} x^\lambda)\partial_\gamma\Big] t^\mu(x) \\
&= \Big[ia^\alpha (P_\alpha)^\mu{}_\nu - \frac{i}{2}\omega^{\alpha\beta}(J_{\alpha\beta})^\mu{}_\nu\Big] t^\nu(x),
\end{aligned} \tag{2.46}$$

where the following notations were used:

$$
\begin{aligned}
(P_\alpha)^\mu{}_\nu &= \delta^\mu{}_\nu(i\partial_\alpha),\\
(J_{\alpha\beta})^\mu{}_\nu &= (M_{\alpha\beta})^\mu{}_\nu + (S_{\alpha\beta})^\mu{}_\nu,
\end{aligned}
\tag{2.47}
$$

$$
\begin{aligned}
(M_{\alpha\beta})^\mu{}_\nu &= \delta^\mu{}_\nu\Big(x_\alpha P_\beta - x_\beta P_\alpha\Big),\\
(S_{\alpha\beta})^\mu{}_\nu &= i\big(\delta^\mu{}_\alpha\eta_{\beta\nu} - \delta^\mu{}_\beta\eta_{\alpha\nu}\big).
\end{aligned}
\tag{2.48}
$$

The operator $(P_\alpha)^\mu{}_\nu$ is the generator of spacetime translations in the contravariant vector representations, and the operator $(J_{\alpha\beta})^\mu{}_\nu$ is the generator of Lorentz rotations in the covariant vector representation. Finally, P_α and $J_{\alpha\beta}$ are the Poincaré group generators in vector representations.

Similar considerations can be made for any tensor. For instance, in the particular case of the $(1,1)$-type tensor, the result can be written in the symbolic form

$$
\delta t^{A'}{}_{B'}(x) = \Big[ia^\alpha(P_\alpha)^{A'}{}_{B'A}{}^B - \frac{i}{2}\omega^{\alpha\beta}(J_{\alpha\beta})^{A'}{}_{B'A}{}^B\Big]t^A{}_B(x),
$$

where $A' \equiv (\mu'_1,\dots,\mu'_m)$, $B' \equiv (\nu'_1,\dots,\nu'_n)$, $A \equiv (\mu_1,\dots,\mu_m)$ and $B \equiv (\nu_1,\dots,\nu_n)$. As before, the operators P_α and $J_{\alpha\beta}$ are the generators of translations and the Lorentz rotations of the Poincaré group representation in the space of tensors $t^A{}_B$. The explicit form can be found in a way similar to that used in the case of a vector. If the vector $a^\alpha = 0$, one gets the tensor transformation law under the infinitesimal homogeneous Lorentz transformations.

2.5 Spinor representation

Along with tensors, there are other objects associated with the Lorentz group, called spinors. As we will see, in some sense, they are simpler than tensors.

Consider a set of 2×2 complex matrices N with unit determinants. Since N is not degenerate, there is an inverse matrix N^{-1}, and $\det N^{-1} = (\det N)^{-1} = 1$. For the two matrices N_1 and N_2 with $\det N_1 = \det N_2 = 1$, we have $\det(N_1N_2) = \det N_1 \times \det N_2 = 1$. The set of such matrices N forms a group which is called the two-dimensional special complex linear group and is denoted as $SL(2|\mathbb{C})$. We will show that there is a map of group $SL(2|\mathbb{C})$ into group $L_+{}^\uparrow$. Namely, for each matrix $N \in SL(2|\mathbb{C})$, there exists a matrix $\Lambda \in L_+{}^\uparrow$, and vice versa. Moreover,

$$
\Lambda(N_1N_2) \,=\, \Lambda(N_1)\Lambda(N_2) \quad\text{and}\quad \Lambda(N_1) = \Lambda(N_2) \quad\Longrightarrow\quad N_1 = \pm N_2.
$$

The construction of matrix $\Lambda(N)$ consists of several steps:

1. Consider a linear space of Hermitian 2×2 matrices X, $X = X^+$, where $X^+ = (X^*)^T$, where $*$ means the complex conjugation. A basis in this space can be taken as $\sigma_\mu = (\sigma_0, \sigma_j)$, where

$$
\sigma_0 = \begin{pmatrix}1&0\\0&1\end{pmatrix}, \qquad \sigma_1 = \begin{pmatrix}0&1\\1&0\end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix}0&-i\\i&0\end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix}1&0\\0&-1\end{pmatrix}. \tag{2.49}
$$

Here σ_0 is the unit 2×2 matrix, and σ_j are the Pauli matrices. It is evident that all matrices σ_μ are Hermitian. Also, we introduce the matrices $\tilde{\sigma_\mu} = (\sigma_0, -\sigma_j)$. Then it is easy to check that the following relation takes place:

$$\mathrm{tr}\left(\tilde{\sigma_\mu}\sigma_\nu\right) = 2\eta_{\mu\nu}. \tag{2.50}$$

Let X be an arbitrary 2×2 Hermitian matrix, written in the basis σ_μ as

$$X = x^\mu \sigma_\mu. \tag{2.51}$$

Since the matrices X and σ_μ are Hermitian, the x^μ are real numbers which can be identified with coordinates in Minkowski space. The relations (2.50) and (2.51) lead to

$$x^\mu = \frac{1}{2}\,\mathrm{tr}\left(\tilde{\sigma}^\mu X\right). \tag{2.52}$$

2. Let $N \in SL(2|\mathbb{C})$. Consider the matrix

$$X' = NXN^+.$$

It is easy to see that X' is a Hermitian matrix. In addition, $\det X' = \det(NXN^+) = \det X$. Matrices X' and X can be expanded in the basis of σ_μ, providing $X' = x'^\mu \sigma_\mu$ and $X = x^\mu \sigma_\mu$. The coefficients x'^μ can be obtained according to (2.51) as

$$x'^\mu \;=\; \frac{1}{2}\,\mathrm{tr}\left(\tilde{\sigma}^\mu X'\right) \;=\; \frac{1}{2}\,\mathrm{tr}\left(\tilde{\sigma}^\mu N \sigma_\nu N^+\right) x^\nu.$$

Denoting

$$\Lambda^\mu{}_\nu \;=\; \frac{1}{2}\,\mathrm{tr}\left(\tilde{\sigma}^\mu N \sigma_\nu N^+\right) \;\equiv\; \Lambda^\mu{}_\nu(N), \tag{2.53}$$

we arrive at the transformation law

$$x'^\mu = \Lambda^\mu{}_\nu(N) x^\nu. \tag{2.54}$$

3. The $\det X$ can be exactly calculated using an explicit form of the matrices σ_μ,

$$\begin{aligned} X = x^\mu \sigma_\mu &= \begin{pmatrix} x^0 & 0 \\ 0 & x^0 \end{pmatrix} + \begin{pmatrix} 0 & x^1 \\ x^1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -ix^2 \\ ix^2 & -0 \end{pmatrix} + \begin{pmatrix} x^3 & 0 \\ 0 & -x^3 \end{pmatrix} \\ &= \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix}. \end{aligned}$$

Therefore,

$$\begin{aligned} \det X &= (x^0 + x^3)(x^0 - x^3) - (x^1 - ix^2)(x^1 + ix^2) \\ &= (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = \eta_{\mu\nu} x^\mu x^\nu. \end{aligned}$$

Similarly, $\det X' = \eta_{\mu\nu} x'^\mu x'^\nu$. Hence, since the two determinants are equal, we get

$$\eta_{\mu\nu} x'^\mu x'^\nu = \eta_{\alpha\beta} x^\alpha x^\beta. \tag{2.55}$$

This is just the condition of preserving the interval (2.3).

4. Substitute (2.53) into (2.55) and get $\Lambda^T \eta \Lambda = \eta$. Therefore, the matrices $\Lambda(N)$ (2.53) satisfy the basic relation, and hence they realize the Lorentz transformations. Thus, the numbers x^μ from (2.52) can indeed be treated as the Minkowski space coordinates. Also, the matrix $\Lambda^\mu{}_\nu(N)$ (2.53) satisfies the relation $\det \Lambda(N) = 1$. In addition,

$$\Lambda^0{}_0(N) = \frac{1}{2}\,\mathrm{tr}\,\tilde{\sigma}^0 N \sigma_0 N^+ = \frac{1}{2}\,\mathrm{tr}\,(NN^+) > 0.$$

As a result, the matrices $\Lambda(N)$ (2.53) belong to the proper Lorentz group $L_+{}^\uparrow$.

5. We proved that, for each matrix $N \in SL(2|\mathbb{C})$, there is a matrix $\Lambda(N) \in L_+{}^\uparrow$. One can also prove the inverse statement. For each matrix $\Lambda(N) \in L_+{}^\uparrow$, there exists some matrix $N \in SL(2|\mathbb{C})$. Let

$$\Lambda(N_1) = \Lambda(N_2) \quad \Longleftrightarrow \quad \mathrm{tr}\,\tilde{\sigma}^\mu N_1 \sigma_\nu N_1{}^+ \;=\; \mathrm{tr}\,\tilde{\sigma}^\mu N_2 \sigma_\nu N_2{}^+ \quad \Longleftrightarrow \quad N_1 = \pm N_2.$$

There are two matrices N corresponding to a given matrix Λ. Performing two consequent transformations $X' = N_1 X N_1^+$ and $X'' = N_2 X' N_2^+$, one gets $x'^\alpha = \Lambda^\alpha{}_\nu(N_1)x^\nu$ and $x''^\mu = \Lambda^\mu{}_\alpha(N_2)x'^\alpha$. Therefore $x''^\mu = \Lambda^\mu{}_\alpha(N_2)\Lambda^\alpha{}_\nu(N_1)x^\nu$. On the other hand,

$$X'' = N_2 N_1 X N_1{}^+ N_2{}^+ = (N_2N_1)X(N_2N_1)^+.$$

Hence, $x''^\mu = \Lambda^\mu{}_\nu(N_2N_1)x^\nu$. As a result, $\Lambda^\mu{}_\nu(N_2N_1) = \Lambda^\mu{}_\alpha(N_2)\Lambda^\alpha{}_\nu(N_1)$. Thus, we have the mapping

$$\begin{aligned} SL(2|\mathbb{C}) \quad &\longrightarrow \quad L_+{}^\uparrow, \\ N_2N_1 \in SL(2|\mathbb{C}) \quad &\longrightarrow \quad \Lambda(N_2N_1) = \Lambda(N_2)\Lambda(N_1) \in L_+{}^\uparrow, \\ \Lambda_2 = \Lambda_1 \quad &\longrightarrow \quad N_2 = \pm N_1. \end{aligned}$$

We see that the group $SL(2|\mathbb{C})$ is closely related to the proper Lorentz group $L_+{}^\uparrow$. The coordinate transformations in Minkowski space, generated by the elements from the proper Lorentz group, create the transformations $X' = NXN^+$, which are generated by the matrices $N \in SL(2|\mathbb{C})$. Thus, we can consider the transformations $X' = NXN^+$ to be equal footing with the Lorentz transformations $x'^\mu = \Lambda^\mu{}_\nu x^\nu$.

The matrices N act in the two-dimensional complex space formed by the elements $\varphi \equiv \{\varphi_a\}$, with $a = 1, 2$. The rule looks like

$$\varphi'_a = N_a{}^b \varphi_b. \tag{2.56}$$

Since, for each matrix N, there is a matrix $\Lambda(N) \in L_+{}^\uparrow$, one can say that (2.56) is the transformation law of a complex two-component vector under the Lorentz transformation. The vectors $\varphi = \{\varphi_a\}$, transforming according to (2.56), are called left Weyl spinors, and a is called the spinor index. The representation of the $SL(2|\mathbb{C})$ group in the linear space of the left Weyl spinors is called the fundamental representation of the Lorentz group. Next, we introduce the matrix $\varepsilon = (\varepsilon_{ab})$ by the rule

$$\varepsilon = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \Longrightarrow \qquad \varepsilon^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{2.57}$$

and consider the expression

$$f_{ab} = N_a{}^c N_b{}^d \varepsilon_{cd} = N_a{}^1 N_b{}^2 \varepsilon_{12} + N_a{}^2 N_b{}^1 \varepsilon_{21} = N_a{}^2 N_b{}^1 - N_a{}^1 N_b{}^2.$$

It is evident that $f_{11} = f_{22} = 0$, $f_{12} = -f_{21}$ and we can calculate

$$f_{12} = N_1{}^2 N_2{}^1 - N_1{}^1 N_2{}^2 = -\det N = -1.$$

Then $f_{21} = 1$, and hence $f_{ab} = \varepsilon_{ab}$. As a result, one gets $N_a{}^c N_b{}^d \varepsilon_{cd} = \varepsilon_{ab}$, or $N\varepsilon N^T = \varepsilon$. This relation shows that the matrix ε (2.57) is an invariant matrix for group $SL(2|\mathbb{C})$.

Let us introduce the inverse matrix ε^{-1} with elements ε^{ab}, $\varepsilon^{ab}\varepsilon_{bc} = \delta^a{}_c$, $\varepsilon_{ab}\varepsilon^{bc} = \delta_a{}^c$. If ε in (2.57) is an invariant quantity, ε^{-1} is also invariant, and one can prove that

$$\varepsilon^{-1} = N^T \varepsilon^{-1} N. \tag{2.58}$$

The matrices ε and ε^{-1} can be used for raising and lowering the spinor indices

$$\varphi^a = \varepsilon^{ab}\varphi_b, \qquad \varphi_a = \varepsilon_{ab}\varphi^b. \tag{2.59}$$

Let us show that the expression $\varphi_1{}^a \varphi_{2a}$ is invariant. Indeed,

$$\begin{aligned}\varphi_1'{}^a \varphi_{2a}' &= \varepsilon^{ab}\varphi_{1b}'\varphi_{2a}' = \varepsilon^{ab} N_b{}^d \varphi_{1d} N_a{}^c \varphi_{2c} \\ &= (N^T)^c{}_a \varepsilon^{ab} N_b{}^d \varphi_{1d}\varphi_{2c} = \varepsilon^{cd}\varphi_{1d}\varphi_{2c} = \varphi_1{}^c \varphi_{2c}.\end{aligned}$$

Thus, we have learned how to construct a Lorentz invariant object from spinors.

Let N be an arbitrary matrix from $SL(2|\mathbb{C})$ and let the matrix N^* have complex conjugate elements. The elements of this matrix are denoted by definition as $N^*{}_{\dot a}{}^{\dot b}$; $\dot a, \dot b = \dot 1, \dot 2$. This matrix realizes the transformation

$$\chi'_{\dot a} = N^*{}_{\dot a}{}^{\dot b} \chi_{\dot b}. \tag{2.60}$$

The two-dimensional complex vector $\chi \equiv \{\chi_{\dot a}\}$, which transforms according to (2.60), is called the right Weyl spinor, and $\dot a$ is a spinor index. The representation of the $SL(2|\mathbb{C})$ group in the linear space of right Weyl spinors is called the conjugate representation.

It proves useful to introduce the matrices

$$\varepsilon_{\dot a\dot b} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad \varepsilon^{\dot a\dot b} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

which satisfy the relations

$$\varepsilon^{\dot a\dot b} = N^*{}_{\dot a}{}^{\dot c} N^*{}_{\dot b}{}^{\dot d} \varepsilon_{\dot c\dot d}\,, \qquad \varepsilon^{\dot a\dot b} = N^*{}_{\dot c}{}^{\dot a} N^*{}_{\dot d}{}^{\dot b} \varepsilon^{\dot c\dot d}. \tag{2.61}$$

This means that $\varepsilon_{\dot a\dot b}$ and $\varepsilon^{\dot a\dot b}$ are invariant tensors of the group $SL(2|\mathbb{C})$. These matrices can be used for raising and lowering the dotted spinor indices:

$$\chi^{\dot a} = \varepsilon^{\dot a\dot b}\chi_{\dot b}, \qquad \chi_{\dot a} = \varepsilon_{\dot a\dot b}\chi^{\dot b}.$$

It is evident that the expression $\chi_{1\dot a}\chi_2{}^{\dot a}$ is Lorentz invariant; hence, we obtain, once again, a recipe for how to construct the Lorentz invariants from spinors.

Generalization of the undotted φ_a and dotted $\chi_{\dot{a}}$ spinors is a general spin tensor $\varphi_{a_1,\ldots,a_m,\dot{a_1},\ldots,\dot{a_n}}$, defined by the transformation law under the Lorentz transformation as follows:

$$\varphi'_{a_1,\ldots,a_m,\dot{a_1},\ldots,\dot{a_n}}(x') = N_{a_1}{}^{b_1}\ldots N_{a_m}{}^{b_m} N^*{}_{\dot{a}_1}{}^{\dot{b}_1}\ldots N^*{}_{\dot{a}_n}{}^{\dot{b}_n}\varphi_{b_1,\ldots,b_m,\dot{b_1},\ldots,\dot{b_n}}(x). \quad (2.62)$$

Consider the matrices $X' = x'^\mu\sigma_\mu$ and $X = x^\mu\sigma_\mu$, where $X' = NXN^+$ and $x'^\mu = \Lambda^\mu{}_\nu(N)x^\nu$. This means $\Lambda^\mu{}_\nu(N)x^\nu\sigma_\mu = N\sigma_\nu N^+ x^\nu$. Therefore,

$$\sigma_\mu = (\Lambda^T)^{-1}{}_\mu{}^\nu N\sigma_\nu N^+. \quad (2.63)$$

This relation shows that the matrices σ_μ form invariant objects of the group $SL(2|\mathbb{C})$. The matrix elements of σ_μ are denoted as $(\sigma_\mu)_{a\dot{a}}$. Rising the spinor indices, one gets

$$(\sigma_\mu)^{a\dot{a}} = \varepsilon^{ab}\varepsilon^{\dot{a}\dot{b}}(\sigma_\mu)_{b\dot{b}} \equiv (\tilde{\sigma}_\mu)^{\dot{a}a}. \quad (2.64)$$

Using the explicit form of $\sigma_\mu = (\sigma_0, \sigma_j)$ and that of ε^{ab}, $\varepsilon^{\dot{a}\dot{b}}$, one can find $\tilde{\sigma}_\mu = (\sigma_0, -\sigma_j)$. This matrix has already been introduced, at the beginning of section 2.5.

Starting from the spinors φ_a, $\chi_{\dot{a}}$, the matrices σ_μ and the relations (2.56), (2.60) and (2.63), we can construct (complex) vectors under Lorentz transformations, e.g.,

$$\begin{aligned}\varphi'^a(\sigma^\mu)_{a\dot{a}}\chi'^{\dot{a}} &= \varepsilon^{ab}\varphi'_b(\sigma^\mu)_{a\dot{a}}\varepsilon^{\dot{a}\dot{b}}\chi'_{\dot{b}} \\ &= \Lambda^\mu{}_\nu\varepsilon^{ab}N_b{}^c\varphi_c(N\sigma^\nu N^\dagger)_{a\dot{a}}\varepsilon^{\dot{a}\dot{b}}(N^*)_{\dot{b}}{}^{\dot{c}}\chi_{\dot{c}} = \Lambda^\mu{}_\nu\varphi^d(\sigma^\nu)_{d\dot{d}}\chi^{\dot{d}}.\end{aligned} \quad (2.65)$$

Thus, we have obtained the transformation law for the contravariant vector under the Lorentz transformation. Analogously, one can prove that the expression $\chi_{\dot{a}}(\tilde{\sigma}^\mu)^{\dot{a}a}\varphi_a$ is also the contravariant vector. As a result, we arrive at a prescription for how to construct Lorentz vectors from spinors and σ-matrices.

The matrices σ_μ, $\tilde{\sigma_\mu}$ possess many useful properties that can be established by direct calculations, e.g.,

$$\begin{aligned}&(\sigma_\mu\tilde{\sigma}_\nu + \sigma_\nu\tilde{\sigma}_\mu)_a{}^b = 2\eta_{\mu\nu}\delta_a{}^b, \qquad &&(\tilde{\sigma}_\mu\sigma_\nu + \tilde{\sigma}_\nu\sigma_\mu)^{\dot{a}}{}_{\dot{b}} = 2\eta_{\mu\nu}\delta^{\dot{a}}{}_{\dot{b}} \\ &\operatorname{tr}\sigma_\mu\tilde{\sigma}_\nu = 2\eta_{\mu\nu}, &&\sigma^\mu{}_{a\dot{a}}\tilde{\sigma}_\mu{}^{\dot{b}b} = 2\delta_a{}^b\delta_{\dot{a}}{}^{\dot{b}}.\end{aligned} \quad (2.66)$$

The matrices σ_μ, $\tilde{\sigma}_\mu$ make it possible to convert the vector indices into a pair of spinor indices, and vice versa. For example, if we have a vector t_μ, we can construct the object $t_{a\dot{a}} \sim (\sigma^\mu)_{a\dot{a}}t_\mu$. Consider an arbitrary spin tensor with an equal number of dotted and undotted indices, $\varphi_{a_1\ldots a_n\dot{a_1}\ldots\dot{a_n}}$. Then we can construct a tensor of rank n as

$$\varphi_{\mu_1\ldots\mu_n} = \sigma_{\mu_1}{}^{a_1\dot{a_1}}\sigma_{\mu_2}{}^{a_2\dot{a_2}}\ldots\sigma_{\mu_n}{}^{a_n\dot{a_n}}\varphi_{a_1\ldots a_n\dot{a_1}\ldots\dot{a_n}} \quad (2.67)$$

and vice versa,

$$\varphi_{a_1\ldots a_n\dot{a_1}\ldots\dot{a_n}} = \frac{1}{2^n}\,\tilde{\sigma}^{\mu_1}{}_{\dot{a_1}a_1}\ldots\tilde{\sigma}^{\mu_n}{}_{\dot{a_n}a_n}\,\varphi_{\mu_1\ldots\mu_n}. \quad (2.68)$$

For the spin tensors $\varphi_{a_1\dots a_n a \dot{a}_1\dots\dot{a}_n}$ and $\chi_{a_1\dots a_n \dot{a}_1\dots\dot{a}_n}{}^{\dot{a}}$, one can construct the quantities

$$\begin{aligned}\varphi_{\mu_1\dots\mu_n a} &= \sigma_{\mu_1}{}^{a_1\dot{a}_1}\dots\sigma_{\mu_n}{}^{a_n\dot{a}_n}\varphi_{a_1\dots a_n a\dot{a}_1\dots\dot{a}_n},\\ \chi_{\mu_1\dots\mu_n^{\dot{a}}} &= \sigma_{\mu_1}{}^{a_1\dot{a}_1}\dots\sigma_{\mu_n}{}^{a_n\dot{a}_n}\chi_{a_1\dots a_n\dot{a}_1\dots\dot{a}_n}{}^{\dot{a}}.\end{aligned} \tag{2.69}$$

A tensor with the spinor components

$$\psi_{\mu_1\dots\mu_n} = \begin{pmatrix}\varphi_{\mu_1\dots\mu_n}\\ \chi_{\mu_1\dots\mu_n}\end{pmatrix} \tag{2.70}$$

is called the Dirac tensor spinor.

The relations (2.56) and (2.60) make it possible to derive the generators of the Lorentz group in the fundamental and conjugate representations. In the vicinity of the unit element, the matrices $N \in SL(2|\mathbb{C})$ have the form $N = E+T$, or $N_a{}^b = \delta_a{}^b + T_a{}^b$, where E is the unit 2×2 matrix, and $T_a{}^b$ is a 2×2 matrix with infinitesimal elements. It is known that $\det N = 1 + \operatorname{tr} T$ is linear in T; therefore, $\operatorname{tr} T = 0$, since $\det N = 1$. The matrix T can be expanded in the basis σ_μ; however, since T is traceless, we can write $T = z_i\sigma_i$, where z_i are the tree complex numbers. Thus, in the vicinity of the unit element, each matrix N is parametrized by six real numbers, as it should be for the Lorentz transformations.

On the other hand, the elements of the Lorentz group representation in the vicinity of the unit element have the form $e^{-\frac{i}{2}\omega^{\alpha\beta}J_{\alpha\beta}}$, where $\omega^{\alpha\beta} = -\omega^{\beta\alpha}$ and $J_{\alpha\beta}$ are the Lorentz group generators in the given representation. We will need these generators in the fundamental and conjugate representations. For the fundamental representation in the infinitesimal vicinity of the unit element, we have $E + z_i\sigma_i = \mathbf{1} - \frac{i}{2}\omega^{\alpha\beta}J_{\alpha\beta}$, where, in the case under consideration, $\mathbf{1} = E$. Let us parameterize the complex numbers z_i as follows:

$$z_1 = -\frac{1}{2}(\omega^{01} + i\omega^{23}),\quad z_2 = -\frac{1}{2}(\omega^{02} + i\omega^{31}),\quad z_3 = -\frac{1}{2}(\omega^{03} + i\omega^{12}). \tag{2.71}$$

In this notation, $z_i\sigma_i = -\frac{i}{2}\omega^{\alpha\beta}(i\sigma_{\alpha\beta})$, where

$$(\sigma_{\alpha\beta})_a{}^b = \frac{1}{4}(\sigma_\alpha\tilde{\sigma}_\beta - \sigma_\beta\tilde{\sigma}_\alpha)_a{}^b\ . \tag{2.72}$$

Therefore, the generators of the Lorentz group in the fundamental representation are

$$J^{(F)}_{\alpha\beta} = i\sigma_{\alpha\beta}. \tag{2.73}$$

Here, (F) labels the fundamental representation. Similar consideration shows that the generators of the Lorentz group in the conjugate representation have the form

$$J^{(\bar{F})}_{\alpha\beta} = i\tilde{\sigma}_{\alpha\beta}, \tag{2.74}$$

$$\text{where}\qquad (\tilde{\sigma}_{\alpha\beta})^{\dot{a}}{}_{\dot{b}} = \frac{1}{4}(\tilde{\sigma}_\alpha\sigma_\beta - \tilde{\sigma}_\beta\sigma_\alpha)^{\dot{a}}{}_{\dot{b}}. \tag{2.75}$$

As a result, the transformations laws under infinitesimal Lorentz transformation for two-component spinors have the form

$$\delta\varphi = -\frac{i}{2}\omega^{\alpha\beta}(i\sigma_{\alpha\beta})\varphi, \qquad \delta\chi = -\frac{i}{2}\omega^{\alpha\beta}(i\tilde{\sigma}_{\alpha\beta})\chi. \tag{2.76}$$

The matrices $\sigma_{\alpha\beta}$ and $\tilde{\sigma}_{\alpha\beta}$ possesses useful properties:

$$\begin{aligned} &\sigma_{\alpha\beta} = -\sigma_{\beta\alpha}, \qquad \tilde{\sigma}_{\alpha\beta} = -\tilde{\sigma}_{\beta\alpha}, \\ &(\sigma_{\alpha\beta})^{ab} = (\sigma_{\alpha\beta})^{ba}, \qquad (\tilde{\sigma}_{\alpha\beta})^{\dot{a}\dot{b}} = (\tilde{\sigma}_{\alpha\beta})^{\dot{b}\dot{a}}. \end{aligned} \tag{2.77}$$

The generators of the Lorentz group that is the representation of arbitrary tensor spinors are found quite analogously by using the transformation law (2.62) and considering the infinitesimal forms of the matrices N and N^*. Also, we can derive the transformations of tensor spinors under non-homogeneous Lorentz transformations. We will not discuss this in detail but only formulate the final form. Consider an arbitrary spin tensor $\varphi_{a_1\dots a_m\dot{a}_1\dots\dot{a}_n} \equiv \varphi_{a(m)\dot{a}(n)}$. The variation of $\varphi_{a(m)\dot{a}(n)}(x)$ under the infinitesimal transformations $x'^\mu = x^\mu + \omega^\mu{}_\nu x^\nu + a^\mu$ is written as follows:

$$\delta\varphi_{a(m)\dot{a}(n)} = \Big[ia^\alpha (P_\alpha)_{a(m)\dot{a}(n)}{}^{b(m)\dot{b}(n)} - \frac{i}{2}\omega^{\alpha\beta}(J_{\alpha\beta})_{a(m)\dot{a}(n)}{}^{b(m)\dot{b}(n)}\Big]\varphi_{b(m)\dot{b}(n)}, \tag{2.78}$$

with Poincaré group generators $(P_\alpha)_{a(m)\dot{a}(n)}{}^{b(m)\dot{b}(n)}$ and $(J_{\alpha\beta})_{a(m)\dot{a}(n)}{}^{b(m)\dot{b}(n)}$, which can be found in an explicit form in the same way as we derived (2.76).

2.6 Irreducible representations of the Poincaré group

Irreducible representations of the Poincaré group determine relativistic physical systems with given mass and spin. These systems can be called elementary and are associated with elementary particles. The aim of this subsection is to formulate the basic notions and results concerning these systems. Let us note that we do not pretend to consider a complete description of the representations of the Poincaré group, which can be found in more specialized literature.

In section 2.3 we denoted the elements of the Poincaré group as (Λ, a). The multiplication rule in this group has the form (2.32), where the matrix Λ satisfies the basic relation (2.7). Let us denote the operators of some representation of the Poincaré group as $\mathcal{U}(\Lambda, a)$. According to the definition of a representation, we have

$$\mathcal{U}(\Lambda_1, a_1)\mathcal{U}(\Lambda_2, a_2) = \mathcal{U}((\Lambda_1, a_1)\cdot(\Lambda_2, a_2)) = \mathcal{U}(\Lambda_1\Lambda_2, \Lambda_1 a_2 + a_1). \tag{2.79}$$

As we know, the infinitesimal non-homogeneous Lorentz transformation is (2.33), with the parameters a^μ and $\omega^\mu{}_\nu$. Denoting the generators of the Poincaré group as P_α and $J_{\alpha\beta}$, in the vicinity of the unit element, one can write

$$\mathcal{U}(\Lambda, a) = e^{ia^\alpha P_\alpha - \frac{i}{2}\omega^{\alpha\beta}J_{\alpha\beta}}. \tag{2.80}$$

The operator P_α is the generator of spacetime translations, and the operator $J_{\alpha\beta}$ is the generator of Lorentz rotations. Examples of these generators for scalar and vector representations were given in section 2.3. Eqs. (2.79) and (2.80) allow one to derive, in a purely algebraic way, the commutation relations

$$
\begin{aligned}
&[P_\alpha, P_\beta] = 0, \\
&[P_\mu, J_{\alpha\beta}] = i(\eta_{\mu\alpha} P_\beta - \eta_{\mu\beta} P_\alpha), \\
&[J_{\mu\nu}, J_{\alpha\beta}] = i(\eta_{\mu\alpha} J_{\nu\beta} + \eta_{\nu\beta} J_{\mu\alpha} - \eta_{\mu\beta} J_{\nu\alpha} - \eta_{\nu\alpha} J_{\mu\beta}).
\end{aligned}
\tag{2.81}
$$

These equations define the Lie algebra of the Poincaré group, which is sometimes called the Poincaré algebra. The last line defines the Lie algebra of the Lorentz group.

One can show that the algebra (2.81) has two operators, C_1, C_2, which commute with all generators. These are the Casimir operators

$$
C_1 = P^\mu P_\mu \quad \text{and} \quad C_2 = W^\mu W_\mu\,, \tag{2.82}
$$

where W^μ is the Lubanski-Pauli vector

$$
W^\mu = -\frac{1}{2}\varepsilon^{\mu\nu\alpha\beta} P_\nu J_{\alpha\beta}. \tag{2.83}
$$

One can check that the invariants (2.82) satisfy the relations

$$
[P_\mu, C_1] = [P_\mu, C_2] = 0, \qquad [J_{\alpha\beta}, C_1] = [J_{\alpha\beta}, C_2] = 0, \qquad [C_1, C_2] = 0. \tag{2.84}
$$

One can distinguish two kinds of physically acceptable irreducible representations of the Poincaré group: massive and massless. The basis vectors $|p, m, s\rangle$ of a massive irreducible representation are defined by the following equations:

$$
\begin{aligned}
P_\mu |p, m, s\rangle &= p_\mu |p, m, s\rangle, \\
P^\mu P_\mu |p, m, s\rangle &= m^2 |p, m, s\rangle, \\
W^\mu W_\mu |p, m, s\rangle &= -m^2 s(s+1) |p, m, s\rangle.
\end{aligned}
\tag{2.85}
$$

The first two equations mean that $p^2 = p_0^2 - \mathbf{p}^2 = m^2$, which is the well-known relation (1.1). Thus, massive irreducible representations of Poincaré group are naturally associated with free massive relativistic particles. The parameter m plays a role in mass and should be positive. One can prove that the parameter s takes the values $s = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$. For each fixed s, there are $2s+1$ different vectors $|p, m, s\rangle$, and s is called the spin of a massive relativistic particle.

For the massless irreducible representation, the basis vectors of the state $|p, \lambda\rangle$ are defined by the equations

$$
P_\mu |p, \lambda\rangle = p_\mu |p, \lambda\rangle, \qquad P^\mu P_\mu |p, \lambda\rangle = 0, \qquad W^\mu W_\mu |p, \lambda\rangle = 0. \tag{2.86}
$$

The first two of these equations mean that $p^2 = p_0^2 - \mathbf{p}^2 = 0$. This is the relation between the energy and the three-dimensional momentum of a free massless relativistic particle. Thus, the massless irreducible representations of the Poincaré group are naturally associated with free massless relativistic particles. One can prove that, in the case under consideration, $W_\mu = \lambda P_\mu$, where λ takes the values $\lambda = 0, \pm\frac{1}{2}, \pm 1, \pm\frac{3}{2}, \pm 2, \ldots$.

This parameter is called the helicity of the massless elementary particle. Sometimes the value $|\lambda|$ is called the spin of a massless particle.

One can show that the Poincaré algebra can be realized in the linear space of the tensor fields $\varphi_{\mu_1\dots\mu_n}$ or the tensor spinor fields $\psi_{\mu_1\dots\mu_n}$, as defined by Eq. (2.70). The fields $\varphi_{\mu_1\dots\mu_n}$ or $\psi_{\mu_1\dots\mu_n}$, corresponding to the massive irreducible representations of the Poincaré algebra, are characterized by their masses and spin. The relativistic field with the given mass m and an integer spin $s = n$, is defined by the system of equations

$$\begin{aligned} &\varphi_{\mu_1\dots\mu_s}(x) = \varphi_{(\mu_1\dots\mu_n)}(x), &&(\Box + m^2)\varphi_{\mu_1\dots\mu_s}(x) = 0, \\ &\partial^{\mu_1}\varphi_{\mu_1\mu_2\dots\mu_s}(x) = 0, &&\eta^{\mu_1\mu_2}\varphi_{\mu_1\mu_2\mu_3\dots\mu_s}(x) = 0. \end{aligned} \tag{2.87}$$

Consider the following few examples:

$$1)\ s = 0 \quad (\Box + m^2)\varphi = 0, \tag{2.88}$$

$$2)\ s = 1 \quad (\Box + m^2)\varphi_\mu = 0, \quad \partial^\mu\varphi_\mu = 0, \tag{2.89}$$

$$3)\ s = 2 \quad (\Box + m^2)\varphi_{\mu\nu} = 0, \quad \partial^\mu\varphi_{\mu\nu} = 0, \ \ \eta^{\mu\nu}\varphi_{\mu\nu} = 0, \ \ \varphi_{\mu\nu}(x) = \varphi_{(\mu\nu)}. \tag{2.90}$$

In the last formula, $\varphi_{(\mu_1\dots\mu_n)}$ means a total symmetrization of the indices.

Eq. (2.88) is the Klein–Gordon equation (1.5), corresponding to a free massive scalar field. Eqs. (2.89) define the free massive vector field equations of motion. Eqs. (2.90) define the equation of motion of a massive symmetric second-rank tensor field. Equations for massless relativistic fields with integer spin can be obtained from Eqs. (2.87) at $m = 0$.

The relativistic fields with given mass m and given half-integer spin $s = n + \frac{1}{2}$ are defined in terms of the Dirac tensor spinors $\psi_{\mu_1\dots\mu_n}$ (2.70) by the system of equations

$$\begin{aligned} &\psi_{\mu_1\dots\mu_n} = \psi_{(\mu_1\dots\mu_n)}, &&(i\gamma^\mu\partial_\mu - m)\psi_{\mu_1\dots\mu_n} = 0, \\ &\partial^{\mu_1}\psi_{\mu_1\mu_2\dots\mu_n} = 0, &&\gamma^{\mu_1}\psi_{\mu_1\mu_2\dots\mu_n} = 0, \end{aligned} \tag{2.91}$$

where Dirac matrices γ^μ are defined in terms of the matrices σ^μ and $\tilde{\sigma}^\mu$, as defined in section 2.5,

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \tilde{\sigma}^\mu & 0 \end{pmatrix}, \tag{2.92}$$

satisfying the basic relation (also called the Clifford algebra)

$$\begin{aligned} \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu \ &= \begin{pmatrix} 0 & \sigma^\mu \\ \tilde{\sigma}^\mu & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^\nu \\ \tilde{\sigma}^\nu & 0 \end{pmatrix} + \begin{pmatrix} 0 & \sigma^\nu \\ \tilde{\sigma}^\nu & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^\mu \\ \tilde{\sigma}^\mu & 0 \end{pmatrix} \\ &= \begin{pmatrix} \sigma^\mu\tilde{\sigma}^\nu + \sigma^\nu\tilde{\sigma}^\mu & 0 \\ 0 & \tilde{\sigma}^\mu\sigma^\nu + \tilde{\sigma}^\nu\sigma^\mu \end{pmatrix} = \begin{pmatrix} 2\eta^{\mu\nu} & 0 \\ 0 & 2\eta^{\mu\nu} \end{pmatrix} = 2\eta^{\mu\nu} I, \end{aligned} \tag{2.93}$$

where I is the four-dimensional unit matrix.

Some examples of the construction described above, are

$$s = \frac{1}{2}, \quad (i\gamma^\mu\partial_\mu - m)\psi(x) = 0, \tag{2.94}$$

$$s = \frac{3}{2}, \quad (i\gamma^\mu\partial_\mu - m)\psi_\nu(x) = 0\,, \quad \text{with } \partial^\nu\psi_\nu(x) = 0, \ \ \gamma^\nu\psi_\nu(x) = 0. \tag{2.95}$$

Equation (2.94) is a free, massive, spin-$\frac{1}{2}$ equation of motion. It is called the Dirac equation. The four-component field ψ is called the Dirac spinor field, or the Dirac fermion. The relations (2.95) define the massive spin-$\frac{3}{2}$ field equation of motion, which is called the Rarita-Schwinger equation. Massless relativistic fields with half-integer spins are described by Eqs. (2.91) in the limit $m = 0$.

Exercises

2.1. Let G be the group of $n \times n$ matrices O such that $O^T O = E$, where T means transposition, and E is the unit matrix. Prove that this set forms a group where the multiplication law is the matrix product. This group is called the rotation group and is denoted $O(n)$. Consider a subset of the matrices from $O(n)$ with a unit determinant. Show that this subset is a subgroup of $O(n)$, called $SO(n)$.

2.2 Prove that the groups $SO(2)$ and $U(1)$ are isomorphic.

2.3. Consider the coordinate transformation $x'^{\mu} = f^{\mu}(x)$. Prove that the Eq. (2.3) leads to $f^{\mu}(x) = \Lambda^{\mu}{}_{\nu} x^{\nu} + a^{\mu}$, where the matrix $\Lambda^{\mu}{}_{\nu}$ satisfies the basic relation (2.7).

2.4. Let Λ be a matrix realizing the Lorentz transformation. Show that $\Lambda^{-1} = \eta^{-1}\Lambda^T\eta$.

2.5. Consider a set 5×5 of matrices of the form

$$\begin{pmatrix} \Lambda^{\mu}{}_{\nu} & a^{\mu} \\ 0 & 1 \end{pmatrix}.$$

Show that the group of these matrices is a realization of the Poincaré group.

2.6. Consider a matrix $A = E + \alpha X$, where α is an infinitesimal parameter, and $\det A = 1$. Show that $\operatorname{tr} X = 0$.

2.7. Let $A \in SO(n)$ and $A = E + \alpha X$, where α is an infinitesimal real parameter. Show that $X^T = -X$.

2.8. Let $A \in SU(n)$ and $A = E + i\alpha X$, where α is an infinitesimal real parameter. Show that $X^{\dagger} = X$.

2.9. Consider the Lie algebra with generators T_i, where $[T_i, T_j] = i f_{ijk} T_k$ and the structure constants are totally antisymmetric. Prove that $C = T_i T_i$ is the Casimir operator.

2.10. Prove that the vector product in the three-dimensional Euclidean space possesses all of the properties of the multiplication law in the Lie algebra.

2.11. Consider a phase space of some dynamical system with phase coordinates q^i, p_i. Assuming that $f(q,p)$ and $g(q,p)$ are functions on the phase space, prove that the Poisson bracket

$$\{f, g\} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} \right)$$

possesses all of the properties of the multiplication law in the Lie algebra.

2.12. Prove that the commutator of the operators possesses all of the properties of the multiplication law in the Lie algebra.

2.13. Let A be an $n \times n$ matrix and $\det A \neq 0$. Prove that $\det A = e^{\mathrm{tr}\,\log A}$.

2.14. Let $J^{\alpha\beta} = x^\alpha P^\beta - x^\beta P^\alpha + S^{\alpha\beta}$. Show that $W_\alpha = -\frac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}P^\beta S^{\gamma\delta}$.

2.15. Calculate the commutator $[J_{\alpha\beta}, J_{\mu\nu}]$ in the scalar representation.

2.16. Calculate the commutator $[S_{\alpha\beta}, S_{\mu\nu}]$, where $S_{\alpha\beta}$ are defined by (2.48).

2.17. Show that all matrices σ_μ are Hermitian.

2.18. Prove the identity $\sigma_i\sigma_j = \sigma_0\delta_{ij} + i\varepsilon_{ijk}\sigma_k$.

2.19. Prove the identity $(\sigma\mathbf{n_1})(\sigma\mathbf{n_2}) = (\mathbf{n_1}\mathbf{n_2}) + i\big(\mathbf{n_1} \times \mathbf{n_2}\big)\sigma$.

2.20. Prove the identity $\mathrm{tr}\,(\tilde{\sigma}_\mu\sigma_\nu) = 2\eta_{\mu\nu}$.

2.21. Prove that the matrices $i\sigma_{\mu\nu}$ satisfy the commutation relation for the generators of Lorentz rotations.

2.22. Let $v_{a\dot{a}} = (\sigma^\mu)_{a\dot{a}}v_\mu$. Find the v_μ from this relation.

2.23. Prove the relation $W^2 = -\frac{1}{2}\big(S_{\beta\gamma}S^{\beta\gamma}P^2 + S_{\beta\gamma}S^{\alpha\beta}P_\alpha P^\gamma + S_{\beta\gamma}S^{\gamma\alpha}P_\alpha P\beta\big)$.

2.24. Consider $W^2\varphi^\mu(x)$, where $\varphi^\mu(x)$ is an arbitrary vector field. Using the result of the previous exercise and the relation (2.48), formulate the conditions for $W^2\varphi^\mu(x) = -2m^2\varphi^\mu(x)$. Explain on the basis of the last relation how the spin of a massive vector field is equal to 1.

Comments

There are many excellent books on special relativity, e.g., the eminent book by L.D. Landau and E.M. Lifshitz [202] as well as the ones by W. Rindler [264], P.M. Schwarz and J.H. Schwarz [275] and G.L. Naber [228]. Many details of special relativity are usually considered in books on general relativity and gravitation, e.g., in the books by S. Weinberg [340], C.W. Misner, K.S. Thorne and J.A. Wheeler [219] and J.B. Hartle [175].

Group theory for physicists is considered in many books, e.g., in those by W. K. Tung [320], A.O. Barut, R. Raczka [29] and P. Ramond [257]. There are also excellent lecture notes on group theory for physicists available on-line (see e.g., the notes by H. Osborn [236]).

Representations of Lorentz and Poincaré groups are considered with different levels of detail in, e.g., the books by I.M. Gelfand, R.A. Minlos and Z.Ya. Shapiro [156], W. K. Tung [320], A.O. Barut and R. Raczka [29], I.L. Buchbinder and S.M. Kuzenko [81].

Group theory is considered in the physical context in books on quantum field theory, e.g., S. Gasiorowicz [152], S. Schweber [276] and S. Weinberg [345]. Our considerations here followed, in a simplified form, those in [81].

3

Lagrange formalism in field theory

In this chapter, we briefly present the minimal amount of information required about classical fields for the subsequent treatment of quantum theory in the rest of the book.

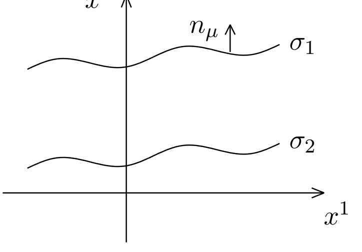
3.1 The principle of least action, and the equations of motion

Consider Minkowski space with the coordinates $x = x^\mu$. As we already know, the function of coordinates $\phi \equiv \phi(x)$ defined in Minkowski space is called a field. The field can be real or complex, one component or multi component and it can be a scalar, a tensor or a spinor. In particular, this means that the field can have various indices. Then it can be written as $\phi^i(x)$, where i is a set of all indices (tensor, spinor or any other). Usually, we will not use indices if there is no special reason to do so.

It is supposed that the dynamics of the field ϕ is described in terms of the action functional $S = S[\phi]$. It is postulated that the action has the following form:

$$S = \int_\Omega d^4x \, \mathcal{L} \, . \tag{3.1}$$

Here Ω is a domain in Minkowski space bounded by two space-like hypersurfaces, $\sigma(x) = \sigma_1$ and $\sigma(x) = \sigma_2$, as shown in the figure below. Remember that a hypersurface is called space-like if its normal vector $n_\mu(x) = \frac{\partial \sigma(x)}{\partial x^\mu}$ is time-like at any point x^μ, i.e., $n_\mu n^\mu > 0$. In this case, there is an inertial reference frame such that these two hypersurfaces are written as $t = t_1$, and $t = t_2$, where t is a time coordinate.



Usually, it is assumed that the domain Ω coincides with the whole Minkowski space, meaning $t_1 \to -\infty$ and $t_2 \to \infty$. It is postulated that the function $\mathcal{L}$ in (3.1) is a real scalar field under the Lorentz transformations. This guarantees that the action $S[\phi]$ is a real Lorenz invariant.

It is generally assumed that the model of field theory is defined when the set of fields $\phi^i(x)$ and the function $\mathcal{L}$ are specified.

Introduction to Quantum Field Theory with Applications to Quantum Gravity. Iosif L. Buchbinder and Ilya L. Shapiro, Oxford University Press (2021).
DOI: 10.1093/oso/9780198838319.003.0003

The function $\mathcal{L}$ is called Lagrangian. In the special reference frame described above, the action (3.1) can be presented in the form

$$S = \int_{t_1}^{t_2} dt \int d^3x\mathcal{L} \,=\, \int_{t_1}^{t_2} dt\, L,$$

where $L = \int d^3x\mathcal{L}$ is the Lagrange function, similar to the one in classical mechanics. In this framework, the Lagrangian is nothing else but the density of the Lagrange function.

Let us discuss the analogy with classical mechanics. The field $\phi(x) = \phi^i(x)$ can be regarded as $\phi^i(t, \mathbf{x}) \equiv \phi^i{}_{\mathbf{x}}(t)$, with space coordinates $\mathbf{x}$ playing the role of indices. Thus, one can understand a relativistic field as a mechanical system with generalized coordinates $\phi^i{}_{\mathbf{x}}(t)$, characterized by discrete indices i and by the three-dimensional vector $\mathbf{x}$. This means that we can consider a field as a system with an infinite (continuous) number of degrees of freedom.

The Lagrangian is postulated to be a real function of the field and of its spacetime derivatives taken at the same spacetime point x^μ,

$$\mathcal{L} \,=\, \mathcal{L}(\phi(x), \partial_\mu\phi(x),\, \partial_{\mu_1}\partial_{\mu_2}\phi(x),\, \dots, \partial_{\mu_1}\partial_{\mu_2}\dots\partial_{\mu_n}\phi(x)). \tag{3.2}$$

Usually, it is assumed that the Lagrangian includes only first derivatives of the field. However, there are models containing derivatives of order higher than first. Such models are called higher-derivatives theories. We will mainly consider the models with only the first derivatives of the fields in Lagrangian, at least in Part I of this book.

The integral in (3.1) is convergent if it requires that $\mathcal{L} \to 0$ at the space infinity, when $|\mathbf{x}| \to \pm\infty$. In most cases, it is sufficient to consider that $\phi(x) \to 0$ at $|\mathbf{x}| \to \pm\infty$. To get the convergent integral in (3.1), when Ω coincides with the whole Minkowski space, we demand that $\phi(x) \to 0$ at $t \to \pm\infty$. As a result, one gets the standard boundary conditions $\phi(x) \to \infty$ at $x^\mu \to \pm\infty$. However, in some cases, we need to deal with theories of fields that are defined on some domains in Minkowski space, which are bounded in the space directions. In this case, the boundary conditions for the field require a special consideration.

Field dynamics is defined by the least action principle: physically admissible configurations correspond to the minimum of the action. The mathematical formulation of this principle is as follows. Let $\phi^i(x)$ be some field and $\phi'^i(x)$ be another field with the same set of indices. The difference $\delta\phi^i(x) = \phi'^i(x) - \phi^i(x)$ is called a field variation. We assume that the difference $S[\phi + \delta\phi] - S[\phi]$ can be represented in the form

$$S[\phi + \delta\phi] - S[\phi] \,=\, \int_\Omega d^4x\, A(x)\, \delta\phi(x) + \dots\,,$$

where the dots mean the terms with higher than the first power of $\delta\phi$. The expression

$$\delta S[\phi] \,=\, \int_\Omega d^4x A(x)\, \delta\phi(x) \tag{3.3}$$

is called a variation of the functional $S[\phi]$. The function $A(x)$ is called a variational or functional derivative and it is denoted $\frac{\delta S[\phi]}{\delta\phi(x)}$. Hence, we get

$$\delta S[\phi] \;=\; \int_\Omega d^4x \; \frac{\delta S[\phi]}{\delta \phi^i(x)} \, \delta\phi^i(x). \tag{3.4}$$

Let us use the following theorem from the variational calculus: if the field $\phi(x)$ corresponds to an extremum of the functional $S[\phi]$, then the corresponding variation $\delta S[\phi] = 0$ for any $\delta\phi^i(x)$. Since $\delta\phi^i(x)$ is arbitrary, Eq. (3.4) leads to

$$\frac{\delta S[\phi]}{\delta \phi^i(x)} = 0. \tag{3.5}$$

Eq. (3.5) is called a classical equation of motion or simply the field equation. The solutions of this equation determine the physically admissible field configurations $\phi^i(x)$.

Consider the calculation of the variational derivative of the functional $S[\phi]$ (3.1), assuming that $\mathcal{L} = \mathcal{L}(\phi, \partial_\mu\phi)$. In this case, we have

$$S[\phi] \;=\; \int_\Omega d^4x \, \mathcal{L}\big(\phi(x), \partial_\mu\phi(x)\big). \tag{3.6}$$

Here Ω is a domain in Minkowski space, bounded by $\partial\Omega$, that consists of the space-like hypersurfaces σ_1 and σ_2, where the hypersurface σ_2 lies in the future, relative to the hypersurface σ_1. Let $\phi(x)\Big|_{\sigma_1} = \phi_1(\mathbf{x})$, and $\phi(x)\Big|_{\sigma_2} = \phi_2(\mathbf{x})$, where $\phi(x)\Big|_{\sigma} = \phi(x)\Big|_{x\in\sigma}$, $\sigma = \sigma(x)$ is a space-like hypersurface and $\mathbf{x}$ is the vector formed by independent three-dimensional coordinates on this hypersurface. Let $\phi(x)$ be the field corresponding to an extremum of the functional $S[\phi]$, and $\phi'(x)$ an arbitrary field. We assume that both $\phi(x)$ and $\phi'(x)$ satisfy the same boundary conditions. Then, for the variation $\delta\phi(x)$=$\phi'(x)$-$\phi(x)$, we get $\delta\phi(x)\big|_{\sigma_1} = \delta\phi(x)\big|_{\sigma_2} = 0$. On the top of that, we assume $\delta\phi \to 0$ at $x^i \to \pm\infty$.

Consider

$$S[\phi + \delta\phi] - S[\phi] = \int_\Omega d^4x \big\{\mathcal{L}\big(\phi + \delta\phi, \partial_\mu\phi + \partial_\mu\delta\phi\big) - \mathcal{L}\big(\phi, \partial_\mu\phi\big)\big\}. \tag{3.7}$$

Since the field variation is not related to the change of coordinates, $\partial_\mu\delta\phi = \partial_\mu\phi' - \partial_\mu\phi = \delta\partial_\mu\phi(x)$. Expanding the integral in (3.7) in the Taylor series in $\delta\phi$ up to the first order, we arrive at

$$\begin{aligned}
\delta S[\phi] &= \int_\Omega d^4x \Big\{ \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\partial_\mu\delta\phi \Big\} \\
&= \int_\Omega d^4x \Big\{ \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\delta\phi\Big) - \Big(\partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\Big)\delta\phi \Big\} \\
&= \int_\Omega d^4x \Big[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\Big)\Big]\delta\phi + \int_{\partial\Omega} d\sigma_\mu \Big(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\Big)\delta\phi\,.
\end{aligned} \tag{3.8}$$

Here $d\sigma_\mu$ is an element of the surface $\partial\Omega$, and the Gauss theorem has been used. According to the boundary conditions, $\delta\phi \to 0$ on $\partial\Omega$. Therefore,

$$\delta S[\phi] = \int_\Omega d^4x \Big[\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\Big)\Big]\delta\phi(x) = 0, \tag{3.9}$$

and hence the variation of the action has the form (3.4). Thus, the functional derivative of the action is

$$\frac{\delta S[\phi]}{\delta\phi(x)} = \frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\Big), \tag{3.10}$$

and the equations of motion take the form of the Lagrange equations for the field ϕ,

$$\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\Big) = 0. \tag{3.11}$$

Two observations are in order. First of all, the Lagrangian $\mathcal{L}$ is not uniquely defined. For instance, the two Lagrangians $\mathcal{L}$ and $\mathcal{L} + \partial_\mu R^\mu(\phi)$ lead to the same equations of motion (3.11).

The second observation is that, in some field models, we need to define a field ϕ in the space with boundaries, assuming that at least some of the space coordinates x^i take their values in the finite domains. Then the equations of motion (3.11) take place under the boundary conditions of the modified form. Usually these boundary conditions are defined by the requirement $\delta\phi\big|_{\partial\Omega} = 0$, and then the variation $\delta S[\phi]$ has the standard form (3.3).

3.2 Global symmetries

Following the analogy with classical mechanics, let us explore global symmetries of the Lagrangian approach for the fields, and its relation with the conservation laws.

Consider a theory of the fields $\phi = \phi^i(x)$ with the action (3.6). The infinitesimal transformations of coordinates and fields

$$x'^\mu = x^\mu + \delta x^\mu, \tag{3.12}$$
$$\phi'^i(x') = \phi^i(x) + \Delta\phi^i(x) \tag{3.13}$$

are symmetry transformations if they leave the action invariant, i.e.,

$$S[\phi] \;=\; S'[\phi'], \tag{3.14}$$

or, in the detailed form,

$$\int_\Omega d^4x\,\mathcal{L}\big(\phi(x), \partial_\mu\phi(x)\big) \;=\; \int_{\Omega'} d^4x'\,\mathcal{L}\big(\phi'(x'), \partial'_\mu\phi'(x')\big), \tag{3.15}$$

where Ω' is a domain of integration in terms of coordinates x'^μ. The last means that the equations defining Ω' are obtained from the equations defining Ω by the coordinate transformations $x^\mu = x'^\mu - \delta x^\mu$.

The transformations (3.13) can be reformulated as follows. Rewriting Eq. (3.13) as

$$\phi'^i(x+\delta x) = \phi'^i(x) + \partial_\mu\phi^i(x)\delta x^\mu = \phi^i(x) + \Delta\phi^i(x), \tag{3.16}$$

we arrive at

$$\Delta\phi^i(x) = \phi'^i(x) - \phi^i(x) + \partial_\mu\phi^i\delta x^\mu.$$

The quantity $\phi'^i(x) - \phi^i(x) = \delta\phi^i(x)$, where $\delta\phi^i(x)$ is a field variation, separated from the variation of the independent coordinates. Therefore,

$$\Delta\phi^i(x) = \delta\phi^i(x) + \partial_\mu\phi^i(x)\delta x^\mu. \tag{3.17}$$

One can assume that the transformations (3.13) are characterized by a finite set of parameters $\xi^1, \xi^2, \dots, \xi^N$, such that

$$\begin{aligned} \delta x^\mu &= X^\mu{}_I(x)\xi^I, \\ \delta\phi^i(x) &= Y^i{}_I\big(x, \phi(x), \partial_\mu\phi(x)\big)\xi^I, \end{aligned} \tag{3.18}$$

where $I = 1, 2, ..., N$ and there is summation over the index I. The transformations (3.18) are called the N-parametric global transformations. The term global means that the parameters ξ^I are coordinate-independent. In the opposite case, the transformations are called local, as it is the case for the gauge transformations. For now, we consider only global transformations.

Taking into account Eq. (3.18), one can rewrite Eq. (3.13) as follows,

$$\begin{aligned} &\phi'^i(x') = \phi^i(x) + \Delta\phi^i(x), \qquad \text{where} \\ &\Delta\phi^i(x) = \big[Y^i{}_I\big(x, \phi(x), \partial_\mu\phi(x)\big) + \partial_\mu\phi^i(x)X^\mu{}_I(x)\big]\xi^I. \end{aligned} \tag{3.19}$$

Thus, in order to specify the global symmetry transformations, one should identify a field model and define the functions $X^\mu{}_I(x)$ and $Y^i{}_I\big(x, \phi(x), \partial_\mu\phi(x)\big)$.

The global symmetry transformations can be classified into spacetime transformations and internal symmetry transformations. In the last case, $\delta x^\mu = 0$ and $Y^i{}_I = Y^i{}_I\big(\phi(x)\big)$. This means that the internal symmetry transformations are transformations of the fields with fixed coordinates.

Spacetime symmetry transformations. Consider this type of symmetry transformations by dealing with two important examples.

As we already mentioned above, the Lagrangian should be a scalar under the Lorentz transformations, i.e., the action must be Lorentz invariant. The infinitesimal Lorentz transformations of coordinates have the form

$$\delta x^\mu = \omega^\mu{}_\nu x^\nu, \tag{3.20}$$

where $\omega_{\mu\nu} = -\omega_{\nu\mu}$ are the transformation parameters.

Usually, it is assumed that the fields are tensors or spinors under the Lorentz transformations. The reason for this is that, for such types of fields, it is easy to control the transformation rules, including constructing the Lagrangian as a Lorentz scalar. In the most general case, one can assume that all the fields are spin tensors $\phi_A(x)$, where A is a collection of tensor and spinor indices. Then, under Lorentz transformations, the fields ϕ_A transform as follows:

$$\delta\phi_A = -\frac{i}{2}\omega^{\alpha\beta}(J_{\alpha\beta})_A{}^B\phi_B, \tag{3.21}$$

with the generators of the Lorentz transformations $(J_{\alpha\beta})_A{}^B$ in the corresponding representation. The transformations, (3.20) and (3.21), represent an example of transformations related to the spacetime symmetry.

Another example is the symmetry under spacetime translations $x'^\mu = x^\mu + a^\mu$, which is equivalent to a shift of the reference frame as a whole. Since all points of the Minkowski space are physically equivalent, a transformed field in a transformed point must coincide with the initial field in the initial point, which is $\phi'^i(x') = \phi^i(x)$, and the action $S[\phi]$ should be invariant. Then $\phi'^i(x) + \partial_\mu \phi^i a^\mu = \phi'(x)$ or $\delta\phi^i(x) = -\partial_\mu \phi^i(x) a^\mu$. This transformation, together with $x'^\mu = x^\mu + a^\mu$, gives us another example of a spacetime symmetry transformation.

Internal symmetries. Consider the field $\phi^r{}_A(x)$, where A is a set of all Lorentz indices (tensor or spinor) and $r = 1, 2, \dots, n$. The field $\phi^r{}_A(x)$ can be treated as a vector in some n-dimensional linear space. Let this vector transform according to

$$\begin{aligned}\phi'^r{}_A(x) &= \phi^r{}_A(x) + \delta\phi^r{}_A(x),\\ \delta\phi^r{}_A(x) &= i(T^I)^r{}_s \phi^s{}_A(x)\xi^I,\end{aligned} \tag{3.22}$$

where ξ^I are constant parameters, $I = 1, 2, \dots, N$. The matrices $(T^I)^r{}_s$ are assumed to satisfy the condition

$$[T^I, T^J] = i f^{IJ}{}_K T^K, \tag{3.23}$$

where $f^{IJ}{}_K = -f^{JI}{}_K$, and $f^{IJ}{}_K$ are constants. The relations (3.22) and (3.23) mean that we have a representation of some Lie group, in the linear space of vectors $\phi^r{}_A$. The matrices $(T^I)^r{}_s$ are generators of the group, and the quantities $f^{IJ}{}_K$ are the corresponding structure constants, while $\xi^1, \xi^2, \dots, \xi^N$ are the group parameters. One can compare this relation to Eq. (2.22).

If the action $S[\phi]$ is invariant under the transformations (3.22), the last correspond to internal symmetry. Later on, we shall discuss the concrete examples of internal symmetries.

3.3 Noether's theorem

Noether's theorem gives a general method for finding the conserved quantities corresponding to the symmetries of the theory. Additive conserving quantities are also called dynamical invariants.

Theorem. Each N-parameters continuous symmetry transformation corresponds to N dynamical invariants.

Proof. Consider the transformations (3.12), (3.13) in the form (3.18), and assume the action is invariant, $\delta S[\Phi] = 0$. The variation of action $\delta S[\phi] = S'[\phi'] - S[\phi]$ follows from

$$S[\phi] = \int\limits_\Omega d^4x \mathcal{L}(\phi(x), \partial_\mu \phi(x)) \quad \text{and} \quad S'[\phi'] = \int\limits_{\Omega'} d^4x' \mathcal{L}(\phi'(x'), \partial'_\mu \phi'(x')). \tag{3.24}$$

Making the change of variables $x'^\mu = x^\mu + \delta x^\mu$ in the *r.h.s.* of the expression for $S'[\phi']$, the domain Ω' is transformed into the domain Ω. In the linear order in δx^μ, the Jacobian of this change of variables is

$$\det\left(\frac{\partial x'^{\mu}}{\partial x^{\nu}}\right) \;=\; \det\left(\delta^{\mu}{}_{\nu}+\partial_{\nu}\delta x^{\mu}\right) \;=\; 1+\frac{\partial\delta x^{\mu}}{\partial x^{\mu}}. \tag{3.25}$$

Here we have taken into account that all the terms out of the main diagonal of the matrix in the *l.h.s.* enter the determinant in at least second order in ξ^I, and therefore these terms are irrelevant for the first-order expression. Thus, we get

$$S'[\phi'] \;=\; \int_{\Omega} d^4x\,(1+\partial_{\mu}\delta x^{\mu})\,\mathcal{L}\big(\phi'(x+\delta x),\;\frac{\partial x^{\nu}}{\partial x'^{\mu}}\partial_{\nu}\phi'(x+\delta x)\big). \tag{3.26}$$

According to the definition of symmetry transformations, $\phi'(x+\delta x)=\phi(x)+\Delta\phi(x)$. Furthermore, in the lowest order,

$$\frac{\partial x^{\nu}}{\partial x'^{\mu}}=\delta^{\nu}{}_{\mu}-\partial_{\mu}\delta x^{\nu} \qquad \text{and} \qquad \frac{\partial x'^{\mu}}{\partial x^{\alpha}}=\delta^{\mu}{}_{\alpha}+\partial_{\alpha}\delta x^{\mu}. \tag{3.27}$$

Finally,

$$S'[\phi'] \;=\; \int\limits_{\Omega} d^4x\,(1+\partial_{\mu}\delta x^{\mu})\,\mathcal{L}\big(\phi+\Delta\phi,(\delta^{\nu}{}_{\mu}-\partial_{\mu}\delta x^{\nu})(\partial_{\nu}\phi+\partial_{\nu}\Delta\phi)\big). \tag{3.28}$$

The Lagrangian in the *r.h.s.* of the last expression can be expanded into the power series up to the linear terms in $\Delta\phi$ and δx^{ν}, to give

$$\begin{aligned}
S'[\phi'] &= \int\limits_{\Omega} d^4x\Big\{\mathcal{L}(\phi,\partial_{\mu}\phi)\,\big[1+\partial_{\mu}\delta x^{\mu}\big]+\frac{\partial\mathcal{L}}{\partial\phi^i}\Delta\phi^i+\frac{\partial\mathcal{L}}{\partial\partial_{\mu}\phi^i}\,\big(\partial_{\mu}\Delta\phi^i-\partial_{\nu}\phi^i\partial_{\mu}\delta x^{\nu}\big)\Big\}\\
&= \int\limits_{\Omega} d^4x\Big\{\mathcal{L}+\mathcal{L}\partial_{\mu}\delta x^{\mu}+\frac{\partial\mathcal{L}}{\partial\phi^i}\delta\phi^i+\frac{\partial\mathcal{L}}{\partial\partial_{\mu}\phi^i}\partial_{\mu}\delta\phi^i\,\frac{\partial\mathcal{L}}{\partial\phi^i}\partial_{\mu}\phi^i\delta x^{\mu}+\frac{\partial\mathcal{L}}{\partial\partial_{\nu}\phi^i}\partial_{\nu}\partial_{\mu}\phi^i\delta x^{\mu}\Big\}\\
&+\int\limits_{\Omega} d^4x\,\Big\{\mathcal{L}+\mathcal{L}(\partial_{\mu}\delta x^{\mu})+(\partial_{\mu}\mathcal{L})\delta x^{\mu}+\frac{\partial\mathcal{L}}{\partial\phi^i}\delta\phi^i+\frac{\partial\mathcal{L}}{\partial\partial_{\mu}\phi^i}\partial_{\mu}\delta\phi^i\Big\},
\end{aligned} \tag{3.29}$$

where we used the feature of coordinate-independent variations, $\delta\partial_{\mu}\phi^i=\partial_{\mu}\delta\phi^i$. Using the equations of motion (3.11) in the expression (3.29), we obtain

$$S'[\phi'] \;=\; \int\limits_{\Omega} d^4x\,\left\{\mathcal{L}+\partial_{\mu}(\mathcal{L}\delta x^{\mu})+\partial_{\mu}\left(\frac{\partial\mathcal{L}}{\partial\partial_{\mu}\phi^i}\delta\phi^i\right)\right\}. \tag{3.30}$$

Next, substituting (3.18) and (3.19) into the last expression yields, for the first variation of the action,

$$\begin{aligned}
\delta S[\phi] = S'[\phi']-S|\phi| &= \int\limits_{\Omega} d^4x\partial_{\mu}\left\{\frac{\partial\mathcal{L}}{\partial\partial_{\mu}\phi^i}\delta\phi^i+\mathcal{L}\delta x^{\mu}\right\}\\
&= \int\limits_{\Omega} d^4x\;\partial_{\mu}\Big\{\frac{\partial\mathcal{L}}{\partial\partial_{\mu}\phi^i}Y^i{}_I+\mathcal{L}X^{\mu}{}_I\Big\}\xi^I \;=\; 0,
\end{aligned} \tag{3.31}$$

since (3.18) corresponds to the symmetry transformations. Then, taking into account that the parameters ξ^I are linear independent and an arbitrariness of the domain Ω, one gets

$$\partial_\mu J^\mu_{\ I} = 0, \quad \text{for} \quad I = 1, 2, \ldots, N\,, \tag{3.32}$$

where we have introduced the notation for the Noether's current, or generalized current,

$$J^\mu_{\ I} \;=\; -\Big(\frac{\partial\,\mathcal{L}}{\partial\,\partial_\mu\phi^i}\,Y^i_{\ I} + \mathcal{L}\,X^\mu_{\ I}\Big). \tag{3.33}$$

The relation (3.32) is the local conservation law of generalized current. Starting from this identity, we can get the integrated form of the conservation law. Using the Gauss theorem, one gets

$$\int_{\partial\Omega} d\sigma_\mu\, J^\mu_{\ I} = 0. \tag{3.34}$$

Since the fields $\phi^i(x)$ are vanishing at the space infinity, the identity (3.34) implies that

$$\int_{\sigma_2} d\sigma_\mu\, J^\mu_{\ I} - \int_{\sigma_1} d\sigma_\mu\, J^\mu_{\ I} \;=\; 0, \tag{3.35}$$

where the change of sign is stipulated by the change of direction of the vector n_μ normal to the hypersurface $\sigma(x) = \sigma_1$.

One can introduce the functionals depending on the hypersurface σ,

$$C_I[\sigma] = \int_\sigma d\sigma_\mu\, J^\mu_{\ I}, \qquad\qquad I = 1, 2, \ldots, N. \tag{3.36}$$

Then Eq. (3.35) gives, for the two space-like hypersurfaces σ_1 and σ_1,

$$C_I[\sigma_1] \;=\; C_I[\sigma_2], \qquad I = 1, 2, \ldots, N. \tag{3.37}$$

Thus, the functionals $C_I[\sigma]$ do not depend on choice of the hypersurface σ, such that $C_I[\sigma] = const$. Choosing a constant time hypersurface, $\sigma(x) = t$, we arrive at

$$C_I[t_2] \;=\; C_I[t_1], \qquad \text{where} \qquad C_I[t] \;=\; \int d^3x\, J^0_{\ I}(x). \tag{3.38}$$

The conditions $C_I[\sigma] = const$ mean that the functionals (3.36) are conserved quantities. Thus, we have shown that there is a conservation law for each continuous symmetry transformation with a fixed parameter ξ^I. Since the functionals $C_I[\sigma]$ are given by the integrals over hypersurfaces, they are additive quantities, which completes the proof.

Remark 1. Since Eq. (3.37) was derived using the equations of motion, the quantities $C_I[\sigma]$ in (3.37) are conserved only *on shell* (or on the mass shell), when the fields $\phi^i(x)$ are solutions to the equations of motion.

Remark 2. It s important that the generalized current is defined in a non-unique way. For instance, let $J^\mu{}_I$ be a generalized current. One can introduce the quantity

$$\tilde{J}^\mu_I \,=\, J^\mu_I \,+\, \partial_\nu f^{\mu\nu}{}_I, \tag{3.39}$$

where $f^{\mu\nu}{}_I = -f^{\nu\mu}{}_I$ is an arbitrary function of the fields and their derivatives, that is antisymmetric in the indices μ and ν. Obviously,

$$\partial_\nu \tilde{J}^\mu_I \,=\, \partial_\mu J^\mu_I \,+\, \partial_\mu\partial_\nu f^{\mu\nu}{}_I \,=\, \partial_\mu J^\mu_I. \tag{3.40}$$

In other words, if $\partial_\mu J^\mu{}_I = 0$, then $\partial_\mu \tilde{J}^\mu{}_I = 0$ too. Thus, the local conservation law for the current (3.32) does not change under the modification of the current (3.39). Consequently, the dynamical invariants $C_I[\sigma]$ remain conserved quantities under the same operation. This arbitrariness can be used to impose additional conditions on the generalized current.

As an application of the general Noether's theorem, consider the conservation law corresponding to internal symmetries, when $\delta x^\mu = 0$. According to (3.18) and (3.22),

$$X^\mu_I \,=\, 0, \qquad Y^r_I \,=\, i(T^I)^r{}_s\phi^s. \tag{3.41}$$

Then, the Noether's current is

$$J^\mu_I \,=\, -\,i\,\frac{\partial\mathcal{L}}{\partial\,\partial_\mu\phi^r}\,(T^I)^r{}_s\phi^s\,. \tag{3.42}$$

The conserving quantities associated with internal symmetries are called charges and are denoted as Q^I. Using (3.36) and (3.42), one gets

$$Q^I \,=\, \int d^3x\, J^0{}_I \,=\, -i\int d^3x\,\frac{\partial\mathcal{L}}{\partial\,\dot\phi^r}\,(T^I)^r{}_s\phi^s. \tag{3.43}$$

3.4 The energy-momentum tensor

One of the most important conservation laws is related to the invariance under the spacetime translations $x'^\mu = x^\mu + a^\mu$, where a^μ is an arbitrary constant four-vector. It is clear that if the Lagrangian does not depend explicitly on the coordinates, the spacetime translations are the symmetry transformations. In this case,

$$\delta x^\mu = a^\mu = \delta^\mu{}_\nu a^\nu, \tag{3.44}$$

i.e., $X^\mu_I = \delta^\mu_\nu$ in the general relation $\delta x^\mu = X^\mu{}_I\xi^I$, and a^μ plays the role of parameters ξ^I. As we already know, the field transforms under translation as

$$\delta\phi^i \,=\, -\partial_\mu\phi^i a^\mu \,=\, -\delta^\mu{}_\nu\partial_\mu\phi^i a^\nu. \tag{3.45}$$

It means that the role of the function $Y^i{}_I$ in the general relation $\delta\phi^i = Y^i{}_I\xi^I$ is played by $-\delta^\mu{}_\nu\partial_\mu\phi^i$. The generalized current, corresponding to the symmetry under the spacetime translations described above, is called the canonical energy-momentum tensor and is denoted as $T^\mu{}_\nu$. Let us note in passing that in Part II we shall introduce another (dynamical) definition of the energy-momentum tensor. The two definitions are equivalent in all known cases, but the general proof of this fact is not known yet.

The expression for generalized current (3.33), in the case of (3.44), yields

$$T^{\mu}{}_{\nu} = \frac{\partial \mathcal{L}}{\partial \partial_\mu \phi^i} \partial_\nu \phi^i - \mathcal{L} \delta^{\mu}{}_{\nu}, \tag{3.46}$$

while the local conservation law has the form

$$\partial_\mu T^{\mu}{}_{\nu} = 0. \tag{3.47}$$

The dynamical invariants corresponding to the symmetry under the spacetime translations are denoted P_ν. According to (3.36), they have the form

$$P_\nu = \int_\sigma d\sigma_\mu \; T^{\mu}{}_{\nu}. \tag{3.48}$$

Consider the expression (3.48) in more detail. Let the hypersurface σ in (3.48) be a surface of a constant time t. In this case,

$$P_\nu = \int d^3x T^{0}{}_{\nu} = \int d^3x \Big(\frac{\partial \mathcal{L}}{\partial \partial_0 \phi^i} \partial_\nu \phi^i - \mathcal{L} \delta^{0}{}_{\nu} \Big). \tag{3.49}$$

In particular, the component P_0 has the form

$$P_0 = \int d^3x \Big(\frac{\partial \mathcal{L}}{\partial \dot{\phi}^i} \dot{\phi}^i - \mathcal{L} \Big). \tag{3.50}$$

By analogy with classical mechanics, one defines the momenta $\pi_i = \frac{\partial \mathcal{L}}{\partial \dot{\Phi}^i}$, canonically conjugate to the fields ϕ^i. Then, the component P_0 becomes

$$P_0 = \int d^3x (\pi_i \dot{\phi}^i - \mathcal{L}) = H. \tag{3.51}$$

The expression H is analogous to the classical Hamilton function, or energy. Thus, the component P_0 of the vector P_ν is energy. Then, due to relativistic covariance, P_ν is the energy-momentum vector.

Exercises

3.1. Consider the higher-derivative theory with the Lagrangian depending on higher derivatives of the fields,

$$\mathcal{L} = \mathcal{L}(\phi, \partial_\mu \phi, \partial_{\mu_1} \partial_{\mu_2} \phi, \dots \partial_{\mu_1} \partial_{\mu_2} \dots, \partial_{\mu_n} \phi). \tag{3.52}$$

Formulate the boundary conditions for the variations of the field and its derivatives, which enable one to derive the Lagrange equations from the least action principle. Calculate the variational derivative of the action and obtain the equations of motion.

3.2. Prove, without taking a variational derivative of the actions, that the Lagrangians $\mathcal{L}(\phi, \partial_\alpha \phi)$ and $\mathcal{L}(\phi, \partial_\alpha \phi) + \partial_\mu R^\mu(\phi)$ lead to the same equations of motion.

3.3. Prove that the Lagrangians $\mathcal{L}(\phi, \partial_\mu \phi)$ and $\tilde{\mathcal{L}} = \mathcal{L}(\phi, \pi_\mu) + \varrho^\mu(\pi_\mu - \partial_\mu \phi)$ lead to the same equations of motion. Here $\pi_\mu = \pi_\mu(x)$ and $\varrho^\mu = \varrho^\mu(x)$ are arbitrary vector functions.

3.4. Formulate the conditions under which the equations of motion for a theory with the Lagrangian $\mathcal{L}(\phi^i, \partial_\mu \phi^i)$ have the form

$$A_{ij}^{\mu\nu} \partial_\mu \partial_\nu \phi^j + B_{ij}^\mu \partial_\mu \phi^j + C_i \,=\, 0$$

with constant coefficients, and find the explicit form for the $A_{ij}^{\mu\nu}$, B_{ij}^μ, C_i in this case.

3.5. Let the symmetry transformations be $\delta x^\mu = X^\mu(x, \phi, \partial_\alpha \phi)$. Construct the proof of Noether's theorem in this case.

3.6. Let the condition of invariance (3.15) have the alternative form

$$\int_{\Omega'} d^4x' \, \mathcal{L}\big(\phi'(x'), \partial'_\mu \phi'(x')\big) \,=\, \int_\Omega d^4x \, \Big[\mathcal{L}\big(\phi(x), \partial_\mu \phi(x) + \partial_\mu R^\mu(\phi)\Big],$$

with an arbitrary function $R^\mu(\phi)$. Construct a proof of Noether's theorem in this case. Explore whether the conserved charges depend on $R^\mu(\phi)$.

Comments

Different aspects of Lagrange formalism in field theory are considered in practically all books on relativistic field theory, e.g., in [187], [250], [304], [274], [256], [215], [155], [345], [105] [109], [293], [276], [59].

4
Field models

In this chapter, we consider the constructions of Lagrangians for various field models and discuss the basic properties of these models.

4.1 Basic assumptions about the structure of Lagrangians

Consider an arbitrary theory with a set of fields $\phi^i(x)$ and with the action (3.6). In what follows, we do not need to specify the choice of Ω, and deal with the action

$$S[\phi] \,=\, \int d^4x \, \mathcal{L}\big(\phi(x), \partial_\mu \phi(x)\big). \tag{4.1}$$

The choice of the field model is related by the specification of the set of fields and the Lagrangian $\mathcal{L}$. Usually, it is assumed that fields ϕ^i are the spin tensors, e.g., we will explore scalar, vector and spinor field models, higher-rank tensors, etc. Models with different types of fields in the same Lagrangian are also possible.

As to the choice of Lagrangian, it is assumed that it should be a function of fields and their derivatives, being taken in the same point x^μ (this is called an assumption of locality), that can be always divided into the sum of the two terms

$$\mathcal{L} \,=\, \mathcal{L}_0 + \mathcal{L}_{int}, \tag{4.2}$$

where $\mathcal{L}_0$ is bilinear in fields and their derivatives, while $\mathcal{L}_{int}$ contains powers of the fields and the derivatives higher than the second. The part $\mathcal{L}_0$ is called the free Lagrangian, and $\mathcal{L}_{int}$ is called the interaction Lagrangian.

The equations of motion for the theory with the Lagrangian $\mathcal{L}$ can be written in the form

$$\frac{\partial \mathcal{L}_0}{\partial \phi^i} - \partial_\mu \frac{\partial \mathcal{L}_0}{\partial(\partial_\mu \phi^i)} \,=\, -\Big[\frac{\partial \mathcal{L}_{int}}{\partial \phi^i} - \partial_\mu \frac{\partial \mathcal{L}_{int}}{\partial(\partial_\mu \phi^i)}\Big]. \tag{4.3}$$

Since the Lagrangian $\mathcal{L}_0$ is quadratic in fields and their derivatives, the *l.h.s.* of the equations (4.3) is a linear equation, containing no more than two derivatives of fields. If $\mathcal{L}_{int} = 0$, the corresponding equations of motion will be linear partial differential equations, typically of the order not higher than the second. If $\mathcal{L}_{int} \neq 0$, then the equations of motion will be non-linear. This feature explains the terms "free Lagrangian" and "interacting Lagrangian."

Equations to the free Lagrangian are called free equations of motion, which are linear partial differential equations for spin tensors. Requiring that the fields transform under irreducible spin-tensor representations of the Poincaré group, the corresponding

Introduction to Quantum Field Theory with Applications to Quantum Gravity. Iosif L. Buchbinder and Ilya L. Shapiro, Oxford University Press (2021). © Iosif L. Buchbinder and Ilya Shapiro.
DOI: 10.1093/oso/9780198838319.003.0004

equations of motion must be compatible with the relations (2.87) and (2.91) or with their massless versions, defining the irreducible representations of the Poincaré group in the linear space of fields. Due to the Lorentz covariance, the Lagrangian can be constructed from fields and their derivatives, and other covariant objects of the Lorentz group, such as $\eta_{\mu\nu}$ and spinor quantities, such as $(\sigma^\mu)_{a\dot{a}}$, ε_{ab} and so on.

In principle, the free Lagrangian for any kind of spin-tensor fields can be restored based on the above relations. The main problem in the construction of the Lagrangian $\mathcal{L}$ consists of finding the $\mathcal{L}_{int}$. There is no general prescription for this. The construction of an interacting Lagrangian for a concrete field model is based on the use of additional physical and mathematical assumptions, including the arguments based on the quantum consistency of the theory.

4.2 Scalar field models

Consider the simplest example of a field model, namely, a scalar field.

4.2.1 Real scalar fields

According to (2.87), the real scalar field φ describes the massive or massless irreducible representation of the Poincaré group with spin $s = 0$ under the Klein–Gordon equation

$$(\Box + m^2)\varphi \,=\, 0. \tag{4.4}$$

The free Lagrangian $\mathcal{L}_0$ for the field φ is constructed as follows. Since the equation (4.4) is linear, the corresponding Lagrangian should be quadratic in φ and $\partial_\mu\varphi$. Hence the most general expression for $\mathcal{L}_0$ is

$$\mathcal{L}_0 = \frac{1}{2}c_1\eta^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi + \frac{1}{2}c_2 m^2\varphi^2, \tag{4.5}$$

where c_1, c_2 are some arbitrary numerical coefficients. The term $\varphi\partial_\mu\varphi$ is ruled out by Lorentz covariance. Derive the equations of motion for the Lagrangian (4.5) gives

$$\frac{\partial\mathcal{L}_0}{\partial\varphi} - \partial_\mu\frac{\partial\mathcal{L}_0}{\partial(\partial_\mu\varphi)} \,=\, c_2 m^2\varphi - c_1\Box\varphi \,=\, 0. \tag{4.6}$$

The comparison of this equation with Eq. (4.4) shows that $c_2 = -c_1$, and we obtain the Lagrangian $\mathcal{L}_0$ in the form

$$\mathcal{L}_0 = \frac{c_1}{2}\Big(\eta^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi - m^2\varphi^2\Big). \tag{4.7}$$

To fix the coefficient c_1, one has to derive the energy (3.50) corresponding to the Lagrangian (4.7),

$$E \,=\, P_0 \,=\, \int d^3x\left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}^i}\dot{\varphi}^i - \mathcal{L}\right) \,=\, \frac{c_1}{2}\int d^3x\left(\dot{\varphi}^2 + \partial_j\varphi\partial_j\varphi + m^2\varphi^2\right), \tag{4.8}$$

where $j = 1, 2, 3$. Requiring that the energy is positively defined, we arrive at a positive value of c_1. The absolute value of this constant can be modified by rescaling $\varphi \to k\varphi$