# GAUGE FIELDS

N. P. Konopleva
and
V. N. Popov

# GAUGE FIELDS

N. P. Konopleva
and
V. N. Popov

*Translated from the Second Russian Edition*
*and Edited by*
N. M. Queen
Department of Mathematical Physics
University of Birmingham, England

harwood academic publishers
chur • london • new york

Published under license by:

Harwood Academic Publishers GmbH
Poststrasse 22
CH-7000 Chur, Switzerland

Editorial Office for the United Kingdom:
61 Grays Inn Road
London, WC1X 8TL

Editorial Office for the United States of America:
Post Office Box 786
Cooper Station
New York, NY 10276

**Originally published in Russian in 1980 as КАЛИБРОВОЧНЫЕ ПОЛЯ by Atomizdat**

Library of Congress Cataloging in Publication Data
Konopleva, N.P. (Nelli Pavlovna)
Gauge fields

Translation of Kalibrovochnye polia.
Bibliography: p.
Includes index.
1. Gauge fields (Physics) I. Popov, V. N.
(Viktor Nikolaevich) II. Queen, N.M. III. Title.
QC793.3F5K6613 530.1'43 81-20209
ISBN 3-7186-0045-5 AACR2



Printed in the United States of America

## PREFACE TO THE SECOND RUSSIAN EDITION

Since the appearance of the first edition of this book, which was devoted largely to the construction of the classical and quantum theory of gauge fields, interest in unified theories of the various interactions has grown appreciably. This is due primarily to the solution of two major theoretical problems which stood in the way of the construction of realistic gauge models of elementary particles: 1) renormalizability of gauge theories; 2) the origin of the masses of the vector particles. The mechanism of spontaneous breaking of local gauge symmetry which had already been proposed by Higgs in 1964 not only made it possible to assign mass to the quanta of gauge fields, but also ensured the renormalizability of the resulting theory of massive fields. The latter was demonstrated by 't Hooft in 1971 for the example of the Weinberg—Salam model (1967), which provides a unified description of the weak and electromagnetic interactions. The correctness of the Weinberg—Salam model was confirmed experimentally by the discovery in 1973 of neutral currents, which were predicted by this model in first-order perturbation theory. Subsequently, a large number of unified gauge models of the strong, weak, and electromagnetic interactions required the existence of new quarks possessing a new quantum number ("charm"), as well as new types of elementary particles. In 1974 bound systems of two charmed quarks $c\bar{c}$ — the mysterious $\psi$ particles — were actually observed experimentally. At the present time, these particles, which manifest themselves as extremely narrow long-lived resonances, have already been well studied. There has even appeared a spectroscopy of the family of $\psi$ particles, which are excited states of the $c\bar{c}$ system.

In 1977 an analogous system of two b quarks ($b\bar{b}$) was discovered — the $\Upsilon$ particle. Charmed mesons and baryons in which the new quantum number is not compensated were also discovered. Many unified gauge models of the weak and electromagnetic interactions predicted the existence of heavy leptons, and for a long time this was regarded as an argument against such models. However, in 1976 the heavy $\tau$ lepton with mass ~1.8 GeV was discovered experimentally. Thus, unified gauge models of the interactions are leading to a new physics of elementary particles, which is rich in discoveries. Therefore a solution to the problem of finding a unified description of all forms of interaction (strong, weak, electromagnetic, and gravitational) is not only of mathematical interest, but is becoming practically essential. For the first time since the creation of quantum electrodynamics, unified gauge models of the weak and electromagnetic interactions provide a theory in which calculations can be carried through to completion to arbitrary order of perturbation theory.

Asymptotically free gauge models of the strong interactions are free from ultraviolet divergences and ensure "confinement" for quarks in the infrared region. The next problem is to include quantum gravity in the unified scheme of interactions. An idea which is very promising in this respect is to make use of dual models ("strings") in conjunction with gauge invariance, and possibly also supergravity.

The classical theory of gauge fields is being developed with equal success. The nonlinearity of the classical equations of non-Abelian gauge fields has given birth to a new industry among theoreticians. We have in mind the study of particle-like solutions of these equations (solitons, kinks, monopoles, and vortices). Particle-like solutions possess a new type of charge — topological charge, which one can attempt to associate with the quantum numbers that characterize the elementary particles. Therefore the theory of gauge fields raises the question of the relationship between classical and quantum physics in a new way. Unfortunately, the volume of this book does not enable us to give a sufficiently complete treatment of all the problems. However, we present here the basic mathematical apparatus (with the exception of renormalization theory): the Lagrangian and geometrical formulations of the classical theory of gauge fields, and the quantum theory using the method of functional integration. In addition, we analyze the role of the principles of relativity and symmetry in the construction of a physical theory.

We shall make use of contemporary mathematical methods in the book: the variational formalism and Noether's theorems — in the Lagrangian formulation of field theory invariant with respect to an infinite group (Chapter II); the coordinate-free method of exterior forms on a manifold and the concept of a fiber space — in the analysis of the geometrical picture of interaction (Chapter III); the path-integral method — in the construction of a quantum theory of gauge fields (Chapter IV). In particular, we shall show that the classical theory of a gauge field can be regarded as an aspect of geometry, and in this sense we have a realization of the profound physical and philosophical idea of Einstein that the geometry of space-time does not in itself exist, since it is determined by the interaction of physical bodies. In other words, each form of interaction creates its own geometry.

This book is based on original work of the authors, and it also contains a survey of the most important results on gauge fields by Soviet and foreign authors.

All the chapters of the book are relatively self-contained and may be read independently. The first chapter is introductory in character. To make the exposition of the other chapters more accessible, it introduces, in particular, geometrical and physical terminology in parallel. For an understanding of the remaining chapters, it is desirable to be acquainted with group theory, Riemannian geometry, and field theory at the level of courses given in physics and mathematics departments at Universities. Chapters I—III and the Preface were written by N. P. Konopleva, and Chapter IV by V. N. Popov.

The authors are grateful to Academicians M. A. Markov, L. D. Faddeev, and A. G. Iosif'yan for supporting the second edition of this book and for valuable remarks.

## TRANSLATOR'S PREFACE

This is a translation of the second Russian edition of the book Kalibrovochnye polya, which was published in Moscow in 1980 as an updated version of the first edition of 1972. While no attempt has been made to revise the text of the translation, a number of minor misprints and erroneous references have been corrected, and many references to Russian translations of works published in the West have been replaced by the references to the original sources.

N. M. Queen

CONTENTS

# CHAPTER I. INTERACTION OR GEOMETRY?

## §1. PRINCIPLES OF RELATIVITY, GEOMETRY, AND INTERACTION

Introduction. In the 1960s a peculiar situation arose in elementary-particle theory: on the one hand, there was not a single experimental fact for which a theoretical basis could not be found, and on the other hand, there was no consistent theory which provided a unified description of the entire diversity of properties and species of elementary particles. The gulf between the "internal" symmetries (hypercharge, isospin, etc.) and the "external" (space-time) symmetries of elementary particles was felt particularly acutely. It became increasingly clear that the construction of a unified theory of interactions would require modification of the fundamental principles on which physical theories are based and would lead to the use of new ideas about the structure of space-time and the nature of elementary-particle interactions.

The symmetry properties of elementary particles are usually formulated in terms of the invariants of the symmetry groups* of space-time, which specify the principle of relativity of the theory (for example, Lorentz invariance), and the internal symmetry groups (for example, isospin invariance of the strong interactions, which is a consequence of the fact that the nuclear forces are independent of the electric charges of the particles). Thus, the problem of finding a natural unification of the internal and external symmetries is intimately related to the use of new principles of relativity and symmetry in elementary-particle theory. Such a fundamental principle is the requirement

---

*A *group* is a class of transformations (or operations) over the elements of the given set satisfying the following conditions (axioms): 1) the product of two transformations A and B (two transformations performed in succession) gives some transformation C from the same class, i.e., $A \cdot B = C$; 2) this law of multiplication is associative, i.e., $A \cdot (B \cdot C) = (A \cdot B) \cdot C$; 3) an identity transformation E is defined; 4) each of the transformations A has an inverse transformation $A^{-1}$, i.e., $A \cdot A^{-1} = E$. A group is said to be *finite* if its transformations depend on a finite number of numerical parameters, and *infinite* if the transformations of the group depend on a finite number of functions or on an infinite number of parameters.

of local invariance of the theory, and it is this that is related to the ideas of universal interactions and gauge fields.

The work of Yang, Mills, Utiyama, and Sakurai,[1-3] who first discussed gauge fields, was based on the assertion that all the internal symmetry properties of elementary particles are essentially local in character. It follows from this statement that finite gauge symmetry groups must be replaced by corresponding local groups, the parameters of whose transformations vary from point to point. This makes it possible to endow the theory with a new physical object — a gauge field, the interaction with which ensures invariance of the theory with respect to the local symmetry group. Thus, the principle of local gauge invariance is a deep physical principle, which permits the introduction of an interaction purely axiomatically, its form being determined in accordance with the symmetry properties of the theory. Therefore, the properties of gauge fields can be studied even independently of experiment. The problem of the realization of the theoretical concepts in observable phenomena is in itself quite complicated and is thereby distinguished from the mathematical apparatus of the theory. We note that local invariance was used for the first time as a fundamental physical principle in Einstein's general theory of relativity.[4] This idea was subsequently developed by Weyl, who introduced the electromagnetic field through the requirement of invariance of the theory with respect to local, i.e., point-dependent, expansions of the interval: $ds^{2\prime} = \lambda(x)\, ds^2$.[5] But the principle of local gauge invariance took its final form as a physical principle in the above-mentioned work of Yang, Mills, Utiyama, and Sakurai (see §2).

The gravitational and electromagnetic fields, with which the idea of gauge invariance was first associated, refer to universal interactions. The gravitational field interacts universally with all massive particles, and the electromagnetic field with all charged particles. Local gauge invariance led to the discovery of universal nuclear interactions mediated by unstable vector particles — resonances, which interact identically with all particles that carry isospin. Universality of certain weak interactions was also observed, and in this connection attempts were also made to apply the method of gauge fields to this case.[6] For a number of years these attempts had no success, but the final result exceeded all expectations. After the discovery of the mechanism of spontaneous generation of the masses of vector mesons (the Higgs mechanism[7] (1964)) and the formulation of a renormalization procedure for gauge models with spontaneous symmetry breaking ('t Hooft[8] (1971), Slavnov[9] (1972), and Taylor[10] (1971)), it became possible to construct a unified renormalizable theory of the weak and electromagnetic interactions of elementary particles, the simplest variant of which was the Weinberg—Salam model[11] (1967). This model predicted that neutral currents necessarily exist, and until they were discovered experimentally in 1973 this was regarded as an argument against the theory. Subsequent experiments confirmed more complicated quark gauge models which afford a unified description of the strong, weak, and electromagnetic interactions of hadrons.[12] At the present time, the incorporation of gravity into the general scheme of renormalizable interactions is under consideration.

The basis of the theory of gauge fields comprises symmetry principles and the hypothesis of locality of the fields, which converts global symmetries into local symmetries.

The principle of local gauge invariance reflects a deep relationship between the universality of the various interactions, conservation of the vector currents, and the existence of the interactions themselves. This principle determines the form of all interactions, irrespective of their physical nature, and thereby opens the way to the construction of a unified and consistent theory of the interactions of elementary particles. At the same time, the principle of local gauge invariance, like Einstein's general principle of relativity, gives the theory a form which admits a purely geometrical interpretation. As a result, it becomes possible to develop and generalize Einstein's idea that the geometry of space is not specified ad hoc, but is determined by the interaction of physical bodies.[13] In other words, geometry acquires a dynamical character and effectively reflects the influence on a distinguished test particle (or field) of all the remaining matter in the world.

The geometrization of gauge fields shows that 4-dimensional space-time is merely a particular case of possible dynamical geometries. An arbitrary gauge field corresponds to the geometry of a fiber space obtained from ordinary space-time by replacing its points by "internal" spaces in which the gauge group acts. Thus, the classical theory of gauge fields, like general relativity, becomes a purely geometrical theory. The resulting unified theory of the various interactions (strong, weak, electromagnetic, and gravitational) is also a geometrical theory. Its unity consists in the existence of a *general principle* according to which a geometry corresponding to each of the interactions is constructed.[14] In terms of the geometry of a fiber space, the motion of particles interacting with any gauge field becomes free (forceless). As in general relativity, this eliminates the distinction between inertial (or free) motions and noninertial motions (which take place under the action of external forces). This makes it possible to describe gauge fields by means of simple geometrical concepts (connection coefficients and curvature tensors) and renders geometry experimentally testable. The transition from 4-dimensional space-time to a fiber space implies the recognition of an astonishing possibility: the physical space determined by the interactions may be multidimensional or even infinite-dimensional. From this point of view, however, the description of microprocesses in ordinary space-time terms implies a certain projection of the "true" physical geometry of the interactions into a geometry produced by our macroscopic instruments. Therefore it would be very useful to know what we lose when this projection is made.

Local Symmetries and Geometrization of Interactions. Local Spatial Symmetries and the Gravitational Field. Suppose that we have a square plate of thin glass and a sphere. The flat uniform glass plate will represent flat (Euclidean) space, and the surface of the sphere will represent curved (Riemannian) space. Suppose now that we must "wrap" the glass plate around the sphere.

Let us cut our large glass square into a set of tiny squares and "cover" the sphere with them. This operation is a model of the process of covering a curved surface (or space) by local maps (or coordinate grids). It is easy to see that the whole flat plate can be covered by a single map, while the sphere cannot. It is for this reason that we had to take a set of tiny squares (local maps), in order to fit them as closely as possible to the points of the sphere. Proceeding in this way, we replaced the sphere by a set of small flat surfaces, which are interrelated in a definite manner, for example, rotated with respect to one another by a fixed angle. In other words, we can say that the difference between the set of tiny flat squares assembled into a single flat plate and the same set of squares assembled into a sphere is that the angle of rotation between their planes is zero in the first case but nonzero in the second. Translated into geometrical language, this means that a curved space can be represented as a set of flat spaces "joined" by connection coefficients. The connection coefficients determine the magnitude of the mutual "rotation" or "displacement" of neighboring local flat spaces (Fig. 1). Therefore the connection coefficients are zero when the small squares are stuck together into a

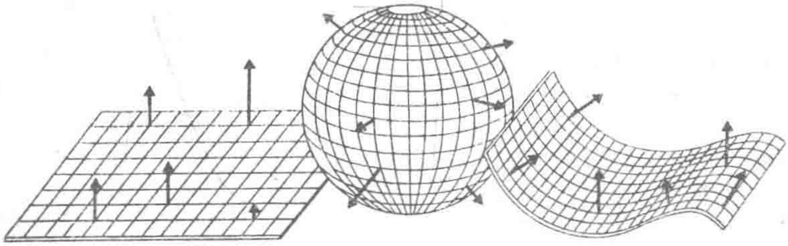

FIGURE 1

plane, but nonzero when they are stuck together into a sphere. Thus, the sphere is equivalent to the set of planes + connection coefficients.

Let us now compare the symmetry groups of the plane and the sphere, or more precisely, the groups of motions for which these objects are, as one says, transformed into themselves.

If the large square about which we spoke at the beginning is rotated through a right angle around an axis passing through its center and perpendicular to its surface, it will occupy the same position as it did before the rotation (see Fig. 1). Since for a person who did not observe the process of rotation itself this state is no different from the original one, we say that after such a rotation the square has been transformed into itself. Note that after this transformation all the points of the plate have moved in one and the same plane and have rotated through one and the same angle, i.e., have undergone one and the same motion. Now if we select some tiny square attached to the sphere and rotate the sphere through a right angle around an axis passing through the center of this square and perpendicular to its surface, the sphere will also coincide with itself. But in this case the only points which undergo the previous motion are those belonging to the chosen tiny square. Points of neighboring squares turned slightly on the sphere with respect to one another undergo rotations in different planes, i.e., different motions. This means that while the flat plate was as a whole symmetric with respect to the considered rotations, on the sphere the previous symmetry has become only local, i.e., it exists for each tiny square individually, but not for all of them collectively. Note that this fact does not rule out the possibility that the sphere as a whole has an intrinsic symmetry different from that of the flat plate.

Thus, locally the sphere possesses the same symmetry as the plane as a whole. Localization of the symmetry reduces to the fact that, while having the same structure (i.e., type of motion) at each point, the transformations must have parameters which vary in going from one point to another. In our example, in going from one tiny square on the sphere to another, we are performing rotations around a new axis, whereas the flat plate was rotated as a whole in one plane around a single axis.

Now let us imagine that both the plane and the sphere are very large and that the observer is very small. Suppose that the observer has the possibility of learning something about the space in which he finds himself, but that all his observations are "tied" to the point at which he is situated and to the instant of time at which he makes the measurements. Obviously, all results of the measurements will reflect only local properties of the space surrounding the observer. Can he establish the nature of the space as a whole? Can he, while situated at a point, distinguish a sphere from a plane? This is precisely the question that was raised for the first time in physics by Einstein.[13] Einstein's answer is contained in his *principle of equivalence*. Usually, this principle is formulated as the principle of (local) equality of the inertial and gravitational masses. But the principle of equivalence can also be given another form, namely, *flat space + a gravitational field is locally equivalent to a curved Riemannian space* (i.e., is indistinguishable from it[15]).

It is easy to see that the principle of equivalence in this form is very similar to the local equivalence of a sphere and a plane established in the example given above. For complete correspondence, it is sufficient to identify the connection coefficients (a geometrical concept) with the gravitational field (a physical concept). We then obtain a geometrical interpretation of gravitation.

What is the geometry of the world about us? In a certain sense, the principle of equivalence implies that there can be no unique answer to this question. We can suppose that space is flat and that all bodies are subject to the influence of a universal field that penetrates all matter, or that there is no field but that space is curved. In this case, the question of the geometry of space as a whole is equivalent to the question of the behavior of physical fields at arbitrarily large distances from the source. The symmetry properties of space become symmetry properties of interactions. The topology of space as a whole is reflected in the properties of the interactions. Thus, geometry and physics are interlocked.

Note that the geometrical interpretation of the gravitational field became possible as a result of localization of the space-time symmetry, i.e., the transition from flat space-time to a Riemannian space which is curved but which locally possesses the same symmetries. Other forms of interaction, namely, those mediated by gauge fields, also admit a purely geometrical interpretation. It is only in this case that the local symmetries are internal symmetries of elementary particles.

Local Internal Symmetries and Gauge Fields. To illustrate clearly what local symmetries are, consider the following example. Suppose that a ping-pong ball is moving along some trajectory and that we do not see whether it is rotating around its center of mass, although we know that the law of conservation of angular momentum is satisfied. How can we describe the positions of points of the surface of the ball at an arbitrary instant of time if the angular velocity of its intrinsic rotation can vary?

As is well known from mechanics, the free flight of a ball is determined only by the motion of its center of mass. The free motion of the center of mass is independent of whether the ball is rotating and whether the speed and direction of the axis of rotation are constant. Rotation around the intrinsic center of mass is an additional (internal) degree of freedom which is present for every body (more precisely, there are three degrees of freedom, since rotation is possible in any plane). If the character of rotation changes, to maintain the law of conservation of angular momentum we must assume that during its flight the ball is acted upon by some force field which twists it or brakes its rotation. This force field is an analog of a gauge field.

Gauge transformations is the name given to those transformations of the functions describing the motion of a particle which are not reflected in the observable characteristics of the motion, i.e., do not alter its physical state. In this sense, rotations of a ball around its center of mass are an analog of gauge transformations of an internal symmetry if we are interested only in the trajectories of the motion of the ball. Localization of this internal symmetry leads to a change of the angular-velocity vector of the intrinsic rotation of the ball. Disappearance of the localization, i.e., the presence of symmetry transformations with constant parameters, corresponds here to the establishment of a constant velocity of rotation along the entire trajectory. It is obvious that the gauge field also vanishes in this case. The constant angular-velocity vector of the ball corresponds in the theory of condensed media to constant properties of a medium throughout its volume, which manifests itself as a constant order parameter (for example, magnetization vector). In the theory of superconductivity, the global symmetry is described by a constant phase of the wave function of an electron.

The intrinsic rotations of a ball are unobservable unless some mark is made on the ball, for example, a stripe is painted, making it possible to observed its rotation. But the intrinsic rotations can be made observable only by breaking the internal symmetry, since the stripe renders different rotations of the ball inequivalent. This simple example illustrates another important fact: whatever symmetry is present, it implies the existence of identical, i.e., indistinguishable, states, whereas observation and measurement presuppose a distinction between the states, i.e., symmetry breaking. This symmetry breaking is always associated with an influence on the system, i.e., with the appearance of some force field.[16] In other words, to make a symmetry observable, it must be broken. An internal microscopic symmetry can become macroscopic and, in principle, observable if in a macroscopically large space-time region the local internal symmetry becomes global (order develops). It then becomes possible to observe macroscopic quantum phenomena. Examples of this kind are provided by the quantization of magnetic flux in superconductors and the appearance of coherent emission (lasers). The classical theory of gauge fields describes microscopically disordered systems and, as a rule, its predictions become experimentally observable on macroscopic scales under special conditions (phase transitions).

Invariance with respect to local gauge transformations means that it is impossible to measure the relative phase of the wave function of a particle at two different world points. This assertion is illustrated by means of the following example involving balls. Suppose that a rotating ball is placed at each point of the Universe. If two such balls are situated at points separated from one another by a spacelike interval, it is impossible to establish their angle of rotation with respect to one another simply because the velocity of light is finite. This is true in any space-time $V_4$.

Each local internal symmetry can be associated with its own gauge field, whose source in the case of invariance with respect to an ordinary (i.e., global) gauge group is a conserved quantity — a vector or tensor current density. In the example involving balls, the source of the gauge field is the intrinsic angular-momentum density of the balls.

**Internal Spaces and a Fiber Space over $V_4$.** Internal symmetries can be understood as symmetries of some internal space whose points correspond to different states of a particle, which are not associated with its position in space. An example of an internal symmetry is isospin invariance, or the fact that the nuclear forces are independent of the charges of the particles. As a result of isospin invariance of the nuclear forces, the proton and neutron are indistinguishable in the absence of an electromagnetic field. Two indistinguishable particles can be regarded as two states of the same particle. We label these states by the values of an internal quantum number — the isospin: $\frac{1}{2}$ (the proton) or $-\frac{1}{2}$ (the neutron). This gives an isotopic doublet. It is also possible to have richer isospin multiplets containing three or more particles. The influence of the electromagnetic field on an isospin multiplet leads to breaking of the isospin symmetry and a decomposition of the multiplet into individual components (particles), which behave differently with respect to the electromagnetic field.

Localization of internal symmetries, like localization of space-time symmetries, makes it necessary to introduce a new physical object — a gauge field. The concept of a gauge field was first introduced by Yang and Mills in connection with an attempt to construct a theory of the strong interactions on the basis of the requirement of invariance with respect to the local group of isospin transformations. In 1954 they proposed a method of introducing a vector field which is responsible for the strong interactions between nucleons and which is related to a conserved isospin current. The idea of the method was as follows.[1] Conservation of isospin is identical to the requirement of invariance of all interactions with respect to rotations of the isospin. This means that the orientation of the isospin has no physical significance when electromagnetic interactions can be neglected. In this case, the distinction between the proton and the neutron becomes completely arbitrary. However, it is usually understood that this arbitrariness is limited by the following condition: as soon as one chooses what to call the proton and what to call the neutron at one point of space-time, the freedom of choice disappears at other space-time points, even at points separated from the first point by a spacelike interval.

This situation is incompatible with the hypotheses of short range and locality of the fields, on which ordinary physical theories are based. In fact, suppose that there is no electromagnetic field and that the proton and neutron