Gauge Approach and Quantization Methods in Gravity Theory

The book gives an overview of the geometrical gauge approach to the gravity theory and the methods of quantization of the gravitational field. Gauge-theoretic formalism (universal principle of the local invariance and the mechanism of spontaneous breaking of the gauge symmetry) forms the basis for the modern understanding of fundamental physical interactions and is successfully confirmed by the experimental discoveries of the gauge bosons and the Higgs particle. The carefully selected material of the book provides a minimal but sufficient mathematical introduction to the methods of the gauge gravitational theory, and gives a concise but exhaustive description of all specific physical consequences. In order to describe the many dramatic changes which took place recently in understanding the concepts, issues and motivations of quantum gravitational physics, a special emphasis is put on a comparison of three different quantization methods of the gravitational field – the covariant approach, the Dirac-Wheeler-DeWitt quantization and the method of Arnowitt-Deser-Misner. The basics of the canonical quantization is explained along with the detailed exposition of the path integral approach, and the introduction to the modern BRST technique is given to demonstrate the consistency of the Arnowitt-Deser-Misner and the Dirac-Wheeler-DeWitt quantization schemes.

The book addresses the physicists who specialize in the gravity theory and the high-energy physics, and it can be recommended to the graduate students and senior undergraduate university students in theoretical physics and mathematics.

ISBN 978-5-02-040047-4

© Ponomarev V. N., Barvinsky A. O., Obukhov Yu. N., 2017
© Nuclear Safety Institute of the Russian Academy of Sciences, 2017
Contents

Preface viiii

1 Lagrangian description of gravity theories of Hilbert-Einstein type 1
1.1. Geometrical spacetime structures ................................................................. 1
1.2. Hilbert variational principle and field equations of the general relativity ................................................................. 8
1.3. First-order formalism and field equations of the Einstein-Cartan theory ................................................................. 11

2 Canonical formalism of gravity theories 17
2.1. Systems with singular Lagrangians ................................................................. 17
2.2. Peculiarities of the canonical formalism in curved spacetime ................................................................. 30
2.3. Geometry of (3 + 1)-decomposition of spacetime ................................................................. 31
2.4. Canonical formalism for the fields on the curved spacetime ................................................................. 34

3 Dynamics of gravity theories of Hilbert-Einstein type 45
3.1. The problem of the “frozen” formalism ................................................................. 45
3.2. Arnowitt-Deser-Misner procedure of selection of physical degrees of freedom ................................................................. 46
3.3. Asymptotically flat and closed worlds ................................................................. 49

4 Torsion effects on the structure and evolution of gravitating systems 57
4.1. Matter fields in the Einstein-Cartan theory ................................................................. 57
4.2. Conformal invariance and spacetime torsion ................................................................. 71
4.3. Pre-Friedman stage of Universe's evolution and spacetime torsion ................................................................. 74
4.4. Production of scalar particles by cosmological torsion field ................................................................. 80
Contents

5 Kinematics of gauge theories of gravity 87
5.1. Special aspects of gauge approach in gravitation ........................................... 87
5.2. Bundle of frames and generalized affine connection .......................................................... 89
5.3. Spontaneous symmetry breaking and non-linear realizations ........................................ 97
5.4. Kinematics of the gauge gravity theory ........................................................................ 100

6 Gauge gravity models with dynamical (Г-S)-interaction 107
6.1. Principles for construction of dynamics .............................................................. 107
6.2. Choice of the action in the gauge gravity theory ............................................... 109
6.3. Physical consequences of the gauge gravity theory ........................................ 114
6.4. Model description of microscopic gravitational interactions .................................... 121

7 Quantization of gravity 129
7.1. Methods of quantum gravity ................................................................................... 129
7.2. Arnowitt-Deser-Misner quantization method ...................................................... 133
7.3. Dirac-Wheeler-DeWitt quantum geometrodynamics ........................................ 138
7.4. Dirac quantization from Batalin-Fradkin-Vilkovisky formalism .............................. 142
7.5. Semiclassical approximation .................................................................................. 152
7.6. Problems and prospects of quantum gravity and cosmology ............................. 160

Appendices 163
A1. Geometry of manifolds ......................................................................................... 163
A2. Spinor analysis on an arbitrary manifold ......................................................... 169
A3. Basic notations .................................................................................................. 173
A4. Comments on the literature and general remarks ............................................ 174

References 179

Subject index 197

Bibliography on gauge gravity theory 201
Vladimir Ponomarev was born in Moscow in 1945. He has studied physics at the Moscow State Lomonosov University and at the Warsaw University. His first scientific adviser was Andrzej Trautman, who supervised his internship at the Institute of Theoretical Physics of the Warsaw University in 1969-72. After defending his PhD thesis (scientific advisors: Andrzej Trautman and Dmitry Ivanenko) in 1974, he worked at the Department of Theoretical Physics at Moscow State Lomonosov University till 1982, when he was elected as the head of the Department of Physics for Natural Sciences of the Moscow State Pedagogical University. In 1985, he was awarded the title of full professor. In 1988, he took part in the organization of the Nuclear Safety Institute, where he became the deputy director. He published about 150 scientific papers on the classical and quantum theory of gravity, as well as coauthored the monograph “Geometrodynamical methods and the gauge approach in the theory of gravitational interactions” (1985, Moscow) together with A.O. Barvinsky and Yu.N. Obukhov. During the period of 1991-2009, he was involved in the work on social-economic transformations within the former USSR, including the service in the Government of the Russian Federation as a Deputy Minister. In 2009, he resumed the work in the Nuclear Safety Institute as the Deputy Director for Strategic Development and Innovations, along with continuing the activities in expert groups under the Government and Parliament of the Russian Federation. Throughout his 50-year professional career, he has been active in teaching, also in the recent times as an invited professor at several universities and academic institutions, including the Moscow Physical-Technical University.
ANDREI O. BARVINSKY

Andrei Barvinsky was born in 1955. After graduating from the Physics Department of Moscow State University he has got his PhD in theoretical physics in the field of quantum gravity theory and held a postdoctoral fellowship at the University of Alberta, Canada. Since then he is a leading scientific researcher at the Theory Department of the Lebedev Physics Institute of the Russian Academy of Sciences. The areas of his scientific interests are quantum field theory, quantum gravity and cosmology. In these areas he has more than one hundred publications including the pioneering contributions to the quantum theory of gauge constrained systems, background formalism and the effective action method in quantum field theory and Higgs inflation theory in cosmology of the very early quantum Universe. His teaching activity included supervision of many PhD students and professorship at the Ludwig-Maximillan University in Munich, Germany. He has an ongoing wide scientific collaboration with the physicists of the Landau Institute for Theoretical Physics, the Physics Department of the Ludwig-Maximillan University in Munich, University of Cologne, Department of Physics and Astronomy of the University of British Columbia, the University of Bologna and the Theory Division of CERN. He is also one of the organizers of the long series of the Sakharov and Ginzburg conferences on physics held for more than twenty years at the Theory Department of the Lebedev Physics Institute in Moscow.
YURI N. OBUKHOV

Yuri Obukhov was born in Berlin in 1956 and graduated from the Department of Physics of the Moscow State Lomonosov University. He received the PhD doctoral degree in 1983 from the Moscow State University, where he later worked for many years. He spent the first post-doctoral term in 1986-87 in the relativity group of A. Trautman at the Institute of Theoretical Physics of the Warsaw University. After he was awarded a Humboldt Fellowship and visited the University of Cologne in 1992-94, a long and fruitful collaboration with Friedrich Hehl has started, and subsequently Yuri Obukhov has held several times research positions at the Institute for Theoretical Physics of the University of Cologne. He was also twice (2002 and 2006) a visiting professor at the Institute for Theoretical Physics of the University of Sao Paulo (UNESP), and a research fellow at the Mathematics Department of the University College London (2009-2011). Since 2013 he is a senior scientist at the Nuclear Safety Institute of the Russian Academy of Sciences. Along with research, Yuri Obukhov took part in educational activities, teaching at different universities the lecture courses on classical and quantum field theory, classical gravity theory, relativistic cosmology, and supervising numerous master and PhD students. The scientific interests of Yuri Obukhov encompass the relativistic quantum theory, electrodynamics, classical gravity theory and cosmology. His original results obtained in these areas were published in more than 150 papers. In 2003, together with Friedrich Hehl he published a book “Foundations of Classical Electrodynamics: Charge, Flux, and Metric” (Birkhauser, Boston).
Modern understanding of gravitational phenomena is based on the concept of the spacetime geometry. This is drastically different from Newton’s theory in which the Euclidean space and an absolute time represented a fixed arena for independent physical processes of different kind: mechanical, electric, magnetic, gravitational. Eventually, the development of ideas of Minkowski and Einstein culminated in a unification of space and time into a four-dimensional spacetime manifold. Its geometry became dynamical: it is no longer assumed to be fixed but depends on the motion of matter. In a broad sense, gravitation can be viewed as a geometrodynamics, i.e., as a theory of a dynamical geometry of spacetime. The beautiful and powerful methods of differential geometry are used in Einstein’s general relativity (GR) to describe the gravitational phenomena in terms of the properties of the four-dimensional spacetime manifold. Later, it was recognized that the three other physical interactions (electromagnetic, weak and strong) can also be geometrized by making use of the Yang-Mills gauge-theoretic approach, and a natural question arose whether the gravity theory can be consistently formulated in the gauge framework. The answer is not that simple as it might appear at the first sight. The subtle point is that the Standard Model is based on the fundamental symmetry groups acting in the internal spaces, whereas the gravity is obviously related with the symmetry of the spacetime itself.

In the recent decades, a growing interest of researchers has been attracted to the problem of construction of a unified theory of fundamental interactions. Such a theory, capable to explain the hierarchical structure of physical forces in nature, most probably could be based on the universal principle of the local invariance and on the resulting formalism of the gauge fields. A significant progress in this area is manifest in the development of the unified models of weak and electromagnetic interactions so successfully confirmed by the experimental discoveries of the gauge bosons and the Higgs particle. However, these models do not take the gravity into account, and therefore they should be considered only as approximations to the unification of all interactions which could take place on extremely small scales, or at the very high energies and high temperatures of the early Universe, when the gravitational effects are expected to be significant. Thus, the next step in the construction of a genuine unified theory is a consistent
inclusion of gravity in the general scheme of fundamental physical interactions, reconciling the principles of general relativity theory with those of the quantum theory and of gauge field theory.

This book gives an overview of the geometrical gauge approach to the gravity theory (Chapters 1, 4-6) and methods of the gravitational field quantization (Chapters 2, 3 and 7). Taking into account the long history of the subject and the volume of the relevant literature, the material for the book was carefully selected to provide a minimal but sufficient mathematical introduction to the methods of the gauge gravitational theory (Chapter 4), and to give a concise but exhaustive description of all specific physical consequences (Chapters 3 and 5). The special feature of our book is the use of the invariant language to describe the \((3+1)\)-decomposition of the spacetime manifold (Chapter 2). A particular emphasis is put on a comparison of three different quantization methods of the gravitational field – the covariant approach, the Dirac-Wheeler-DeWitt quantization and the method of Arnowitt-Deser-Misner (Chapter 3).

In this regard, it is worth emphasizing that many dramatic changes took place recently in understanding the concepts, issues and motivations of quantum gravitational physics. Precision cosmology in the form of the cosmic microwave background (CMB) observations actually rendered quantum gravity the status of an experimental subject, inflation theory acquired a dominant role in the picture of the early Universe and we have been put face to face with the mysteries of the dark matter and dark energy phenomena. On top of this progress quantum geometrodynamics method and quantum cosmology have lost a status of acute direction in theoretical physics and, moreover, a widely accepted viewpoint prevailed that this approach is essentially discredited due to numerous but rather unproductive applications of the Wheeler-DeWitt equation. This opinion is, however, very erroneous, in particular, because the analogous situation in quantum field theory is treated completely differently – nobody calls in question the status of the Schrödinger equation, even though almost all field-theoretical results were attained by the relativistic invariant \(S\)-matrix method. Similarly with the Wheeler-DeWitt equation – it lies at the foundation of quantum gravity theory of the Universe as a whole, even though its direct application to concrete problems might be ineffective and should be replaced by such advanced, though technical, tools as path integration. This is the main objective of Chapter 7.

The idea to write this book was born after the reprint volume with commentaries by Friedrich Hehl and Milutin Blagojević “Gauge Theory of Gravitation” was published in 2013, referring to our monograph “Geometrodynamical methods and gauge approach in the theory of gravitational interactions” which appeared in Russian in 1985. An original plan was to translate that old book into English. However, soon it became clear to the authors that after the 30 years some of the material became outdated and should be removed, whereas important new results which appeared in the meantime (also due to the authors) should be included. A need of the serious changes became obvious, and a new deeply revised book was written, now in English. In particular, the discussion of
the classical gauge models and the quantum gravity methods has undergone an essential revision. At the same time, we decided to preserve the general structure of the monograph and to keep the overall presentation style which was, in our opinion, quite successfully chosen in the old book.

We did not intend to describe in detail the modern mathematical methods of the classical gravity theory, which can be found in [2, 3, 4]. The summary of the relevant notions and definitions is given in Appendix A3. At the same time, all the derivations and constructions necessary for understanding of the book are given in full. We also did not try to provide a comprehensive review of the literature on the subject. This is to some extent compensated by the comments to the literature presented at the end of the book in Appendix A4. Moreover, we compiled an exhaustive bibliography of more than 3000 publications on the gauge gravity theory, its physical aspects and related mathematical results. Since no special preselection was made, the readers should clearly understand that some items in this bibliography may be erroneous or misleading, however, we do hope that none of the important publications on the gauge gravity was overlooked, thus making this bibliography a useful resource on the subject.

The book addresses the physicists who specialize in the gravity theory and the high-energy physics, and it can be recommended to the graduate students and senior undergraduate university students in physics and mathematics, who are familiar with the classical and quantum field theory at the level of the textbooks “Classical field theory” of L.D. Landau and E.M. Lifshitz and “Introduction to the theory of quantized fields” of N.N. Bogoliubov and D.V. Shirkov.

The book is written jointly by the authors. Chapters 1, 4-6 are based on original works of Yu.N. Obukhov and V.N. Ponomarev. The material presented in Chapters 2, 3, and 7 is based on the results obtained by A.O. Barvinsky and V.N. Ponomarev.

To help the reader, we give an approximate scheme of dependence of the chapters:

1 2 3 7

4 5 6

The authors express their profound gratitude to all colleagues who took part in the discussion of results presented in the book. We are sincerely grateful to the scientific and administrative staff of the Nuclear Safety Institute of the Russian Academy of Sciences (IBRAE RAN) for the constant and generous support and for the assistance in preparing the manuscript for publication. One of the authors (A.O.B.) is grateful for hospitality to the Theory Division of CERN where his work on the book was accomplished. Y.N.O. acknowledges a partial support from the Russian Foundation for Basic Research (Grant No. 16-02-00844-A).

Authors
1
Lagrangian description of gravity theories of Hilbert-Einstein type

1.1. Geometrical spacetime structures

In the framework of the geometrodynamical approach, we model the spacetime as a smooth four-dimensional manifold $M_4$. On the latter, one can introduce various geometrical structures to describe the gravitational field. In this section, we present a brief overview of the most important ones: the metric and the affine connection structure. The summary of the main notions, definitions and methods of the modern differential geometry can be found in Appendix A1, for a more detailed exposition see [2]-[4]. Our basic notations are given on page 173.

Metric structure ($g$-structure)

A differentiable manifold $M_4$ is called a Riemannian space $V_4$, when the metric is defined on it, that is when a smooth tensor field $g_{\mu \nu}(x^\lambda)$, such that $g_{\mu \nu}(x^\lambda) = g_{\nu \mu}(x^\lambda)$, introduces the metricity relations on the manifold: for any two infinitely close points $x$ and $x + dx$, the interval ("distance") between them is defined by

$$ds^2 = g_{\mu \nu} dx^\mu dx^\nu.$$  

The metric defines a scalar product on the manifold by assigning to any two vectors $A$ and $B$ a real number $AB = g_{\mu \nu} A^\mu B^\nu$, and thereby introduces for each vector $V$ its length $l^2(V) = g_{\mu \nu} V^\mu V^\nu$. By definition, the metric is assumed to be non-degenerate in the sense that the determinant $g = \det g_{\mu \nu} \neq 0$. This allows to determine the matrix $g^{\mu \nu}(x^\lambda)$ inverse to $g_{\mu \nu}(x^\lambda)$ such that $g^{\mu \nu} g_{\nu \lambda} = \delta_\lambda^\mu$, thereby establishing an isomorphism between the covariant and contravariant tensor...
spaces. In local coordinates, this isomorphism is described by the operation of raising and lowering of indices. The metric is called Lorentzian if its signature is $\pm 2$. The Lorentzian metric structure forms the basis of the general relativity theory (GR).

**Connection structure ($\Gamma$-structure)**

Independently of the metric structure, one can define the structure of an affine connection on $M_4$ manifold. The linear (affine) connection $\Gamma$ introduces the local isomorphism of tangent spaces at different points $x \in M_4$ of the manifold, by specifying a rule that maps a vector $V$ (an element of the tangent space) at a point $x$ into a vector $V'$ at an infinitely near point $x + dx$. In the local coordinates, this law is written as

$$\delta V^\mu = V'^\mu - V^\mu = -\Gamma^\mu_{\nu\alpha} V'^\nu dx^\alpha.$$ 

Thus, equivalently, we say that the affine connection $\Gamma$ is introduced on $M_4$ manifold, when the global (i.e., given in each chart of the atlas of the manifold) field $\Gamma^\mu_{\nu\alpha}(x)$ is defined on it. When changing one coordinate chart $\{x\}$ to another one $\{x'\}$, the quantities $\Gamma^\mu_{\nu\alpha}(x)$, which are called coefficients of the linear (affine) connection, are transformed as

$$\Gamma'_{\alpha\beta\mu}(x') = \left[ \frac{\partial x^\rho}{\partial x'^\alpha} \frac{\partial x'^\sigma}{\partial x^\rho} \Gamma^\rho_{\sigma\nu}(x) + \frac{\partial x^\rho}{\partial x'^\nu} \left( \frac{\partial x'^\rho}{\partial x^\sigma} \right) \frac{\partial x'^\nu}{\partial x^\mu} \right] \frac{\partial x'^\mu}{\partial x^\mu}.$$

By defining of the connection structure, one can introduce the notion of a covariant derivative of a tensor. The covariant derivative with respect to an arbitrary connection $\Gamma$ is denoted by $\nabla$ and is defined as follows. Suppose that in the local coordinates, the tensor field has components $T_{\alpha_1...\alpha_p}^{\beta_1...\beta_q}$. Then

$$\nabla_\mu T_{\alpha_1...\alpha_p}^{\beta_1...\beta_q} = \partial_\mu T_{\alpha_1...\alpha_p}^{\beta_1...\beta_q} + \Gamma_{\mu\rho}^{\alpha_1} T_{\beta_1...\beta_q}^{\alpha_2...\alpha_p} + \Gamma_{\mu\rho}^{\alpha_2} T_{\beta_1...\beta_q}^{\alpha_1...\alpha_p} + \Gamma_{\mu\rho}^{\alpha_3} T_{\beta_1...\beta_q}^{\alpha_2...\alpha_p} + \Gamma_{\mu\rho}^{\alpha_4} T_{\beta_1...\beta_q}^{\alpha_3...\alpha_p} + \Gamma_{\mu\rho}^{\alpha_5} T_{\beta_1...\beta_q}^{\alpha_4...\alpha_p} + \cdots - \Gamma_{\beta_1\mu}^{\alpha_1...\alpha_p} T_{\sigma_1...\beta_q}^{\beta_1...\beta_p} - \cdots - \Gamma_{\beta_1\mu}^{\alpha_1...\alpha_p} T_{\beta_1...\sigma_q}^{\alpha_1...\alpha_p}.$$

The connection also defines the notion of a parallel transport of tensors. Let $\gamma(t)$ be a smooth curve in $M_4$, specified by the parametric equations $\gamma = \{x^\mu(t)\}$. Then we define the covariant derivative of a tensor field $T$ along this curve, $DT/dt$. In the local coordinates, it is equal to $DT/dt = \frac{dx^\mu}{dt} \nabla_\mu T$.

The manifold $M_4$ with the connection structure defined on it is called an affinely connected space (denoted by $L_4$), and it is characterized by a number of geometrical objects which we briefly consider below.

In general, we assume that both structures—the metric and the connection—are introduced on the manifold. By definition, they are independent. We call such a space a generalized metric-affine spacetime $G_4$.

Let us consider the parallel transport of a vector $V$ from a point $x \in G_4$ along some closed curve $\gamma \subset G_4$. When returning back to $x$, one finds that in general case the final (parallelly transported) vector $V'$ does not coincide with the original one $V$. The difference is manifest in the three effects: the vector $V'$
1.1. Geometrical spacetime structures

is rotated with respect to \( V \), its length is changed, and the image of the contour \( \gamma \) in the tangent space turns out to be broken by a non-zero vector.

These effects are determined by the following basic geometrical objects of \( G_4 \):

The curvature tensor of the connection

\[
R^\alpha_{\beta\mu\nu} = \partial_\mu \Gamma^\alpha_{\beta\nu} - \partial_\nu \Gamma^\alpha_{\beta\mu} + \Gamma^\alpha_{\rho\mu} \Gamma^\rho_{\beta\nu} - \Gamma^\alpha_{\rho\nu} \Gamma^\rho_{\beta\mu};
\]  

(1.1)

the homothetic curvature tensor

\[
\Omega_{\mu\nu} = \partial_\mu \Gamma_{\nu} - \partial_\nu \Gamma_{\mu},
\]  

(1.2)

where \( \Gamma_{\mu} := \Gamma^\lambda_{\lambda\mu} \);

and the torsion tensor[20]

\[
Q^\alpha_{\mu\nu} = \Gamma^\alpha_{[\mu\nu]} = \frac{1}{2}(\Gamma^\alpha_{\mu\nu} - \Gamma^\alpha_{\nu\mu}).
\]  

(1.3)

In the local coordinates, the aforementioned effects of the parallel transport along a closed loop read as follows:

- the rotation of a vector:

\[
\delta V^\alpha \approx R^\alpha_{\beta\mu\nu} V^\beta \, ds^{\mu\nu},
\]  

(1.4)

- the change of the length:

\[
\delta l \approx l(V) \Omega_{\mu\nu} \, ds^{\mu\nu},
\]  

(1.5)

- the contour image in the tangent space is broken by the vector:

\[
\xi^\alpha \approx 2Q^\alpha_{\mu\nu} \, ds^{\mu\nu}.
\]  

(1.6)

Here \( ds^{\mu\nu} \) is the surface element spanned by the closed contour \( \gamma \).

We introduce now another fundamental tensor which is an important characteristic of \( G_4 \) manifold, even being not entirely independent of the tensors introduced above. This is the nonmetricity tensor that measures the (in)compatibility of the independent metric and connection. It is equal to the covariant derivative of the metric \( g_{\alpha\beta} \):

\[
K_{\mu\alpha\beta} = \nabla_\mu g_{\alpha\beta}.
\]  

(1.7)

We say that the metric and connection are compatible when \( K_{\mu\alpha\beta} = 0 \) vanishes on \( M_4 \), i.e., when the metricity condition is fulfilled

\[
\nabla_\mu g_{\alpha\beta} = 0.
\]  

(1.8)

Let us consider some properties of the fundamental tensors. A closer inspection of the aforementioned non-integrable effects (1.4)-(1.6) shows that they are due to the fact that covariant derivatives \( \nabla \) do not commute with each other. For any vector \( V \) we find

\[
(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu) V^\alpha = R^\alpha_{\beta\mu\nu} V^\beta + 2Q^\beta_{\mu\nu} \nabla_\beta V^\alpha.
\]
This result can be generalized to arbitrary tensors. In particular, by considering the commutator of covariant derivatives of the metric, we obtain the useful identities relating the curvature tensors (1.1) and the nonmetricity (1.7):

\[ 2\nabla_{[\mu}K_{\nu]\alpha\beta} = 2\nabla_{[\mu}\nabla_{\nu]}g_{\alpha\beta} = -2R_{(\alpha\beta)\mu\nu} + 2Q^\sigma_{\mu\nu}K_{\rho\alpha\beta}. \] (1.9)

As a result, we find that the symmetric part of the curvature tensor in the first pair of indices \( R_{(\alpha\beta)\mu\nu} \) is nontrivial when \( K_{\alpha\mu\nu} \neq 0 \). Contracting eq. (1.9) with \( g^{\alpha\beta} \), we recover the homothetic curvature

\[ g^{\alpha\beta}R_{\alpha\beta\mu\nu} = R_{\alpha\mu\nu} = \Omega_{\mu\nu} = -\frac{1}{2}(\partial_\mu K_{\nu} - \partial_\nu K_{\mu}), \] (1.10)

where \( K_{\mu} = K_{\mu\alpha\beta}g^{\alpha\beta} \) is the vector of nonmetricity (or otherwise, the Weyl vector field [5]).

The equation (1.7) can be solved with respect to the connection \( \Gamma^\lambda_{\mu\nu} \) so that the latter is expressed in terms of the metric, torsion and nonmetricity:

\[ \Gamma^\lambda_{\mu\nu} = \{^\lambda_{\mu\nu}\} + D^\lambda_{\mu\nu}. \] (1.11)

Here

\[ \{^\lambda_{\mu\nu}\} = \frac{1}{2}g^{\lambda\sigma}(\partial_\nu g_{\mu\sigma} + \partial_\mu g_{\nu\sigma} - \partial_\sigma g_{\mu\nu}) \]

is the Christoffel symbol, and

\[ D^\lambda_{\mu\nu} = Q^\lambda_{\mu\nu} + Q_{\mu\nu}^\lambda + Q_{\nu\mu}^\lambda - \frac{1}{2}g^{\rho\lambda}(K_{\rho\mu\nu} + K_{\mu\nu\rho} - K_{\rho\mu\nu}) \]

is the connection defect tensor (also called a distortion tensor). The connection \( \{^\lambda_{\mu\nu}\} \) is the only torsion-less connection which is compatible with the metric structure. It is called a Riemannian or a metric connection.

An arbitrary connection with a non-zero torsion, but compatible with metric, is called the Riemann-Cartan connection. Its explicit structure is obtained from (1.11) if we put \( K_{\alpha\mu\nu} = 0 \):

\[ \Gamma^\lambda_{\mu\nu} = \{^\lambda_{\mu\nu}\} + Q^\lambda_{\mu\nu} + Q_{\mu\nu}^\lambda + Q_{\nu\mu}^\lambda = \{^\lambda_{\mu\nu}\} + T^\lambda_{\mu\nu}. \] (1.12)

The combination

\[ T^\lambda_{\mu\nu} = Q^\lambda_{\mu\nu} + Q_{\mu\nu}^\lambda + Q_{\nu\mu}^\lambda \]

is usually called a contortion tensor. From now on we will confine ourselves to the Riemann-Cartan spacetime \( U_4 \) with the metric and the Riemann-Cartan connection (1.12).

The fundamental tensors in \( U_4 \) have a large number of independent components (24 for the torsion and 36 for the curvature), and it is convenient to decompose these geometrical objects into the \( L_6 \)-irreducible pieces (where \( L_6 = SO(3,1) \) is the Lorentz group). The decomposition of the torsion tensor into the three \( L_6 \)-irreducible parts reads as follows:

\[ Q^\lambda_{\mu\nu} = \bar{Q}^\lambda_{\mu\nu} + \frac{2}{3}\delta^\lambda_\mu Q_\nu + g^{\lambda\sigma}\varepsilon_{\mu\nu\sigma\rho} \dot{Q}^\rho. \] (1.13)
Here \( Q^\lambda_{\mu\nu} \) is the traceless part of the torsion tensor, with the properties: \( \overline{Q}^\lambda_{\mu\lambda} = 0 \) and \( \varepsilon^{\lambda\mu\rho\sigma} Q_{\rho\sigma} = 0 \); the torsion trace vector is \( Q_\mu := Q^\lambda_{\mu\lambda} \); and \( Q_\mu = \frac{1}{6} \varepsilon_{\mu\lambda\sigma} Q^{\nu\lambda\sigma} \) is the torsion pseudotrace vector (\( \varepsilon_{\lambda\mu\nu\chi} \) is the totally antisymmetric Levi-Civita tensor: \( \varepsilon_{\lambda\mu\nu\chi} = \varepsilon_{[\lambda\mu\nu\chi]} \)).

In a similar way, one can decompose the curvature tensor in \( U_4 \) spacetime into the 6 irreducible pieces \[1.26\]:

\[
R_{\alpha\beta\mu\nu}(\Gamma) = C_{\alpha\beta\mu\nu}(\Gamma) + \frac{1}{2} \{g_{\alpha\mu} R_{\beta\nu}(\Gamma) - g_{\alpha\nu} R_{\beta\mu}(\Gamma) - g_{\beta\mu} R_{\alpha\nu}(\Gamma)
+ g_{\beta\nu} R_{\alpha\mu}(\Gamma)\} - \frac{1}{6} \{g_{\alpha\mu} g_{\beta\nu} - g_{\alpha\nu} g_{\beta\mu}\} R(\Gamma)
+ \frac{1}{12} \varepsilon_{\alpha\beta\mu\nu} D(\Gamma)
+ \frac{1}{4} \{\varepsilon_{\alpha\beta\mu} D(\lambda\nu)(\Gamma) - \varepsilon_{\alpha\beta\nu} D(\lambda\mu)(\Gamma)
+ \varepsilon_{\mu\nu\alpha} D(\lambda\beta)(\Gamma)\}.
\]

Here \( C_{\mu\nu\chi}(\Gamma) \) is the non-Riemannian analog of the Weyl tensor with the same algebraic properties 
\[
C_{\lambda\mu\nu\chi}(\Gamma) = C_{\nu\chi\lambda\mu}(\Gamma), \quad C^{\lambda}_{\mu\lambda\nu}(\Gamma) = 0, \quad C^{\lambda}_{\chi\nu\mu}(\Gamma) = 0,
C_{\lambda\mu\nu\chi}(\Gamma) = C^{\lambda}_{\mu\nu\chi}(\Gamma) = C^{\lambda}_{\nu\mu\chi}(\Gamma),
\]
\( R_{\mu\nu}(\Gamma) = R^{\lambda}_{\mu\lambda\nu}(\Gamma) \) is the generalized Ricci tensor, \( R(\Gamma) = g^{\mu\nu} R_{\mu\nu}(\Gamma) \) is the curvature scalar, and \( D^{\mu\nu}(\Gamma) = \frac{1}{2} \varepsilon^{\mu\alpha\beta\chi} R_{\alpha\beta\lambda\nu}(\Gamma) \) so that \( D(\Gamma) = D^{\mu\mu}(\Gamma) \).

The list of the 6 irreducible curvature part includes the generalized Weyl tensor, the curvature scalar, the curvature pseudoscalar, the skew-symmetric part of the Ricci tensor, and the two traceless symmetric tensors:

\[
C_{\alpha\beta\mu\nu}(\Gamma), \quad R(\Gamma), \quad R_{(\mu\nu)}(\Gamma) - \frac{1}{4} R(\Gamma) g_{\mu\nu},
R_{[\mu\nu]}(\Gamma), \quad D(\Gamma), \quad D_{(\mu\nu)}(\Gamma) - \frac{1}{4} D(\Gamma) g_{\mu\nu}.
\]

The total number of components is, as expected: \( 36 = 20 + 1 + 9 + 6 + 1 + 9 \). The second line represents the essentially non-Riemannian irreducible parts. When the torsion \( (1.3) \) vanishes, eq. \((1.14)\) reduces to the standard decomposition for Riemann-Christoffel curvature tensor \[27\] into the three pieces shown in the first line.

It is worthwhile to note that the skew-symmetric tensor \( D_{[\alpha\beta]}(\Gamma) \) does not represent an independent part. It is expressed in terms of the antisymmetric part of the Ricci tensor: \( D_{[\mu\nu]}(\Gamma) = -\frac{1}{2} \varepsilon_{\mu\nu\alpha\beta} R^{[\alpha\beta]}(\Gamma) \).

The Riemann-Cartan curvature \((1.12)\) is antisymmetric in the first pair of indices, as it is obvious from \((1.9)\), and it satisfies the following identities:

\[
R^{\alpha}_{\mu\nu\lambda}(\Gamma) = -2 \nabla_{[\mu} Q^\alpha_{\nu\lambda]} + 4 Q^\alpha_{\beta[\mu} Q^\beta_{\nu\lambda]} \] \quad (1.15)

– the generalized Ricci identity;

\[
R_{\alpha\beta\mu\nu}(\Gamma) - R_{\mu\nu\alpha\beta}(\Gamma) = \frac{3}{2} \{R_{[\alpha\beta\mu[\nu}(\Gamma) - R_{[\alpha\beta\nu]\mu]}(\Gamma)
- R_{[\mu\nu\alpha]\beta}(\Gamma) + R_{[\mu\nu\beta]\alpha}(\Gamma)\}, \quad (1.16)
\]
– the “commutator of pairs of indices” identity;
\[ \nabla_{[\mu} R^{\alpha\beta}_{\nu\lambda]}(\Gamma) = 2 R^{\alpha\beta}_{\sigma[\mu}(\Gamma) Q^{\sigma}_{\nu\lambda]} \] (1.17)

– the generalized Bianchi identity.

Contracting a pair of indices (upper and lower) in (1.15), we derive a so-called contracted Ricci identity
\[ R_{[\mu\nu]} = (\nabla_\lambda - 2Q_\lambda) \left( Q^{\lambda}_{\mu\nu} + 2\delta^{\lambda}_{[\mu} Q_{\nu]} \right). \]

Similarly, contracting the two pairs of indices in (1.17), we find the contracted Bianchi identities
\[ (\nabla_\sigma - 2Q_\sigma) G^\sigma_{\mu}(\Gamma) + 2Q^\rho_{\mu\sigma} G^\sigma_{\rho}(\Gamma) + (Q^\sigma_{\alpha\beta} + 2\delta^\sigma_{[\alpha} Q_{\beta]}) R^{\alpha\beta}_{\sigma\mu}(\Gamma) = 0, \] (1.18)

where
\[ G_{\lambda\nu}(\Gamma) = R_{\lambda\nu}(\Gamma) - \frac{1}{2} g_{\lambda\nu} R(\Gamma) \]
is the generalized Einstein tensor.

With the help of the geometrical objects introduced above, one can establish a natural classification of the spacetime theories in accordance with their underlying metric-affine structures, see Table 1.1.

**Tetrad structure**

Along with the so-called world geometrical structures which we described above, in the gravity theory one effectively uses their local Lorentz analogues. In the latter approach, the main object is tangent bundle \( T(M) \), and the related bundle of orthonormal frames \( AO(M) \) which is the principal fibre bundle with the Lorentz \( \mathbb{L}_6 \) structure group.

All the geometrical objects of the spacetime theory had previously been defined in the local coordinates on the manifold with respect to the world or coordinate frame \( \partial_\mu \equiv \partial/\partial x^\mu \) in the tangent space \( T_x(M) \). However, it is often convenient (and when considering spinors in \( M_4 \), it is necessary) to define all the objects with respect to arbitrary orthonormal basis in \( T(M) \), \( \{e_a\}, a = 0, 1, 2, 3 \). Such a frame is orthonormal in the sense that the scalar products of its legs are equal \( (e_a \cdot e_b) = \eta_{ab} = \text{diag}(-1,+1,+1,+1) \) – the Minkowski metric. We call such a basis a local Lorentz frame, and accordingly all the geometrical objects are also called the local Lorentz objects when they are considered with respect to this frame. In order to distinguish the components with respect to \( \{e_a\} \) from the components with respect to \( \{\partial_\mu\} \), we will use the Latin indices \( (a,b,c\ldots) \) for the local Lorentz frames, while the Greek indices \( (\mu,\nu,\lambda\ldots) \) will label the world quantities with respect to \( \{\partial_\mu\} \).

The transition from a coordinate frame to an arbitrary orthonormal frame is described by the tetrad coefficients \( h^a_\mu \). By definition
\[ \partial_\mu = h^a_\mu e_a. \]
Table 1.1: Classification of metric-affine theories

<table>
<thead>
<tr>
<th>Geometrical objects</th>
<th>Type of space</th>
<th>Example of theories</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^a = 0$ $Q^a = 0$ $\Omega^a = 0$</td>
<td>Minkowski space $M_4$</td>
<td>Special relativity theory</td>
</tr>
<tr>
<td>$R^a = 0$ $Q^a \neq 0$ $\Omega^a = 0$</td>
<td>Weitzenböck (or absolute parallelism) space $P_4$</td>
<td>Translational gauge gravity theory [34]</td>
</tr>
<tr>
<td>$R^a \neq 0$ $Q^a = 0$ $\Omega^a = 0$</td>
<td>Riemann space $V_4$</td>
<td>General relativity theory (GR) [27]</td>
</tr>
<tr>
<td>$R^a \neq 0$ $Q^a = 0$ $\Omega^a \neq 0$</td>
<td>Weyl space $W_4$</td>
<td>Weyl’s gravity theory [5]</td>
</tr>
<tr>
<td>$R^a \neq 0$ $Q^a \neq 0$ $\Omega^a = 0$</td>
<td>Riemann-Cartan space $U_4$</td>
<td>Einstein-Cartan gravity theory (ECT) [7, 6]</td>
</tr>
<tr>
<td>$R^a \neq 0$ $Q^a \neq 0$ $\Omega^a \neq 0$</td>
<td>Generalized metric-affine space $G_4$</td>
<td>Various asymmetric field theories, in particular Einstein-Schrödinger [9]</td>
</tr>
</tbody>
</table>

These quantities are closely related to the spacetime metric, namely

$$g_{\mu\nu} = h^{a}_{\mu} h^{b}_{\nu} \eta_{ab}.$$ (1.19)

In the bundle of the Lorentz frames, one can define a connection $\Gamma^{a}_{b\mu}$ that introduces their parallel transport. The latter is naturally induced by the world affine connection $\Gamma^{\lambda}_{\mu\nu}$. The coefficients of the local Lorentz connection read as follows:

$$\Gamma^{a}_{b\mu} = h^{a}_{\alpha} h^{\beta}_{b} \Gamma^{\alpha}_{\beta\mu} + h^{a}_{\alpha} \partial_{\mu} h^{\sigma}_{b}.$$ (1.20)

Here, the quantity $h^{a}_{\mu}$ is defined as an inverse matrix to $h^{\mu}_{a}$:

$$h^{a}_{\mu} h^{\nu}_{a} = \delta^{\nu}_{\mu}, \quad h^{a}_{\mu} h^{\mu}_{b} = \delta^{a}_{b}.$$  

The frame $e_{a}$ is called nonholonomic, if the corresponding tetrad coefficients $h^{a}_{\mu}$ are no partial derivatives of some four functions $f^{a}(x)$, $a = 0, 1, 2, 3$, i.e., when $h^{a}_{\mu} \neq \partial_{\mu} f^{a}$. Geometrically, it means that there are no such local coordinates $y^{a} = f^{a}(x)$, with respect to which $e_{a}$ becomes a coordinate basis. The degree of deviation from the field of integrable local Lorentz frames, for which such new coordinates exist, is measured by the anholonomity object

$$C^{a}_{\mu\nu} := \partial_{\mu} h^{a}_{\nu} - \partial_{\nu} h^{a}_{\mu}.$$ (1.21)

This object is not a tensor, and it is closely related to the tetrad form of the Riemannian connection induced by the metric. As a result, one can reformulate Einstein’s GR in terms of the tetrad $h^{a}_{\mu}$ and the anholonomity object $C^{a}_{\mu\nu}$ [8].

Computing the commutator of covariant derivatives $D_{\mu} v^{a} = \partial_{\mu} v^{a} + \Gamma^{a}_{b\mu} v^{b}$ with respect to the local Lorentz connection, we obtain the local Lorentz curvature

$$R^{a}_{b\mu\nu} = \partial_{\mu} \Gamma^{a}_{b\nu} - \partial_{\nu} \Gamma^{a}_{b\mu} + \Gamma^{a}_{c\mu} \Gamma^{c}_{b\nu} - \Gamma^{a}_{c\nu} \Gamma^{c}_{b\mu}.$$
It is easy to show that the latter is related to the curvature tensor of the world affine connection,

\[ R_{\alpha \beta \mu \nu} = h_{\alpha}^{\alpha} h_{\beta}^{\beta} R_{\beta \mu \nu}. \]

Analogously, we can prove that it is possible to recast the torsion tensor into a generalized covariant “curl” of the tetrad:

\[ D_{\mu} h_{\nu}^{\alpha} - D_{\nu} h_{\mu}^{\alpha} = -2 h_{\lambda}^{\alpha} Q_{\lambda \mu \nu}. \]

The tetrad formalism represents another framework for the gravity theory. The metric and the affine connection \( (g_{\mu \nu}, \Gamma_{\alpha \beta \mu}) \) in this approach are replaced by the tetrad and the local Lorentz connection \( (h_{\mu}^{a}, \Gamma_{ab \mu}) \). It will be shown later that there is close relation of the latter variables with the structure of the principal bundle of the Lorentz group and it turns out to be possible to develop a consistent gauge approach to the gravity theory using this formulation.

Apparently, the most convenient mathematical formalism for the gauge approach in gravity is the theory of connections in the bundles over the spacetime manifold \( M_4 \). See Chapter 5 for the further details.

1.2. Hilbert variational principle and field equations of the general relativity

Einstein [22] formulated the fundamental equations of GR describing the dynamics of the interacting gravitational and matter fields on the basis of an idea that matter (its energy-momentum tensor \( T_{\mu \nu} \)) gives rise to the Riemannian curvature of spacetime. An additional requirement was that the covariant conservation law \( T_{\mu \nu ; \nu} = 0 \) should be an automatic consequence of the field equations. The latter assumption almost uniquely fixed the Einstein tensor \( G_{\mu \nu} \) as an appropriate structure describing the geometry of the Riemannian spacetime by making use of the contracted Bianchi identity (1.18) for the case of the vanishing torsion \( Q_{\lambda \mu \nu} = 0 \).

As a result, Einstein’s gravitational field equations read as follows:

\[ G_{\mu \nu} = \kappa T_{\mu \nu}, \tag{1.22} \]

where the value of the Einstein gravitational constant \( \kappa = 8\pi G/c^4 = 2.07 \times 10^{-43} \text{ kg}^{-1} \text{ m}^{-1} \text{ s}^2 \) was derived from the comparison of GR with Newton’s gravitation theory; \( G = 6.67 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2} \) is the Newtonian gravitational constant.

In a vacuum (in the absence of matter), these equations reduce to

\[ G_{\mu \nu} = 0 \quad \text{or} \quad R_{\mu \nu} = 0. \tag{1.22a} \]

At the same time, the field equations (1.22a) were derived by Hilbert [23] on the basis of the fundamental dynamical principle – the principle of the least action,
Gravity models based on Lagrangian which is linear on curvature, are called the gravity theories of the Hilbert-Einstein type.

A short remark is in order concerning the dimensions. The metric components are dimensionless, whereas the connection coefficients have a 1/(length) dimension. Accordingly, the curvature and its contractions have a dimension of 1/(length)$^2$. Consequently, the dimension of the integral $\int d^4x \sqrt{g} R$ is m$^2$. Taking into account that $\frac{1}{\kappa_c} = \text{kg/s}$, we find that the action $S_{HE}$ indeed has the dimension of the action (recall that the Planck constant is a “quantum of an action”).

By substituting the expression of the Christoffel symbols $\{_{\lambda}^{\mu
u}\}$ in terms of the metric $g_{\mu\nu}$ into $L_{HE}$, we obtain

\[
L_{HE} = L_{HE}[g_{\mu\nu}, g_{\mu\nu,\lambda\kappa}, g_{\mu\nu,\kappa\lambda}] = g^{\mu\nu} \left(2\{_{\lambda}^{\mu\nu}\}_{,\lambda\kappa} + 2\{_{\sigma\lambda}^{\mu\nu}\}_{,\sigma\mu}\right)
\]

Commas denote the usual partial derivatives, $,\lambda := \partial_{\lambda}$, etc. Let us vary the action $S_{HE}$ with respect to the metric $g_{\mu\nu}$. Then

\[
\delta g S_{HE} = \frac{1}{2\kappa_c} \int d^4x \sqrt{g} G_{\mu\nu} \delta g^{\mu\nu}
\]

\[
+ \frac{1}{2\kappa_c} \int_{\partial \Omega} d\sigma_{\lambda} \sqrt{g} g^{\mu\nu} \left[ -\delta_{g} \{_{\lambda}^{\mu\nu}\} + \delta_{\mu} \delta_{g} \{_{\rho}^{\nu}\} \right] = 0, \quad (1.23)
\]

\[
d\sigma_{\lambda} = \frac{1}{3!} \varepsilon_{\lambda\mu\nu\chi} dx^\mu dx^\nu dx^\chi.
\]

Einstein’s equations (1.22a) follow from (1.23), if and only if the surface term vanishes due to the conditions imposed by the respective boundary problem

\[
\int_{\partial \Omega} d\sigma_{\lambda} \sqrt{g} g^{\mu\nu} \left[ -\delta_{g} \{_{\lambda}^{\mu\nu}\} + \delta_{\mu} \delta_{g} \{_{\rho}^{\nu}\} \right] = 0, \quad (1.24)
\]

that is when the variational principle is consistent with the boundary value problem of the resulting Euler-Lagrange equations. The boundary conditions specify the appropriate test functions, and thereby restrict the class of the variational problems [10].

Einstein’s equations are the second-order partial differential equations with respect to metric $g_{\mu\nu}$ in the spacetime domain $\Omega$, hence an appropriate Dirichlet boundary value problem for them fixes the values of the metric at the boundary $\partial \Omega$ of the spacetime domain, $g_{\mu\nu}\mid_{\partial \Omega} = C_{\mu\nu}$. However, the corresponding condition on the variations, $\delta g_{\mu\nu}\mid_{\partial \Omega} = 0$ is clearly not sufficient to make the surface term (1.24) vanish, in general. The only exceptions are the gravitational fields of isolated matter configurations (the asymptotically flat metrics).
Strictly speaking, the statement that a physical theory is based on the variational principle $\delta S = 0$ is valid only when the system of the Euler-Lagrange equations is solved under the boundary conditions obtained from the stationarity condition of the action $S$. In other words, the boundary conditions obtained from the stationarity conditions should not contradict the order of Euler-Lagrange equations, i.e., the corresponding Lagrangian is a proper (nondegenerate) one.

For a given boundary value problem, a proper (nondegenerate) Lagrangian is not linear in the higher derivatives of the field functions. Making use of this definition, we conclude that the Hilbert Lagrangian of the gravitational field $L_{HE} = R$ is degenerate [11].

Let us illustrate the above by the example from mechanics. Consider the two Lagrangians $L_1 = -\frac{1}{2}q\ddot{q}$, and $L_2 = \frac{1}{2}\dot{q}^2$ ($q$ is the generalized coordinate, the dot denotes the time differentiation). Variations of the respective action functionals are given by

$$\delta S_1 = -\int_{t_1}^{t_2} \dot{q}\delta\dot{q} \, dt - \frac{1}{2}q\delta\dot{q}\bigg|_{t_1}^{t_2} + \frac{1}{2}\dot{q}\delta\dot{q}\bigg|_{t_1}^{t_2}, \quad (1.25)$$

$$\delta S_2 = -\int_{t_1}^{t_2} \dot{q}\delta\dot{q} \, dt + \frac{1}{2}\dot{q}\delta\dot{q}\bigg|_{t_1}^{t_2}. \quad (1.26)$$

Both actions yield the same Euler-Lagrange equation $\ddot{q} = 0$. However, eq. (1.25) tells us that in order to obtain the equation of motion, it is necessary to fix the values of the coordinate $q(t_1), q(t_2)$, and the values of the velocities $\dot{q}(t_1), \dot{q}(t_2)$ at the boundary. But this is impossible in view of the order of these equations. The attempt in (1.25) to impose only two conditions, for example ($q/\dot{q})\big|_{t_1}^{t_2}$, does not lead to a physically meaningful result. As for (1.26), the equation of motion $\ddot{q} = 0$ follows from the usual variational problem with the fixed endpoints $\delta q(t_1) = \delta q(t_2) = 0$, consistent with the correct boundary value problem.

In a sense, $L_1$ in this mechanical example is similar to the Hilbert Lagrangian $L_{HE}$, and $L_2$ is similar to the so-called truncated Lagrangian $L_D$. The latter is a proper Lagrangian that leads to the self-consistent boundary value problem $G_{\mu\nu} = 0, g_{\mu\nu}\big|_{\partial\Omega} = C_{\mu\nu} = \text{const.}$

It is worthwhile to note that $L_D$ (not $L_{HE}$) up to a boundary term coincides with the Lagrangian in the gauge theory of the translation group [13]:

$$L_D = G = g^{\mu\nu} \left( -\{\lambda^\sigma\}_\mu \{\lambda^\nu\}_\rho + \{\lambda^\lambda\}_\mu \{\lambda^\rho\}_\nu + \{\lambda^\rho\}_\mu \{\lambda^\sigma\}_\nu \right),$$

see [40]. The latter is a proper Lagrangian that leads to the self-consistent boundary value problem $G_{\mu\nu} = 0, g_{\mu\nu}\big|_{\partial\Omega} = C_{\mu\nu} = \text{const.}$

In Hawking’s approach [14], the truncated action is written as

$$S_D = S_{HE} + 2 \int_{\partial\Omega} d^4x \sqrt{g} K,$$
where the ordinary Hilbert-Einstein action is explicitly supplemented by the “surface term”. The boundary term is an integral of $K$ which is the difference of traces of the second fundamental form of boundary $\partial \Omega$ for the metric $g$ and for the flat space metric. The role of the surface term is to compensate the contribution of the second derivatives of the metric in $R$, so that the Lagrangian becomes quadratic in the first derivatives (which means a nondegenerate), as required in the path integration method.

Along with Hilbert’s method of derivation of Einstein’s field equations from the variational principle with the action linear in the curvature, the gravitational field equations can be obtained within the framework of the first-order formalism, commonly called a Palatini principle [28], which we present in the next section.

1.3. First-order formalism and field equations of the Einstein-Cartan theory

In the derivations above, the Riemannian character of the spacetime structure was essential. However, we now recall that the metric and connection are defined as a priori independent structures on a differentiable manifold. Therefore, from the geometrical point of view, it would be more natural to consider $g_{\mu\nu}$ and $\Gamma^\lambda_{\mu\nu}$ as the independent dynamical variables of the gravitational theory. This idea underlied the variational principle of Palatini, who did not impose the metricity condition (1.8) as a constraint equation on $g_{\mu\nu}$ and $\Gamma^\lambda_{\mu\nu}$. However, he kept the torsion tensor equal zero: $Q^\lambda_{\mu\nu} = 0$.

In this case, the field equations are obtained from the independent variation of the Hilbert action $S_{HE} = \frac{1}{2\kappa c} \int_\Omega d^4x \sqrt{g} R(g_{\mu\nu}, \Gamma^\lambda_{(\mu\nu),\alpha}, \Gamma^\lambda_{(\mu\nu)})$ with respect to the metric $g_{\mu\nu}$ and the symmetric connection $\Gamma^\lambda_{(\mu\nu)}$. The variation of $S_{HE}$ with respect to the metric yields

$$G_{\mu\nu} (g_{\mu\nu}, \Gamma^\lambda_{(\mu\nu),\alpha}, \Gamma^\lambda_{(\mu\nu)}) = 0, \quad (1.27)$$

and from the variation with respect to $\Gamma^\lambda_{\mu\nu}$ one obtains the metricity condition (1.8). If the latter is solved with respect to $\Gamma^\lambda_{\mu\nu}$ (under the condition $Q^\lambda_{\mu\nu} = 0$) and the result is plugged back into (1.27), we obtain Einstein’s field equation (1.22a) as before.

The vanishing torsion condition $Q^\lambda_{\mu\nu} = 0$ is the constraint equation for the dynamical variables $(g_{\mu\nu}, \Gamma^\lambda_{(\mu\nu)})$. This fact was taken into account from the beginning of the derivation of the field equations by assuming $\Gamma^\lambda_{\mu\nu} = \Gamma^\lambda_{(\mu\nu)}$. If the constraints are not resolved from the very beginning, they should be added to the original Lagrangian with the undetermined Lagrange multipliers. Then the problem of derivation of the field equations becomes a variational problem on the conditional extremum of the action [15, 16, 17]. If the constraints are imposed after the field equations are derived, the resulting system would not be equivalent to the case, when the field equations are obtained by taking into
account constraints even before the variation. The exceptions arise when the Lagrange multipliers vanish due to field equations\(^1\).

We thus have demonstrated that Einstein’s field equations in vacuum can be obtained by the Palatini variational method provided the torsion is equal zero. As we show below, by removing the restriction \(Q^\lambda_{\mu\nu} = 0\) it is possible to obtain more general gravitational field theories.

Let us consider the gravitational field dynamics based on the Palatini principle [18], assuming the full independence of the metric tensor from the general affine connection. The variation of Hilbert-Einstein type action (the subscript \(P\) stands for “Palatini”)

\[
S_P = \frac{1}{2\kappa c} \int d^4x \sqrt{g} L_P = \frac{1}{2\kappa c} \int d^4x \sqrt{g} R(\Gamma)
\]

yields, in the absence of matter (in vacuum):

\[
\frac{\delta S_P}{\delta g_{\mu\nu}} = -G^{(\mu\nu)}(\Gamma) = 0, \quad (1.28)
\]

\[
\frac{\delta S_P}{\delta \Gamma^\lambda_{\mu\nu}} = -\nabla_\lambda g^{\mu\nu} + \delta^\lambda_\varphi \nabla_\varphi g^{\mu\nu} + 2Q^\rho_{\mu\nu} + g^{\mu\nu}D_\lambda - D^\mu \delta^\nu_\lambda = 0, \quad (1.29)
\]

where \(D_\lambda = D^\rho \lambda_\rho\) is the connection defect trace.

By splitting (1.29) into the symmetric and antisymmetric parts, we obtain:

\[
Q^\lambda_{\mu\nu} = -\frac{2}{3} \delta^\lambda_\nu Q_\mu, \quad (1.30)
\]

\[
\nabla_\lambda g_{\mu\nu} = \frac{4}{3} Q_\lambda g_{\mu\nu}. \quad (1.31)
\]

By solving the last equation with respect to \(\Gamma^\lambda_{\mu\nu}\), we obtain the relation between the connection, the metric and the torsion

\[
\Gamma^\lambda_{\mu\nu} = \{^\lambda_{\mu\nu}\} - \frac{2}{3} \delta^\lambda_\mu Q_\nu. \quad (1.32)
\]

Substituting (1.32) in (1.28), we find

\[
G_{\mu\nu} = 0. \quad (1.33)
\]

Summarizing, the following results were obtained by using the consistent Palatini principle:

− For the metric, the field equations (1.33) coincide with Einstein’s equations of GR in vacuum.

− The torsion is expressed in terms of its trace \(Q_\lambda\) (1.30) (i.e., the connection (1.32) is semi-symmetric [24]) and remains arbitrary since the contraction of the equation (1.30) yields an identity.

\(^1\)If the Lagrangian is supplemented by the term \(\lambda^{\lambda\mu\nu} Q_{\lambda\mu\nu}\) containing the Lagrange multipliers \(\lambda\), the result is not changed, since \(\lambda^{\lambda\mu\nu} = 0\) due to the field equation.
1.3. First-order formalism and field equations of the Einstein-Cartan theory

- The covariant derivative of the metric (the nonmetricity) does not vanish (1.31) (the semi-metric parallel transport [24]).

However, the results obtained are not satisfactory since the torsion trace \( Q_\lambda \) is not fixed, therefore, all the main geometrical characteristics of the manifold are not determined. To find the way out of this situation, it is necessary to consider the variation principle with a different Lagrangian (see Chapter 5).

If we consider the theory with a Hilbert-Einstein Lagrangian, but require the preservation of vector lengths under the parallel transport, we obtain the Einstein-Cartan theory of gravity. The constraints (1.8) can be resolved from the very beginning and taken into account in action [7] or, without resolving them explicitly, they can be added to the Lagrangian with the undetermined Lagrange multipliers, and then the problem of derivation of the field equations becomes a variational problem on the constrained extremum of action [15, 16].

We consider the second procedure, including now into the dynamical scheme, along with the gravity, the matter fields interacting with the gravitational field. The gravitational interaction is introduced into the matter Lagrangian \( L_m \) on the basis of the minimal coupling principle by replacing the Minkowski metric \( \eta_{\mu\nu} \) with the metric of the curved space \( g_{\mu\nu}(x) \), and of the partial derivatives

\[
\nabla_\mu \Phi^A = \partial_\mu \Phi^A - \Gamma^\alpha_{\beta\mu} \Omega^\alpha_{\beta\mu} \Phi^B.
\]

(1.35)

The total action reads

\[
S_{ECT} = \int d^4x \sqrt{g} \left\{ \frac{1}{2\kappa c} R(\Gamma) + \frac{1}{c} L_m(\Phi^A, \nabla_\mu \Phi^A, g_{\mu\nu}) + \frac{1}{2c} \Lambda^\lambda_{\mu\nu} \nabla_\lambda g_{\mu\nu} \right\},
\]

(1.34)

where \( \Phi^A \) are the matter field variables (\( A \) is the generalized index), \( \Lambda^\lambda_{\mu\nu} \) are the Lagrange multipliers. The matter Lagrangian depends on the connection \( \Gamma^\alpha_{\beta\mu} \) via the covariant derivatives only

By varying (1.34) with respect to \( g_{\mu\nu}, \Gamma^\lambda_{\mu\nu}, \Lambda^\lambda_{\mu\nu}, \) and \( \Phi^A \), we obtain the following field equations:

\[
\frac{\delta S_{ECT}}{\delta g_{\mu\nu}} = \frac{1}{2c} \left\{ \frac{1}{\kappa} G(\mu\nu)(\Gamma) + (\nabla_\lambda - 2Q_\lambda)\Lambda^\lambda_{\mu\nu} - T_{\mu\nu} \right\} = 0,
\]

(1.36)

\[
\frac{\delta S_{ECT}}{\delta \Gamma^\alpha_{\mu\nu}} = \frac{1}{c} \left\{ \frac{1}{\kappa} (Q^\nu_{\alpha\beta} + 2\delta^\nu_{[\alpha} Q_{\beta]} ) g^{\beta\mu} + N_\alpha^{\mu\nu} - \Lambda^\nu_{\alpha\mu} \right\} = 0,
\]

(1.37)

\[
\frac{\delta S_{ECT}}{\delta \Lambda^\lambda_{\mu\nu}} = \frac{1}{2c} \nabla_\lambda g_{\mu\nu} = 0,
\]

(1.38)

\[
\frac{\delta S_{ECT}}{\delta \Phi^A} = \frac{1}{c} \left\{ \frac{\partial L_m}{\partial \Phi^A} - (\nabla_\lambda - 2Q_\lambda) \frac{\partial L_m}{\partial \nabla_\lambda \Phi^A} \right\} = 0,
\]

(1.39)
where (1.36) and (1.37) are given in the final form after taking into account the metricity constraint (1.38). Here, in a usual way we defined

\[ T_{\mu\nu} := \frac{2}{\sqrt{g}} \frac{\partial \left( \sqrt{g} L_m \right)}{\partial g_{\mu\nu}} \]

the metrical energy-momentum tensor of matter, and we denoted

\[ N_\alpha^{\mu\nu} = \frac{\partial L_m}{\partial \Gamma^\alpha_{\mu\nu}}. \]

Making use of the eq. (1.35), we find

\[ N_\alpha^{\mu\nu} = \frac{\partial L_m}{\partial \nabla_{\nu} \Phi A} \frac{\partial \nabla_{\lambda} \Phi^A}{\partial \Gamma^\alpha_{\mu\nu}} = - \frac{1}{\kappa} \left( Q^{\nu}_{\alpha\beta} + 2 \delta^\nu_{[\alpha} Q_{\beta]} \right) + N_{\alpha\beta}^\nu - \Lambda_{\beta\alpha}^\nu = 0. \]

Let us decompose it into the symmetric and antisymmetric parts. From the symmetric one, we find the Lagrange multipliers explicitly

\[ \Lambda_{\alpha\beta}^\nu = N_{(\alpha\beta)}^\nu, \]

so that (1.36) is recast into

\[ G_{(\mu\nu)} (\Gamma) = \kappa \left\{ T_{\mu\nu} - (\nabla_{\lambda} - 2Q_{\lambda}) N_{(\mu\nu)}^\lambda \right\}. \] (1.41)

The remaining antisymmetric part of (1.37) reads as follows:

\[ Q^{\nu}_{\alpha\beta} + 2 \delta^\nu_{[\alpha} Q_{\beta]} = \kappa c S^{\nu}_{\alpha\beta}, \] (1.42)

where

\[ c S^{\nu}_{\alpha\beta} = N_{[\alpha\beta]}^\nu = - \frac{\partial L_m}{\partial \nabla_{\nu} \Phi A} \Omega_{[\alpha\beta]}^A B \Phi^B \]

coincides with the canonical tensor of the spin density arising from the Noether theorem.

For the matter whose Lagrangian does not depend on the connection (for the scalar fields, for example), or for the vacuum gravitational field, the equations (1.42) yield \( Q^\lambda_{\mu\nu} = 0 \) and then (1.41) reduce to the usual Einstein’s equations.

Otherwise, the torsion is non-zero and (1.41) differ from the GR equations. One can show that the term \((\nabla_{\lambda} - 2Q_{\lambda}) N_{(\mu\nu)}^\lambda \) leads to the fact that the sources of the gravitational field are the canonical tensor of spin \( S^\lambda_{\mu\nu} \), and the symmetrical part of the canonical tensor of the energy-momentum of matter

\[ t_{\mu\nu} : = \delta^\mu_{\nu} L_m - \frac{\partial L_m}{\partial \nabla_{\mu} \Phi A} \nabla_{\nu} \Phi^A. \]
Indeed, let us consider an infinitesimal transformation of local coordinates $x^\mu \rightarrow \overline{x}^\mu = x^\mu + \xi^\mu \delta \tau$, where $\xi^\mu(x^\lambda)$ is an arbitrary smooth vector field, and $\delta \tau$ is an infinitesimal parameter. The Lie derivative of the Lagrangian density $\sqrt{g} L_m(\Phi^A, \nabla_\mu \Phi^A, g_{\mu\nu})$ along the vector field $\xi^\mu$, which induces these transformations, is computed as follows:

$$L_\xi (\sqrt{g} L_m) = \frac{\partial (\sqrt{g} L_m)}{\partial \Phi^A} L_\xi \Phi^A + \frac{\partial (\sqrt{g} L_m)}{\partial \nabla_\mu \Phi^A} L_\xi \nabla_\mu \Phi^A + \frac{\partial (\sqrt{g} L_m)}{\partial g_{\mu\nu}} L_\xi g_{\mu\nu}. \tag{1.43}$$

Using the definition of Lie derivative (which is introduced on the spacetime manifold even in the absence of the metrics and connection [3]), it is straightforward to see that in the Riemann-Cartan space

$$L_\xi \Phi^A = \xi^\lambda \nabla_\lambda \Phi^A + \zeta_\beta^\alpha \Omega_{\alpha \beta} \Phi^B, \tag{1.43a}$$

$$L_\xi \nabla_\mu \Phi^A = \xi^\lambda \nabla_\lambda \nabla_\mu \Phi^A + \zeta_\beta^\alpha \Omega_{\alpha \beta} \nabla_\mu \Phi^B + \zeta_\mu^\nu \nabla_\nu \Phi^A, \tag{1.43b}$$

$$L_\xi g_{\mu\nu} = 2 \zeta_{(\mu\nu)}, \tag{1.43c}$$

where $\zeta_{\mu\nu} = \nabla_\mu \xi^\nu + 2 Q^\nu _{\mu \lambda} \xi^\lambda$.

On the other hand, since $(\sqrt{g} L_m)$ is a scalar density of the weight +1, we have

$$L_\xi (\sqrt{g} L_m) = \xi^\lambda \nabla_\lambda (\sqrt{g} L_m) + \zeta_\mu^\mu (\sqrt{g} L_m). \tag{1.43d}$$

Substituting (1.43a)-(1.43c) into (1.43) and comparing with (1.43d), we find

$$-t^\mu_\nu + T^\mu_\nu - (\nabla_\lambda - 2 Q_\lambda) N_\nu^\mu \lambda + c \frac{\delta S_{ECT}}{\delta \Phi^A} \Omega_{\nu^A \mu^B} = 0, \tag{1.44}$$

$$(\nabla_\mu - 2 Q_\mu) t^\mu_\nu + 2 Q^\lambda _{\nu \mu} t^\mu_\lambda + c S^\lambda _{\alpha \beta} R^{\alpha \beta} \nabla_\nu \Phi^A = c \frac{\delta S_{ECT}}{\delta \Phi^A} \nabla_\nu \Phi^A. \tag{1.45}$$

By virtue of the field equations (1.39), we get from (1.44)

$$t_{(\mu\nu)} = T_{\mu\nu} - (\nabla_\lambda - 2 Q_\lambda) N_{(\mu\nu)}^\lambda, \tag{1.46}$$

$$t_{[\mu\nu]} = c (\nabla_\lambda - 2 Q_\lambda) S^\lambda_{\mu\nu}. \tag{1.47}$$

Thus we see that eq. (1.46) establishes the relation between the symmetrical part of the canonical and the metrical tensors of energy-momentum of matter. Making use of this relation in (1.41), we find the complete system of the field equations of the Einstein-Cartan theory:

$$G_{\mu\nu} (\Gamma) = \kappa t_{\mu\nu}, \tag{1.48}$$

$$Q^\lambda _{\mu\nu} + \delta^\lambda _{\mu} Q_{\nu} - \delta^\lambda _{\nu} Q_{\mu} = \kappa c S^\lambda_{\mu\nu}. \tag{1.49}$$

It is worthwhile to note that to determine the dynamics of the metric it is sufficient to solve only the symmetric part of (1.48) which is equivalent to (1.41). The skew-symmetric part of (1.48) is redundant and is automatically satisfied in view of (1.49) and the contracted Ricci identity.
Summarizing, within the framework of the Einstein-Cartan theory, the matter source of gravitational field \((g, \Gamma)\) is described by the canonical energy-momentum tensor \(t^\lambda_\mu\) and the spin density tensor \(S^\lambda_\mu\). The invariance of the theory with respect to the general coordinate transformations and the local Lorentz group leads to the conservation law of energy-momentum (1.45) and of the total angular momentum (1.47).

The important feature of the Einstein-Cartan theory (ECT) is the algebraic nature of the equation (1.49), which will be called a Palatini equation (sometimes it is also called a Cartan equation). As a consequence of such an algebraic coupling, the torsion disappears in the absence of a spinning matter, which confirms Cartan’s prediction \([19]\). Furthermore, the entire Einstein-Cartan theory can be written in terms of the Riemann-Einstein objects. Indeed, using the expression for the connection (1.12), we can explicitly separate the Riemann-Einstein part in the action

\[
S_{ECT} = \frac{1}{2\kappa c} \int d^4x \sqrt{g} R(\Gamma) = \frac{1}{2\kappa c} \int d^4x \sqrt{g} (R + 2T^\alpha_\mu [\nu T^\nu_\beta]) \tag{1.50}
\]

plus the surface integral.

Next, we solve eq. (1.49) expressing the torsion in terms of the spin

\[
Q^\lambda_\mu = \kappa c \left( S^\lambda_\mu + \frac{1}{2} \delta^\lambda_\mu S + \frac{1}{2} \delta^\lambda_\mu S \right),
\]

with \(S_\mu = S^\lambda_\mu \gamma^\lambda\). Inserting this into (1.50), we obtain the action

\[
S_{ECT} = \int d^4x \sqrt{g} \left( \frac{1}{2\kappa c} R + \frac{1}{c} L_{eff}^m \right), \tag{1.51}
\]

where the effective matter Lagrangian reads

\[
L_{eff}^m = L_m - \kappa c \left[ S^\mu_\nu S_{\mu} + \frac{1}{2} S^\mu_\nu (S_{\nu}^\lambda + S^\lambda_\mu + S^\mu_\lambda) \right].
\]

The field equations (1.48) and (1.49) are then recast into

\[
G^\mu_\nu = \kappa T^\mu_\nu, \tag{1.52}
\]

where the effective metrical tensor of the energy-momentum of matter

\[
T^\mu_\nu_{\text{eff}} = \frac{2}{\sqrt{g}} \frac{\delta (\sqrt{g} L_{eff})}{\delta g_{\mu\nu}}
\]

satisfies the covariant conservation law, \(T^\mu_\nu_{\text{eff}} = 0\), of GR \([188]\).

Thus, the spin-torsion interaction in the Einstein-Cartan theory predicts the appearance of the (effective) matter Lagrangian of the terms describing the contact spin-spin interaction, which is \(\kappa c^2\) times smaller than the interaction between the energy-momentum and the metric in GR. The change of the matter current \(T^\mu_\nu \rightarrow T^\mu_\nu_{\text{eff}}\) affects the spacetime metric via the effective Einstein equation (1.52). In those areas, where the spinning matter is absent, the influence of the spin-spin interaction manifests itself through the matching conditions for the spin discontinuity at the boundary \([21]\).
2
Canonical formalism of gravity theories

2.1. Systems with singular Lagrangians

The gravitational field theory is invariant with respect to the group of general coordinate transformations, as a result, the dynamics of this field is essentially hidden due to the arbitrary choice of a coordinate system. The dynamical contents of any theory as such (i.e., the calculation of the number of independent degrees of freedom, their explicit selection from the full set of the field variables, the construction of such dynamical characteristics as the energy, etc.) can be revealed by the methods of the canonical formalism. The solution of the problems listed above is necessary for the canonical quantization of the theory, which leads to the unitary $S$-matrix.

Therefore, here and in the next chapter, we will study the specific features of the Hamiltonian gravitational dynamics. We begin with considering of the dynamical consequences of the invariance of the theory with respect to the local group.

Degeneracy of the dynamical system as a result of invariance with respect to the local group

Let us find out what are the dynamical consequences of the invariance of an action of the field system $\Phi^A(x)$ ($A = 1, \ldots, k$)

$$S[\Phi] = \int d^4x \mathcal{L}[\Phi^A(x), \partial_\mu \Phi^A(x)],$$

(2.1)
with respect to transformations of the field variables and of the spacetime co-
ordinates, constituting the local group:

$$
\Phi^A(x) \rightarrow \Phi'^A(x'), \quad x^\mu \rightarrow x'^\mu(x).
$$

If

$$
\Delta^I \Phi^A(x) \equiv \Phi'^A(x) - \Phi^A(x),
\quad \Delta^I x^\mu(x) \equiv x'^\mu - x^\mu
$$

are the infinitesimal transformations of this $m$-parameter local group with the
local infinitesimal parameters $f^\alpha(x) \ (\alpha = 1, \ldots, m)$, then the condition of the
invariance of the action (2.1) reads

$$
\int d^4x \left\{ \left( \frac{\partial L}{\partial \Phi^A} - \partial_\mu \frac{\partial L}{\partial \partial_\mu \Phi^A} \right) \Delta^I \Phi^A + \partial_\mu \left( L \Delta^I x^\mu + \frac{\partial L}{\partial \partial_\mu \Phi^A} \Delta^I \Phi^A \right) \right\} = 0.
$$

(2.4)

Let us consider such transformation parameters that the variations (2.2) and
(2.3) vanish outside some four-dimensional spacetime area $\Omega$, as well as its
boundary. After the integration in (2.4), we then obtain

$$
\int_{\Omega} d^4x \left( \frac{\partial L}{\partial \Phi^A} - \partial_\mu \frac{\partial L}{\partial \partial_\mu \Phi^A} \right) \Delta^I \Phi^A = 0.
$$

(2.5)

It is convenient to introduce the new condensed notation of the field variables
$\Phi^A(x) \equiv Q^i(t)$, where the dependence on time $t$ is explicitly distinguished, and the
generalized index $i = (A, x)$ labels both the discrete internal components and
the spatial coordinates. The summation over $i$ thus will also imply an
integration over $x = \{x^1, x^2, x^3\}$.

We assume that the group transformations are arbitrarily complicated local
functions of the fields, of the group parameters $f^\alpha$, and of their spacetime
derivatives up to the $N$-th order. The parameters of the group $f^\alpha$ will also be
labeled by the condensed indices $\alpha$, including the spatial coordinates.

Then we can recast (2.2) into

$$
\Delta^I Q^i(t) = \sum_{n=0}^{N} \frac{(n)}{a_{\alpha}} \partial_0^n f^\alpha(t), \quad \partial_0 \equiv \frac{\partial}{\partial t},
$$

(2.6)

where $\frac{(n)}{a_{\alpha}}$ are local in time functions of the fields ($n$ denotes the order of time
derivative in the transformation), and summation over $\alpha$ in field systems also
implies integration over $x$.

Substituting (2.6) into (2.5) and integrating by parts, we obtain

$$
\int dt f^\alpha(t) \sum_{n=0}^{N} (-1)^n \partial_0^n \left\{ \left( \frac{\partial L}{\partial Q^i(t)} - \partial_0 \frac{\partial L}{\partial Q^i(t)} \right) \frac{(n)}{a_{\alpha}} \right\} = 0,
$$
2.1. Systems with singular Lagrangians

where $\dot{Q}^i \equiv \partial Q^i / \partial t$, and $L = \int d^3 x \mathcal{L}$ is system’s Lagrange function. Due to arbitrariness of $f^\alpha(t)$, this yields the following identity:

$$\sum_{n=0}^{N} (-1)^n \partial_0^n \left\{ \left( \frac{\partial L}{\partial Q^i(t)} - \partial_0 \frac{\partial L}{\partial \dot{Q}^i(t)} \right)^{(n)}_\alpha \right\} = 0. \quad (2.7)$$

We will consider the second-order theories, then the following relation is fulfilled

$$\frac{\partial L}{\partial Q^i} - \partial_0 \frac{\partial L}{\partial \dot{Q}^i} = -\frac{\partial^2 L}{\partial \dot{Q}^i \partial \dot{Q}^k} \ddot{Q}^k + \mathcal{F} \left( Q^k, \dot{Q}^k \right), \quad (2.8)$$

where the functions $\mathcal{F}(Q^k, \dot{Q}^k)$ do not depend on the second derivatives $\ddot{Q}^i(t)$. Using (2.8) in (2.7), we obtain the identity that is valid for any $Q^i(t)$:

$$\sum_{n=0}^{N} (-1)^n \partial_0^n \left\{ \left( \frac{\partial L}{\partial Q^i} - \partial_0 \frac{\partial L}{\partial \dot{Q}^i} \right)^{(n)}_\alpha \right\} = \sum_{n=0}^{N} (-1)^n \partial_0^n \left\{ \left( \frac{\partial^2 L}{\partial \dot{Q}^i \partial \dot{Q}^k} \ddot{Q}^k + \mathcal{F} \right)^{(n)}_\alpha \right\} = 0, \quad (2.9)$$

where the functions $W_\alpha$ do not depend on the higher derivatives $\partial_0^{N+2} Q^i(t)$. Due to arbitrariness of $Q^i(t)$, both terms in (2.9) are identically zero, since the derivatives of the $(N+2)$-th order do not depend on the lower derivatives. Therefore, $(\sum_{n=0}^{N} (-1)^n \partial_0^n \left\{ \left( \frac{\partial^2 L}{\partial \dot{Q}^i \partial \dot{Q}^k} \ddot{Q}^k + \mathcal{F} \right)^{(n)}_\alpha \right\}) = 0$ and the matrix of the second derivatives of the Lagrangian with respect to the velocities is degenerate:

$$\det \frac{\partial^2 L}{\partial \dot{Q}^i \partial \dot{Q}^k} = 0, \quad \text{rank} \left( \frac{\partial^2 L}{\partial \dot{Q}^i \partial \dot{Q}^k} \right) = k - m. \quad (2.10)$$

The Lagrangians with the property (2.10) are called singular, and the related systems are called degenerate. The equations of motion for such systems

$$\frac{\partial^2 L}{\partial \dot{Q}^i \partial \dot{Q}^k} \ddot{Q}^k - \mathcal{F} \left( Q, \dot{Q} \right) = 0 \quad (2.11)$$

cannot be solved with respect to the higher derivatives, and $m$ linear combinations of these equations, containing only the coordinates $Q^i$ and their velocities $\dot{Q}^i$, vanish on the extremals of the degenerate action (2.1):

$$(\sum_{n=0}^{N} (-1)^n \partial_0^n \mathcal{F} \left( Q, \dot{Q} \right) = 0. \quad (2.12)$$

The latter means that the initial data for the coordinates and the velocities cannot be fixed independently.

Before we consider the Cauchy problem and analyse how to solve the equation (2.11) of an arbitrary degenerate system, it is instructive to illustrate our derivations by the two examples of systems with the singular Lagrangians: the relativistic particle, and Maxwell’s electrodynamics.
The action \( S = m \int dt \sqrt{-\dot{x}^\mu \dot{x}_\mu} \) of a relativistic particle \( \{ Q^i (t) \equiv x^\mu (t) \} \) of the mass \( m \) is invariant with respect to the transformations
\[
\Delta t = f(t), \quad \Delta x^\mu = \left( \frac{0}{a} \right)^{a}_\mu f, \quad \left( \frac{0}{a} \right)^{a}_\mu = x^\mu (t),
\]
and its Lagrangian \( L = m \sqrt{-\dot{x}^\mu \dot{x}_\mu} \) satisfies the relation \( \dot{x}^\mu \partial^2 L / \partial \dot{x}_\mu \partial \dot{x}^\nu = 0 \).

Interesting property of the relativistic particle action is the vanishing of its Hamiltonian,
\[
H = \dot{x}^\mu \left( \partial L / \partial \dot{x}_\mu \right) - L = 0.
\]

The electromagnetic field action with \( \{ Q^i \equiv A_\mu (t, x) \} \),
\[
S = -\frac{1}{4} \int d^4 x F^\mu_\nu F^\nu_\mu,
\]
\( F^\mu_\nu = \partial_\mu A_\nu - \partial_\nu A_\mu \), is invariant with respect to the Abelian one-parameter group \( \Delta f A_\mu = \partial_\mu f(x) \), so that
\[
\left( \frac{1}{a} \right)^{a}_\mu (x, x') = \delta^a_\mu \delta(x - x'), \quad (2.13)
\]
and the degeneracy condition (2.10) of the system reads
\[
\int d^3 x' \left( \frac{1}{a} \right)^{a}_\mu (x, x') \frac{\partial^2 L}{\partial A_\mu (x') \partial A_\nu (y)} = \frac{\partial^2 L}{\partial A_0 (x) \partial A_\nu (y)} = 0.
\]

**Canonical formalism for systems with singular Lagrangians**

Thus far, it is shown that the invariance of a dynamical system with respect to a local group leads to its degeneracy. The dynamics of the system and its Cauchy problem are usually studied within the framework of the canonical formalism, so let us consider the specific features of the canonical formalism for systems with singular Lagrangians.

Define the momenta conjugated to the canonical coordinates \( Q^i \):
\[
\mathcal{P}_i = \frac{\partial L}{\partial \dot{Q}^i}.
\]

Since the matrix of the second derivative of the Lagrangian with respect to the velocities is degenerate (2.10), the equations (2.14) cannot be solved with respect to the velocities in terms of the momenta. Technically, this means that not all \( k \) momenta \( \mathcal{P}_i \) can be taken as independent arguments. Looking at the rank of the matrix \( \partial^2 L / \partial \dot{Q}^i \partial \dot{Q}^k \), we conclude that there are \( m \) identities connecting momenta and coordinates,
\[
\varphi_\mu (Q, \mathcal{P}) = 0. \quad (2.15)
\]
These identities are called the primary constraints [40]. Now let us consider the function of the three sets of arguments \( Q^i, \dot{Q}^i \) and \( \mathcal{P}_i \):
\[
H_0 = \mathcal{P}_i \dot{Q}^i - L(Q^i, \dot{Q}^i).
\]

Let us calculate the variation of this quantity with all the arguments treated as independent ones, and then put it in the space of the variables \( (Q, \dot{Q}, \mathcal{P}) \) on
the surface defined by the equation (2.14):

\[ \delta H_0 \bigg|_{\dot{p}_i = \partial L / \partial \dot{q}^i} = \delta p_i \dot{q}^i - \frac{\partial L}{\partial q^i} \delta q^i. \]

It is obvious that even in spite of the absence of the functional independence of the arguments \( Q \) and \( P \), one can express the quantity (2.16) as a function of coordinates and momenta only, which on the surface (2.14) does not depend on the velocities \( \dot{q}^i \):

\[ H_0 (Q, P) = \left\{ \mathcal{P}_i, \dot{q}^i - L(Q, \dot{Q}) \right\} \bigg|_{\mathcal{P}_i = \partial L / \partial \dot{q}^i}. \tag{2.17} \]

We now develop the first-order variational formalism with the Hamiltonian (2.17). For non-degenerate theories with the independent coordinates and momenta, this formalism is based on the variation with respect to \( Q \) and \( P \) of the action

\[ S[Q, P] = \int dt \left\{ \mathcal{P}_i \dot{q}^i - H_0(Q, P) \right\}. \]

In the theories with singular Lagrangians, it is necessary to study a conditional extremum of the same functional subject to the conditions defined by the constraints (2.15). This is achieved by the inclusion of the constraints with the Lagrange multipliers \( \lambda^\mu \) into the action functional,

\[ S[Q, P, \lambda] = \int dt \left\{ \mathcal{P}_i \dot{q}^i - H_0(Q, P) - \lambda^\mu \varphi^\mu(Q, P) \right\}. \tag{2.18} \]

The equations, obtained by the variation of this action, read as follows:

\[ \begin{align*}
\dot{p}_i &= \left\{ \mathcal{P}_i, H_0 \right\} + \lambda^\mu \left\{ \mathcal{P}_i, \varphi^\mu \right\}, \\
\dot{q}^i &= \left\{ q^i, H_0 \right\} + \lambda^\mu \left\{ q^i, \varphi^\mu \right\}, \\
\varphi^\mu &= 0,
\end{align*} \tag{2.19} \]

where \( \left\{ , \right\} \) is the Poisson bracket determined by the phase variables \((Q, P)\).

However, it is unclear that the total Hamiltonian in the action (2.18) is written correctly, i.e., that all the occurring constraints are taken into account. Indeed, the condition of consistency of the constraints (2.15) with the equations of motion (2.19) means that

\[ \dot{\varphi}^\mu = 0, \quad \left\{ \varphi^\mu, H_0 \right\} + \lambda^\alpha \left\{ \varphi^\mu, \varphi^\alpha \right\} = 0. \tag{2.20} \]

Demanding that this equation is satisfied, we have one the following options:

1. \( \det \left\{ \varphi^\mu, \varphi^\alpha \right\} \neq 0 \). Then the equations (2.20) explicitly determine the Lagrange multipliers, and substituting the latter into (2.19) we get the system of equations that have a unique solution. The time dependence of the field variables \( q^i(t) \) is then uniquely defined.
Among the set \( (2.15) \), there are such constraints \( \varphi_q (q = 1, \ldots, r \leq m) \) that \( \{ \varphi_q, \varphi_\alpha \} = C_{q}^{\mu} \varphi_\mu \), where \( C_{q}^{\mu} \) are some functions of the phase coordinates. Then \( \det \{ \varphi_\mu, \varphi_\alpha \}_{\varphi=0} = 0 \) and the equations \( (2.20) \) yield new \( r \) equations

\[
\chi_q = 0, \quad \chi_q (Q, P) \equiv \{ \varphi_q, H_0 \},
\]

which are called the secondary constraints. Since the phase coordinates must satisfy these new relations, it is necessary to add them to the action with the new Lagrange multipliers. Repeating the process, we again obtain one of the two aforementioned options. Such a sequence of steps can be completed only as follows. We denote the complete set of \( N \) constraints obtained during this process by

\[
\Psi_{\Phi} = 0, \quad \Psi_{\Phi} = (H_\alpha, \Xi_A), \quad (2.21)
\]

where \( R \) constraints \( H_\alpha \) satisfy the relations

\[
\{ H_\alpha, \Psi_A \} = C_{\alpha A}^{\Phi} \Psi_\Phi, \quad \alpha = 1, \ldots, R, \quad (2.22)
\]

and the remaining constraints \( \Xi_A \) are numbered by the index \( A = R+1, \ldots, N \). Then

\[
\{ H_\alpha, H_0 \} = C_{\alpha}^{\Phi} \Psi_\Phi, \quad (2.23)
\]

otherwise the consistency conditions of the form \( (2.20) \) would yield new constraints, and the procedure would not be completed.

The constraints that satisfy \( (2.21)-(2.23) \) are called the constraints of the first kind. They commute in the sense of the Poisson brackets on the surface of constraints in the phase space with each other and with the Hamiltonian \( H_0 \). One can show that the remaining constraints \( \Xi_A \) satisfy

\[
\det \{ \Xi_A, \Xi_B \} \neq 0. \quad (2.24)
\]

These are the constraints of the second kind.

The total canonical action takes the following form:

\[
S [Q, P, \lambda] = \int dt \left\{ P_i \dot{Q}^i - H_0 - \lambda^\alpha H_\alpha - \lambda^A \Xi_A \right\}. \quad (2.25)
\]

The condition \( (2.24) \) provides the uniqueness of the solution for \( \lambda^A \) (in terms of the remaining variables \( Q, P, \lambda^\alpha \)), while the Lagrange multipliers \( \lambda^\alpha \) corresponding to the constraints of the first kind \( H_\alpha \) remain completely arbitrary. This manifests the ambiguity in the solving of the equations of motion of a degenerate system. The origin of this ambiguity lies in the invariance of the theory with respect to the action of the local group, which leads to system’s degeneracy. Later we will demonstrate that variations of the undetermined Lagrange multipliers \( \lambda^\alpha \) correspond to the local group transformations, and that the constraints \( H_\alpha \) are the generators of these transformations in the phase space.
2.1. Systems with singular Lagrangians

We have described the general algorithm of construction of a complete system of constraints for the case of an arbitrary degenerate theory. Let us now specialize to the theories without constraints of the second kind, and give the form of canonical action (2.25) for this important particular case. We assume that by an appropriate choice of the field variables, we can recast the primary couplings (2.15) into a simple form

\[ \varphi_\mu = P_\mu, \quad P_\mu = 0. \] (2.26)

This means that the velocities \( \dot{Q}_\mu \) do not appear in the Lagrangian at all. Let us then exclude \((Q_\mu, P_\mu)\) from the complete set \((Q, P) = (Q_\mu, P_\mu; q^i, p_i)\), \(i = 1, \ldots, n\). Since the Lagrangian does not depend on \( \dot{Q}_\mu \), the variables \( Q_\mu \) play the role of the Lagrange multipliers, \( Q_\mu = \lambda_\mu \), and the canonical action is written as

\[ S[q, p, \lambda] = \int_{t_0}^{t_1} dt \left\{ p_i \dot{q}^i - H_0(q, p) - \lambda_\mu H_\mu(q, p) \right\}, \] (2.27)

where for definiteness we introduced the time parameters \( t_0 \) and \( t_1 \) which determine the initial and the final Cauchy hypersurfaces.

The conservation of the initial constraints (2.26) is a condition of stationarity of the action (2.26) with respect to the variation of \( \lambda_\mu \), or

\[ H_\mu \left( q, p \right) = 0. \] (2.28)

Let us consider degenerate systems with the canonical action of the form (2.27), where \( H_\mu(q, p) \) are the constraints of the first kind, that satisfy the so-called involution relations:

\[ \{ H_\mu, H_\nu \} = U_{\mu\nu}^{\alpha} H_\alpha, \quad \{ H_\mu, H_0 \} = V_\mu^{\nu} H_\nu. \] (2.29)

Here \( U_{\mu\nu}^{\alpha} \) and \( V_\mu^{\nu} \) are some quantities depending generally on the phase variables, and the Poisson brackets are taken in terms of the phase variables of the reduced phase space \((q, p)\).

The equations of motion, obtained from the action (2.27) for the phase variables \( \Phi(q, p) \), read as follows:

\[ \dot{\Phi} = \{ \Phi, H_0 \} + \lambda^\mu \{ \Phi, H_\mu \}. \] (2.28a)

Varying the Lagrange multipliers arbitrarily, one can obtain different solutions \( \Phi(t) \) even for the same initial conditions. This ambiguity corresponds to an

---

\(^1\) The linear dependence of the complete Hamiltonian \( H = H_0 + \lambda T \) on \( \lambda \) is explained by the relation \( \text{rank} \frac{\partial^2 H}{\partial \lambda^\nu \partial \lambda^\mu} = 0 \), the violation of which would mean that it is possible to exclude part of \( \lambda^\mu \) in terms of the other variables, in contradiction with the number \( m \) of the constraints.
uncertainty with which one can fix the state of a physical system, which is
invariant with respect to the action of a local group. If \( Q(t) = (\lambda^\mu(t), q^i(t)) \) is a
solution of the equations of motion of such system, then in view of (2.2), (2.3)
and (2.6), the following functions will satisfy the same equations
\[
\begin{align*}
\lambda^\mu(t) &= \lambda^\mu(t) + \Delta t^\mu \lambda^\mu(t), \\
q^i(t) &= q^i(t) + \Delta t^i q^i(t),
\end{align*}
\]
(2.29)
where the infinitesimal transformations of the group are defined by the formulas
(2.6) with some parameters \( f^\alpha \). Thus, the physical state in the phase space
and in the \( \lambda \)-multipliers' space at a moment \( t \) is determined up to a group
transformation (2.29). The functions \( Q(t) \), as such, do not have a direct physical
meaning. Only the equivalence classes of the field configurations connected by
group transformations (2.2)-(2.3) or the group invariants have physical meaning.

One can verify that the canonical action (2.27) is invariant with respect to
transformations of the phase coordinates and \( \lambda \)-multipliers of the following form:
\[
\begin{align*}
\delta F q^i &= \{ q^i, H_\mu \} F^\mu, \\
\delta F p_i &= \{ p_i, H_\mu \} F^\mu, \quad \text{(2.30)} \\
\delta F \lambda^\mu &= F^\mu - U_\alpha^\mu \lambda^\alpha F^\beta - V_\mu^\alpha F^\alpha, \quad \text{(2.31)}
\end{align*}
\]
with some infinitesimal parameters \( F^\mu \). It turns out that there is a one-to-one
 correspondence between the parameters \( F^\mu \) and \( f^\alpha \), so that the transformations
(2.30), (2.31) are the realizations of the group transformations (2.29) in the
configuration space, under the condition that the transformations (2.30) and
(2.31) are calculated on the extremal values of the momenta \( p_i^0 = p_i(q, \dot{q}, \lambda) \).
The extremal values of the momenta are obtained as a result of solving the
equations
\[
\dot{q}^i = \frac{\partial H_0}{\partial p_i} + \lambda^\mu \frac{\partial H_\mu}{\partial p_i}
\]
with respect to \( p_i \) in terms of variables \((q, \dot{q}, \lambda)\). Therefore, the constraints of the
first kind \( H_\mu(q, p) \) are generators of the phase transformations corresponding
to the local group transformations.

Since the dynamics of the physical state is determined not by a particular
set of the functions \((q, p)\), but by the entire equivalence class of the phase
coordinates connected by the transformations (2.30) and (2.30a), we choose the
evolution of one representative from this class to describe the dynamics. This
is done by means of imposing the additional conditions
\[
\chi^\alpha(q, p) = 0, \quad \text{(2.32)}
\]
which break the invariance with respect to the transformations (2.30), (2.30a).
Obviously, the number of gauge conditions should be equal to the number of
constraints of the first kind or the number of parameters \( m \) of the local group,
and the requirement that no nonzero $\mathcal{F}^\alpha$ would leave the gauge conditions (2.32) invariant, $\delta^\alpha \chi^\alpha = \{ \chi^\alpha, H_\beta \} \mathcal{F}^\beta = 0$, with respect to the transformations (2.30)-(2.30a), implies the invertibility of the following Faddeev-Popov matrix:

$$\det J^\alpha_\beta \neq 0, \quad J^\alpha_\beta = \{ \chi^\alpha, H_\beta \}. \quad (2.33)$$

Following the choice of a gauge, the system becomes non-degenerate and is described by the action

$$S[q, p, \xi] = \int dt (p_i \dot{q}^i - H_0 - \xi_a \psi^a), \quad (2.33a)$$

where the complete set of constraints and the corresponding Lagrange multipliers are defined as

$$\psi^a = (H_\alpha, \chi^\alpha), \quad \xi_a = (\lambda^\alpha, \pi_\beta).$$

The system of constraints $\psi^a$ is of the second kind, because due to (2.33), the matrix

$$Q^{ab} = \{ \psi^a, \psi^b \}, \quad \det Q^{ab} = - (\det \{ \chi^\alpha, H_\beta \})^2 \quad (2.34)$$

is non-degenerate. The Lagrange multipliers can be found from the condition of conservation of the constraints in time

$$\frac{d\psi^a}{dt} = \frac{\partial \psi^a}{\partial t} + \{ \psi^a, H_0 \} + \xi_b \{ \psi^a, \psi^b \} = 0,$$

which yields

$$\xi_a = \{ \psi^b, H_0 \} Q^{-1}_{ba} + \frac{\partial \psi^b}{\partial t} Q^{-1}_{ba}, \quad (2.35)$$

where $Q^{-1}_{ba}$ is the inverse matrix to (2.34). In general case, the additional conditions (2.32) may explicitly depend on time. Hence, there is a $\partial \psi^a/\partial t$ contribution in (2.35). With an account of (2.35), the equations of motion for the phase variables $\Phi$ take on the form:

$$\Phi = \{ \Phi, H_0 \} D - \{ \Phi, \psi^a \} Q^{-1}_{ab} \frac{\partial \psi^b}{\partial t}, \quad (2.36)$$

where $\{ , \} D$ is the so-called Dirac bracket [40] defined as

$$\{ A, B \}_D = \{ A, B \} - \{ A, \psi^a \} Q^{-1}_{ab} \{ \psi^b, B \}. \quad (2.37)$$

Since the evolution of the physical system is subject to the set of $2m$ constraints (2.28) and (2.32), the number of independent (physical) degrees of freedom equals $(2n - 2m)/2 = n - m$. These degrees of freedom $\Phi^* = (q^A, p^*_A)$, $A = 1, \ldots, n - m$ parameterize the initial phase coordinates $\Phi = (q_i, p_i)$ on the subspace of the full set of constraints

$$\Phi = \Phi(\Phi^*) : \quad q^i = q^i(q^*, p^*), \quad p_i = p_i(q^*, p^*), \quad \Psi^a(\Phi(\Phi^*)) \equiv 0, \quad (2.38)$$
so that the following relation holds

\[
\int dt (p_i \dot{q}^i - H_0(q,p)) \bigg|_{\Phi = \Phi^*} = \int dt \left( p_A^* \dot{q}^{*A} - H_{\text{phys}}(q^*, p^*) \right). \tag{2.39}
\]

This relation may be interpreted as a definition of the physical Hamiltonian \(H_{\text{phys}}(q^*, p^*)\) and of the canonical action in terms of \((n - m)\) physical degrees of freedom:

\[
S[q^*, p^*] = \int dt \left( p_A^* \dot{q}^{*A} - H_{\text{phys}}(q^*, p^*) \right). \tag{2.40}
\]

The equations obtained by varying of this action are equivalent to (2.36), since the action (2.40) is the same functional (2.33a), where the equations of the constraints are solved in terms of the smaller number of variables \((q^*, p^*)\). Therefore, each gauge choice (2.32) corresponds to a set of \((n - m)\) physical degrees of freedom, represented by the canonical variables \((q^*, p^*)\). One can generate a change of the gauge \(\delta \chi^\alpha(q,p)\) by transformations (2.30), (2.30a) with the specially chosen parameters \(F^\mu\) (see Chapter 7).

Let us illustrate the above theory on the examples of the relativistic particle and the theory of electromagnetic field. The momenta of the relativistic particle

\[
p_\mu = \frac{\partial L}{\partial \dot{x}^\mu} = -mg_{\mu\nu} \dot{x}^\nu \sqrt{-\dot{x}^\alpha \dot{x}^\alpha}
\]

satisfy the constraint

\[
H \equiv g^{\mu\nu} p_\mu p_\nu - m^2 = 0, \tag{2.41}
\]

which due to the identical vanishing of Hamiltonian \(H_0\) is the constraint of the first kind. Let us choose the following gauge:

\[
\chi \equiv x^0 - t = 0. \tag{2.42}
\]

The condition (2.33) is fulfilled: \(\{H, \varphi\} = 2p^0 \neq 0\). Using these quantities in (2.34) and (2.36), we obtain the equations of a straight-line motion of the free relativistic particle

\[
\dot{x}^i = \frac{p_i}{p^0}, \quad \dot{p}_i = 0.
\]

Taking into account the vanishing of the Hamiltonian \(H_0\), the action (2.33a) for the relativistic particle can be written in the form

\[
S[x^\mu, p_\mu, \lambda, \pi] = \int dt \left\{ p_\mu \dot{x}^\mu - \lambda T - \pi \chi \right\}.
\]

Choosing the spatial coordinates \(x^i = q^*\) and the momenta \(p_i = p^*\) as the physical degrees of freedom, and taking (2.42) into account, we arrive at the canonical action

\[
S[x, p] = \int dt \left( p_i \dot{x}^i - \sqrt{m^2 + p^2} \right)
\]
2.1. Systems with singular Lagrangians

with the Hamiltonian \( H_{\text{phys}} = \sqrt{m^2 + p^2} \) obtained as a result of solving the constraint (2.41) with respect to the momentum \( p_0 = -H_{\text{phys}} \).

For the electromagnetic field, the action (2.33a) can be written as

\[
S[A_i, p^i, A_0] = \int dt d^3x \left\{ p^i \dot{A}_i - \left( \frac{1}{2} p_i p^i + \frac{1}{4} F_{ik} F^{ik} \right) - (-A_0) \partial_i p^i \right\},
\]

where the momentum \( p^i \) conjugated to the vector potential \( A_i \) is equal to

\[
p^i = \partial_0 A_i - \partial_i A_0,
\]

and \( F_{ik} = \partial_i A_k - \partial_k A_i \). The role of the phase variables is played by the set \((A_i, p^i)\), and the zeroth component of the vector potential \( A_0 = \lambda \) is an undetermined Lagrange multiplier for the first-kind constraint \( H = \partial_i p^i \) commuting in the sense of the Poisson brackets with the Hamiltonian

\[
H_0 = \int d^3x \left( \frac{1}{2} p_i p^i + \frac{1}{4} F_{ik} F^{ik} \right). \tag{2.43}
\]

In accordance with the general procedure, we impose the Coulomb gauge

\[
\chi = 0, \quad \chi \equiv \partial_i A^i. \tag{2.44}
\]

The equation for the Lagrange multiplier \( A_0 \) (2.35) reduces to \( \Delta A_0 = 0 \) with \( \Delta = \partial_i \partial^i \), hence we find \( A_0 = 0 \) for the zero boundary conditions, and the equations of motion take the following form: \( \dot{A}_i = p^i \) and \( \dot{p}^i = -\Delta A_i \), equivalent to Maxwell’s equations in the gauge (2.44). The dynamically independent components of the vector potential \( q^A (A = 1, 2) \) are obtained by decomposing \( A_i(x, t) \) into the two transversal polarizations \( A_i^{(A)}(x, t) \) of the electromagnetic field satisfying (2.44).

**The Hamilton-Jacobi theory of degenerate systems**

Let us consider some solution \( \Phi_0(t) \) (obtained in an arbitrary gauge) of the equations of motion (2.28a) and insert it into the action (2.27). The action (2.27) calculated on its own extremal \( \Phi_0(t) \) becomes the function of initial and final values of the phase coordinates \( q_0 \equiv q(t_0) \) and \( q_1 \equiv q(t_1) \):

\[
S[\Phi, \lambda] \bigg|_{\Phi = \Phi_0(t)} = S(q_1, t_1 | q_0, t_0). \tag{2.45}
\]

By denoting the current value of time \( t_1 = t \) and the phase coordinates \( q_1 = q \), we treat the extremal action (2.45) as a function of these variables, \( S(t; q) \). Using the variation procedure, one can show that this function is a potential of the field of the phase momenta \( p = p(t_1) \):

\[
p^i = \frac{\partial S}{\partial q_i}. \tag{2.43a}
\]
Since the extremal momenta and the coordinates at any time satisfy the constraints (2.28), the \( m \) constraint equations are fulfilled for the function of the action \( S(t; q) \):

\[
H_{\mu} \left( q, \frac{\partial S}{\partial q} \right) = 0. \tag{2.44}
\]

Similarly, from the variational procedure it follows that \( S(t; q) \) satisfies the Hamilton-Jacobi equation with the complete Hamiltonian \( H = H_0 + \lambda^\mu H_{\mu} \).

However, in view of (2.44) this equation reduces to

\[
\frac{\partial S}{\partial t} + H_0 \left( q, \frac{\partial S}{\partial q} \right) = 0. \tag{2.45}
\]

The equations (2.44) and (2.45) are the basic equations of the Hamilton-Jacobi theory of the systems with constraints. A solution of this system of equations, i.e., its complete integral, contains \( n - m \) parameters \( \alpha^A \) (\( A = 1, \ldots, n - m \)) [a solution of the equation (2.45) depends on \( n \) nontrivial integration constants, and the remaining \( m \) equations (2.44) impose \( m \) relations on them]. Thus, the function \( S \) depends on \( (n - m) \) additional parameters, the number of which is equal to the number of dynamically independent degrees of freedom \( q^* \):

\[
S = S(t; q, \alpha). \tag{2.46}
\]

Such a number of parameters labeling the solutions of the system of equations (2.44), (2.45) does not contradict to the fact that the function \( S \) in (2.45) depends on the additional \( (n + 1) \) quantities \( (q_0; t_0) \), since the function (2.45) satisfies the same \( (m + 1) \) equations (2.44), (2.45) with respect to the arguments \( (q_0; t_0) \) and the actual number of independent parameters reduces to \( (n + 1) - (m + 1) = n - m \).

We will show that the solution of the equations of our system with constraints for some set of gauges \( \chi^\mu(q, p) \) can be found from the known action function \( S(t; q, \alpha) \) by solving the following system of equations:

\[
\frac{\partial S}{\partial \alpha^A} = \beta_A, \tag{2.47}
\]

\[
\chi^\mu \left( q, \frac{\partial S}{\partial q} \right) = 0, \tag{2.48}
\]

where \( \beta_A \) are some \( m \) arbitrary parameters.

Let us start with demonstrating that if the gauge \( \chi^\mu \) satisfies (2.33), the system of equations (2.47), (2.48) admits the unique solution \( q^i = q^i(t; \alpha, \beta) \).

Indeed, the solvability condition of this system means the absence of zero eigenvectors of the matrix \( B_{ki}, k = (A, \mu) \),

\[
B_{Ai} = \frac{\partial^2 S}{\partial \alpha^A \partial q^i}, \quad B_{\mu i} = \frac{\partial \chi^\mu}{\partial q^i} + \frac{\partial \chi^\mu}{\partial p^k} \bigg|_\rho = \frac{\partial}{\partial q^i} \frac{\partial^2 S}{\partial q^k \partial q^i}. \]
Since the equations (2.44) are fulfilled identically for any \( \alpha^A \) and \( q^i \) if \( S \) is the solution (2.46), then by differentiating (2.44) with respect to \( \alpha^A \) and \( q^i \), we obtain

\[
\frac{\partial H_\mu}{\partial p_i} \bigg|_{\alpha^A} = \frac{\partial^2 S}{\partial q^i \partial \alpha^A} = 0, \tag{2.49}
\]

\[
\frac{\partial H_\mu}{\partial p_k} \bigg|_{\alpha^A} = \frac{\partial^2 S}{\partial q^k \partial q^i} - \frac{\partial H_\mu}{\partial q^i} \bigg|_{\alpha^A}. \tag{2.50}
\]

Hence, \( \partial H_\mu/\partial p_i |_{\alpha^A} = \partial S/\partial q^i \) are the \( m \) zero vectors of the matrix \( \partial^2 S/\partial q^i \partial \alpha^A \), and consequently their linear combination with non-zero coefficients \( \gamma^\mu \) is the zero vector of \( B_{ki} \). However, in view of (2.50), we have

\[
\left. \gamma^\alpha \frac{\partial H_\alpha}{\partial p_i} \right|_{\alpha^A} = \left. \gamma^\alpha \{ \chi^\mu, H_\alpha \} \right|_{\alpha^A},
\]

whence it follows that the matrix \( B_{ki} \) is non-degenerate because of (2.33).

Now we show that the solution of the system of equations (2.47), (2.48) satisfies (2.28a) with the Lagrange multipliers (2.35).

By differentiating (2.47) with respect to time, we find

\[
\frac{\partial}{\partial \alpha^A} \frac{\partial S}{\partial t} + \frac{\partial^2 S}{\partial q^i \partial \alpha^A} \dot{q}^i = 0.
\]

On the other hand, differentiation with respect to \( \alpha^A \) of the Hamilton-Jacobi equation (2.45) gives

\[
\frac{\partial}{\partial \alpha^A} \frac{\partial S}{\partial t} = \left. \frac{\partial H_0}{\partial p_i} \right|_{\alpha^A} = \frac{\partial^2 S}{\partial q^i \partial \alpha^A},
\]

which yields

\[
\frac{\partial^2 S}{\partial q^i \partial \alpha^A} \left( \dot{q}^i - \left. \frac{\partial H_0}{\partial p_i} \right|_{\alpha^A} \right) = 0.
\]

Consequently, the expression in the parentheses is the linear combination of the zero vectors of the matrix \( \partial^2 S/\partial q^i \partial \alpha^A \) with arbitrary coefficients \( \lambda^\mu \):

\[
\dot{q}^i = \left( \frac{\partial H_0}{\partial p_i} + \lambda^\mu \frac{\partial H_\mu}{\partial p_i} \right) \bigg|_{\alpha^A}.
\]

In a similar way, the equation for \( p_i \) is obtained by differentiating of (2.45) with respect to time:

\[
\dot{p}_i = \frac{\partial^2 S}{\partial q^i \partial q^k} \dot{q}^k + \frac{\partial}{\partial q^i} \frac{\partial S}{\partial t},
\]

from which, using (2.45), (2.50) and (2.51), we obtain

\[
\dot{p}_i = -\frac{\partial H_0}{\partial q^i} - \lambda^\mu \frac{\partial H_\mu}{\partial q^i}.
\]
Thus, by requiring the preservation of the gauges (2.48) in time, one can derive
the expression (2.35) for the Lagrange multipliers.

Thus far, a solution of the system of equations (2.48), (2.47), where \( S \) satisfies
the Hamilton-Jacobi equations (2.44), (2.45), fully solves the Cauchy problem
in the gauge (2.32). In addition, it is obvious that the \( 2m \) variables \((\alpha^A, \beta^A)\)
are in one-to-one correspondence with the initial values of the dynamically
independent phase variables \((q^*, p^*)\).

### 2.2. Peculiarities of the canonical formalism
in curved spacetime

To construct the canonical formalism of a dynamical system, one needs to explicitly single out a coordinate which plays the role of time.

It is easy to choose the time in relativistically invariant theories in the flat spacetime. Identifying time with the coordinate \( x^0 \) of any Lorentz coordinate
system, so that the tangent vector to the \( x^0 \)-axis is the Killing vector of the flat spacetime, one can construct the canonical formalism of the theory.

In the curved spacetime, the situation is much more complicated. In general,
there are no time-like Killing vectors, along which the spacetime properties are homogeneous. Hence, in each coordinate system, it is impossible to select a unique time. In addition, the physical theory is invariant with respect to the
group of general coordinate transformations (2.2), (2.3)

\[
\Delta f^A(x) = L_f \Phi^A(x), \quad \Delta f^\mu(x) = f^\mu, \quad (2.52)
\]

where \( L_f \Phi^A(x) \) is the Lie derivative (1.43a) of the tensor field \( \Phi^A(x) \), containing
the term \( f^\mu \partial_\mu \Phi^A(x) \):

\[
L_f \Phi^A(x) = f^\mu \partial_\mu \Phi^A(x) + (\partial_\mu f^\nu) \Omega^{\mu A} \Phi^B(x). \quad (2.53)
\]

Thus, under any transformation of the form (2.52), the question arises, how
to define the time in new coordinate systems when it was already introduced somehow in one of them. In view of the locality of the group (2.52), the time
should be defined locally, i.e., its definition may be changed in the local vicinity
of a point without changing it in the rest of the spacetime.

Therefore, we obtain a coordinate-free definition of time as a system of the
space-like hypersurfaces of the constant parameter

\[
t = \tau(x), \quad (2.54)
\]

where \( \tau(x) \) is a scalar function of coordinates.

If we also want to define a three-dimensional coordinate chart on each hypersurface, this can be done with the help of the four functions

\[
x^\alpha = e^\alpha(x, t) \quad (2.55)
\]
of the three-dimensional coordinates \( x = \{x^a\} \) and the time \( t \), so that their substitution in (2.54) yields an identity. Since the canonical formalism on the curved spacetime essentially relies on its \((3 + 1)\)-decomposition (2.55), we will study its geometrical properties now.

### 2.3. Geometry of \((3 + 1)\)-decomposition of spacetime

The parameters \( x^a \) represent the coordinate system on the hypersurface (2.55). Let us define the normal basis on the hypersurface by the triad of the tangent vectors \( e^a_\alpha \) and the normal \( n_\alpha \), such that

\[
e^a_\alpha \equiv \partial_a e^\alpha(x, t), \quad n_\alpha e^a_\alpha = 0. \tag{2.56}
\]

The normal vector \( n_\alpha \) determines the type of a surface:

\[
n_\alpha n_\alpha = \sigma, \quad \sigma = \pm 1. \tag{2.57}
\]

The value \( \sigma = -1 \) for space-like and \( \sigma = +1 \) for time-like surfaces.

It is natural to define the three-dimensional metric of the hypersurface

\[
g_{ab}(x) = g_{\alpha\beta}(e(x, t)) e^a_\alpha e^b_\beta. \tag{2.58}
\]

Then any 4-vector \( \lambda^\alpha \) can be decomposed with respect to the normal basis:

\[
\lambda^\alpha = n^a \lambda^\perp + e^a_\alpha \lambda^\alpha, \quad \lambda^\perp = \sigma n_\alpha \lambda^\alpha, \quad \lambda^a = e^a_\alpha \lambda^\alpha, \tag{2.59}
\]

The following identities can be immediately derived:

\[
e^a_\alpha e^\alpha = \delta^a_b, \quad e^a_\alpha e^\alpha = \delta^\alpha_b - \sigma n^a n^\beta. \tag{2.60}
\]

The three-dimensional Riemannian connection on the hypersurface can be introduced via the three-dimensional covariant derivative of a 3-vector \( \lambda^\alpha \):

\[
\lambda^{a \beta}_b \equiv (\lambda^c e^\alpha)_{;\beta} e^a_b e^\beta_c, \tag{2.61}
\]

which yields the three-dimensional connection

\[
\gamma^{a \beta}_c = e^a_b e^\beta_c. \tag{2.62}
\]

Here we use the following notation for the four-dimensional covariant derivatives (more exactly, for their projections):

\[
\lambda^a_{;b} \equiv e^\beta_b \lambda^a_{;\beta}, \quad \lambda^a_{;\alpha} \equiv e^\beta_b \lambda^a_{;\alpha}, \quad \lambda^\alpha_{;\beta} \equiv e^a_b \lambda^\alpha_{;\beta}, \quad \lambda^\alpha_{;\alpha} \equiv e^a_b \lambda^\alpha_{;\alpha}, \tag{2.63}
\]

\[
\lambda^\alpha_{;\perp} \equiv n^a \lambda^\alpha_{;\perp}, \quad \lambda^\alpha_{;\parallel} \equiv n^a \lambda^\alpha_{;\parallel}. \tag{2.64}
\]
We define the extrinsic curvature $K_{ab}$ of the hypersurface as

$$K_{ab} = -n_{a;b}. \quad (2.61)$$

Due to the second equation (2.56), $K_{ab} = e^a_{a;b} n_a = K_{ba}$. Comparing (2.61) and (2.59), we see that $K_{ab}$ and $\gamma^{c}_{ab}$ are the projections of the quantity $e^a_{a;b}$ on the normal basis. In view of (2.58), this yields the Gauss-Weingarten formula

$$e^a_{a;b} = \sigma K_{ab} n^a + \gamma^{c}_{ab} e^c_a. \quad (2.62)$$

Let us consider the following vector $N^a$ and its projections on the normal basis:

$$N^a = \frac{de^a(x, t)}{dt}, \quad N = \sigma n^a N_a, \quad N^a = e^a_a N^a. \quad (2.63)$$

We are interested in the deformation of the normal basis when it moves in the four-dimensional spacetime along the vector (2.63). Taking into account that

$$\frac{\partial}{\partial \beta} e^a_{a;\beta}(x, t) \frac{de^a(x, t)}{dt} = \frac{\partial}{\partial \beta} \frac{de^a(x, t)}{dt},$$

we obtain for the quantity $\nabla_N e^a_a \equiv e^a_{a;\beta} N^\beta$ the following expression

$$\nabla_N e^a_a = (N_{|a} + \sigma K_{ab} N^b) n^a + (N_{|a} - NK_{\beta a}) e^a_{\beta}. \quad (2.64)$$

In the derivation we used the decomposition (2.58) for the vector (2.63) and the Gauss-Weingarten formula (2.62).

Taking into consideration (2.56) and (2.57), it is possible to obtain from (2.64) the equation for normal vector deformation

$$\nabla_N n^a = -(\sigma N_{|a} + K_{ab} N^b) e^a_{a}. \quad (2.65)$$

Using (2.64), we obtain the relation between the extrinsic curvature of the hypersurface and the velocity of the change of the metric

$$\dot{g}_{ab} = \nabla_N g_{ab} = -2NK_{ab} + 2N_{(a|b)}. \quad (2.66)$$

Let $\chi^a(x)$ be a vector field in the spacetime with a given hypersurface system (2.55), the projections of which on the normal basis are $\chi^a$ and $\varphi_{\perp}$. Let us analyze how the covariant derivatives of the vector $\varphi^a$ are related to the three-dimensional covariant derivatives of its projections. Calculating the covariant derivative of the projection $\varphi_{\perp} = \sigma n^a \varphi_a$ along the normal vector component (2.63) $n^a N$, we have

$$\delta_{\perp} \varphi_{\perp} \equiv \nabla_N \varphi_{\perp},$$

$$Nn^a \nabla_\beta (\varphi_{\perp}) = -N_{|a} \varphi^a + \sigma \varphi_{\perp;\perp} N,$$

from which

$$N \varphi_{\perp;\perp} = \sigma \delta_{\perp} \varphi_{\perp} + \sigma N_{|a} \varphi^a. \quad (2.67)$$
Doing the same with \( \varphi^a = e^a_\alpha \varphi^\alpha \), we obtain
\[
N \varphi_{a;\perp} = \sigma \delta_\perp \varphi_a + \sigma K_{ab} \varphi_b N - \varphi_{\perp} N_a. \tag{2.68}
\]

The meaning of the operation \( \delta_\perp \) is that it describes the increment caused by the displacement which is purely orthogonal to the hypersurface and is equal to \( n^\alpha N \).

Now let us consider the displacements which are tangent to the hypersurface, \( \delta_{||} \varphi_\perp \) and \( \delta_{||} \varphi_a \). From the point of view of the dependence of the projections of the fixed vector \( \chi^\alpha \) on the normal basis on the embedding of the hypersurface into the four-dimensional spacetime (2.55), \( \varphi_\perp \) and \( \varphi_a \) are functionals of \( e^\alpha(x) \).

Then
\[
\delta_{||} \varphi_a = \int d^3y \left( N^b(y) e_b^\alpha(y) \right) \frac{\delta \varphi_a(x)[e]}{\delta e^\alpha(y)} \tag{2.69}
\]
is nothing but the Lie derivative of the vector field \( \varphi_a(x) \) along the vector \( N^a(x) \):
\[
\delta_{||} \varphi_a (x) = L_N \varphi_a (x). \tag{2.70}
\]
Indeed, the expression (2.69) corresponds to the variation of the form \( \varphi_a(x) \) under the displacement of the three-dimensional coordinate chart \( x = \{ x^a \} \) on the vector \( N = \{ N^a \} \), from which (2.70) follows.

On the other hand, since \( \varphi_a(x) \) is a four-dimensional scalar, the following equation is fulfilled
\[
\delta_{||} \varphi_a = \nabla_N \varphi_a \bigg|_{N^a = e^a_\alpha N^\alpha};
\]
and hence, taking the equation (2.64) into account, with \( N = 0 \), one can derive
\[
\delta_{||} \varphi_a = K_{ab} N^b \varphi_\perp + N_b |A \varphi_b + \varphi_a b N^b. \tag{2.71}
\]

Comparing with the Lie derivative (2.70), \( L_N \varphi_a = \varphi_a |b N^b + \varphi_b N^b |a \), in view of arbitrariness of the vector field \( N^b(x) \), we find
\[
\varphi_a;b = \varphi_a |b - K_{ab} \varphi_\perp. \tag{2.71}
\]

A similar analysis for \( \varphi_\perp \) gives
\[
\varphi_\perp;b = \varphi_\perp |b + \sigma K_{ab} \varphi_\beta. \tag{2.72}
\]

The equations (2.67), (2.68), (2.71), and (2.72) are generalized to the projections of the higher rank tensors in an obvious way, for example:
\[
\varphi_{abc} = \varphi_{ab}c - K_{ac} \varphi_\perp b - K_{bc} \varphi_a \perp,
\]
where
\[
\varphi_a \perp = \sigma n^\alpha e^b_a \varphi^\alpha b, \quad \varphi_\perp b = \sigma n^\alpha e^3_b \varphi^\alpha \beta.
\]

Using the formulas (2.67)-(2.72), one can establish relations between the various projections of the four-dimensional Riemann curvature tensor \( ^4 R^\alpha_{\beta \mu \nu} \) and
the components of the extrinsic curvature of the hypersurface. Projecting the
identity
\[ 2\varphi_{a;[\mu\nu]} = 4R^{\lambda}_{\alpha\mu\nu} \varphi_{\lambda} \]
on the normal basis with an account of the equations (2.67), (2.68), (2.71), and
(2.72), in view of the arbitrariness of the components \( \varphi_\perp \) and \( \varphi_a \) we obtain the
Gauss-Codazzi identities:
\[ 4R_\perp abc = -2\sigma K_{a[b|c]}, \quad 4R_{abcd} = 3R_{abcd} - 2\sigma K_{b[d}K_{c]|a}, \quad (2.73) \]
where \( 3R_{abcd} \) is the Riemann tensor of the extrinsic curvature of the three-
dimensional hypersurface. In a similar way:
\[ N^4R_{a\perp b\perp} = \delta_\perp K_{ab} + NK_{ad}K^d_b - \sigma N_{[ab].} \quad (2.74) \]
In order to rewrite the gravitational field Lagrangian in terms of the quantities
of the (3+1)-decomposition of the spacetime, we will use the following relations:
\[ \delta_\perp (g^{1/2}) = -Ng^{1/2}K, \quad K = K^a_a, \quad g = \text{det } g_{ab}, \]
\[ \delta_\perp (Kg^{1/2}) = \nabla_N (Kg^{1/2}) - \delta_\| (Kg^{1/2}), \]
\[ \delta_\| (Kg^{1/2}) = L_N (Kg^{1/2}) = (Kg^{1/2}N^a|a). \]
With the help of the equations (2.73) and (2.74), we derive
\[ Ng^{1/2}4R = Ng^{1/2} \left\{ 3R - \sigma(K_{ab}K^{ab} - K^2) \right\} + 2\sigma \nabla_N (Kg^{1/2}) \\
- 2(\sigma Kg^{1/2}N^a + g^{1/2}N^a|a). \quad (2.75) \]

2.4. Canonical formalism for the fields
on the curved spacetime

*Canonical formalism of the gravitational field*

Let us clarify the physical meaning of the quantities of the (3+1)-decomposition. The system of space-like hypersurfaces in the physical spacetime is characterized
by the value of the parameter \( \sigma \) equal to \(-1\). However, we will keep \( \sigma \) arbitrary
in all formulas below in order to demonstrate the role of the spacetime signature.

The choice of a family of hypersurfaces (2.55) and of the normal vector (2.56)
must satisfy some requirements. It will be mathematically natural to demand
that the orientation of the normal basis
\[ e_\perp^\alpha = (e_\perp^\alpha, e_a^\alpha), \quad e_\perp^\alpha \equiv n^\alpha, \quad (2.76) \]
coincides with the orientation of the local frames of the four-dimensional coordinate system. Since the metric in the normal basis has the form
\[ e_\perp^\alpha \eta_{\alpha\beta} e_\perp^\beta = \begin{pmatrix} \sigma & 0 \\ 0 & g_{ab} \end{pmatrix}, \]
we find
\[ |4y|^{1/2} \det e^\alpha_{(\mu)} = \pm g^{1/2}, \]
where the upper sign (+) on the right-hand side is fixed by the aforementioned choice of the orientation of the basis,
\[ n^\alpha e^\beta_a e^\nu_b \eta_{\alpha\beta\mu\nu} = \pm g^{1/2}, \quad \varepsilon_{\alpha\beta\mu\nu} = |4y|^{1/2} \eta_{\alpha\beta\mu\nu}, \quad \varepsilon_{abc} = g^{1/2} \eta_{abc}. \]

The parameter \( t \) labels the sequence of the space-like sections of the four-dimensional spacetime in the causal order of increasing of the physical time. This means that the function \( N \) is positive:
\[ N = \sigma n_\alpha \frac{de^\alpha}{dt}, \quad N > 0. \]

This quantity is called a lapse function. The condition (2.78) fixes the direction of the normal vector towards the increase of the time parameter \( t \). In particular, this means that \( N dt \) is the normal interval between the two neighbouring hypersurfaces corresponding to the moments \( t \) and \( t + dt \).

The shift functions \( N^a \) characterize the displacement \( N^a dt \) between the coordinate lines \( x^a \) on the same two neighbouring hypersurfaces, assuming that the coordinate system on the hypersurface corresponding to \( t + dt \), is normally projected on the hypersurface corresponding to the moment \( t \).

An important condition that the curve determined by the equation \( x^a = \text{const} \) is time-like leads to inequality
\[ N^2 - N_a N^a > 0 \quad \text{(2.79)} \]
(it was taken into account that \( \sigma = -1 \) in the physical spacetime). The inequality (2.79) actually means that the coordinate \( t \) is time-like and there may be observers at rest on the three-dimensional coordinate chart on the hypersurface.

Eq. (2.77) yields
\[ |4y|^{1/2} d^4x = dt d^3x \sqrt{|g|}, \quad \text{(2.80)} \]
and hence, in view of (2.75), we obtain
\[
\int_V d^4x |4y|^{1/2} 4R = \int_{t_0}^{t_1} dt d^3x \sqrt{|g|} \left\{ 3R - \sigma (K_{ab} K^{ab} - K^2) \right\}
+ 2\sigma \int d^3x \sqrt{|g|} \left. K \right|_{e_0}^{e_1} - 2 \int_{t_0}^{t_1} dt d\sigma_a \left( \sigma K N^a + N^a \right). \quad \text{(2.81)}
\]

Here \( V \) is a four-dimensional spacetime domain bounded by the “cylindrical” surface, the top and the bottom sides of which are the initial and the final hypersurfaces \( e_0 = e(t_0) \) and \( e_1 = e(t_1) \), and the lateral surface \( \partial V \) is topologically a direct product of the two-dimensional sphere \( \partial e(t) \) of radius \( r \) and the time axis in the interval \( t_0 \leq t \leq t_1 \), \( \partial V = \partial e(t) \times [t_0, t_1] \), see Fig. 2.1.
The last term in (2.81) is a surface integral over the “cylindrical” spatial asymptotics bounding the space-time “tube” in the interval $t_0 \leq t \leq t_1$, and $d\sigma_a = n_a d\sigma$ is an element of two-dimensional surface $\partial e(t)$ with a unit normal $n_a$. When deriving the form of the second term in (2.81), we took into account that

$$\nabla_N (K g^{1/2}) = \frac{d}{dt} (K g^{1/2}).$$

The action of the gravitational field [14] that leads as a result of the variational procedure to Einstein’s equations of the second order, reads as follows:

$$S = \int_V d^4x |g|^{1/2} 4R - \int_{\partial V} d\Sigma \sigma (\tilde{K}). \quad (2.82)$$

The surface integral over the boundary $\partial V$ is determined by the extrinsic curvature $\tilde{K}$ given by the formula (2.61), where $n_\alpha$ is an outer normal to the boundary and $\sigma$ defines the scalar square of the normal vector (2.57). Its role is to compensate in the volume integral of action (2.82) the terms containing the derivatives of the metric which are normal to the boundary. This allows to correctly formulate the boundary value problem for Einstein’s equations as conditions of the extremum of the action (2.82) (see Chapter 1).

---

\[2\] Hawking considered the gravitational field in the Euclidean spacetime where $\sigma = +1$. In the physical spacetime, his formula is modified, taking into account the metric signature on the boundary $\partial V$. 

---

Figure 2.1: Four-dimensional spacetime domain $V$ bounded by “cylindrical” surface $\partial V = \partial \bar{V} \cup e_0 \cup e_1$, $\partial V = \partial e(t) \times [t_0, t_1]$. 

---

36 CANONICAL FORMALISM OF GRAVITY
Substituting (2.81) into (2.82), we obtain the action of the gravitational field in terms of the \((3 + 1)\)-decomposition:

\[
S = \int_{t_0}^{t_1} dt \left\{ \int d^3 x \left( Ng^{1/2} \mathcal{L} \right) - H_0 \right\},
\]

\[
\mathcal{L} = 3R - \sigma (K_{ab}K^{ab} - K^2).
\]

Here the quantity \(H_0\) is determined by the surface integral of the extrinsic curvature of the boundary \(\tilde{K}\) over this two-dimensional boundary of the three-dimensional space,

\[
H_0 = 2 \oint d\tilde{\sigma} \frac{g^{1/2}}{2} \sigma \tilde{K} + 2 \oint d\sigma_a \left( N^{\|a} + \sigma K^a \right),
\]

and the exact form of the measure \(d\tilde{\sigma}\) is obtained if we represent the surface element \(d\Sigma\) of the “cylindrical” lateral boundary of four-dimensional volume as

\[
d\Sigma = dt d\tilde{\sigma},
\]

where \(d\tilde{\sigma}\) is some element of the measure on \(\partial e(t)\) which is in general different from \(d\sigma\).

Let us proceed to the construction of the canonical formalism of the theory described by the action (2.82). Taking into account that according to (2.66)

\[
K_{ab} = \frac{1}{2N} \left( 2N_{(a|b)} - g_{ab} \right),
\]

the action (2.82) is a functional of the set of ten variables \((g_{ab}, N^a, N)\) which replaces the original set of ten metric components \(g_{\mu\nu}\). The action (2.82) does not contain time derivatives of the lapse and shift functions, therefore, \(N\) and \(N^a\) are the Lagrange multipliers from the point of view of the theory of systems with singular Lagrangians. In order to verify this, we will bring the action (2.82) to the form (2.27).

For this purpose, we introduce the canonical momenta \(\pi^{ab}\) conjugated to the 3-metric components \(g_{ab}\):

\[
\pi^{ab} = \frac{\partial (Ng^{1/2} \mathcal{L})}{\partial \dot{g}_{ab}} = \sigma G^{abcd} K_{cd},
\]

\[
G^{abcd} = \frac{1}{2} g^{1/2} \left( 2g^{ab}g^{cd} - g^{ac}g^{bd} - g^{ad}g^{bc} \right).
\]

Here \(G^{abcd}\) is the so-called covariant DeWitt supermetric [53]. Its contravariant components are given by the expression

\[
G_{abcd} = \frac{1}{2} g^{1/2} \left( g_{ac}g_{bd} + g_{ad}g_{bc} - g_{ab}g_{cd} \right),
\]

\[(2.89a)\]
The gravitational field Hamiltonian has the following form:

$$H = \int d^3x \left( \pi^{ab}(x) \dot{g}_{ab}(x) - Ng^{1/2}L \right).$$  \hspace{1cm} (2.90)

Since from (2.89) the extrinsic curvature can be expressed in terms of the momenta

$$K_{ab} = \sigma G_{abcd} \pi^{cd},$$  \hspace{1cm} (2.91)

the Hamiltonian (2.90) is recast into

$$H = \int dx^3 N^{(\mu)} \mathcal{H}_\mu + H_0,$$  \hspace{1cm} (2.92)

where we introduced the unified notation for the lapse and shift functions

$$N^{(\mu)} = (N^\perp, N^a), \quad N^\perp = N,$$  \hspace{1cm} (2.93)

and the quantities $\mathcal{H}_\mu = (\mathcal{H}_\perp, \mathcal{H}_a)$ are given by the following expressions:

$$\mathcal{H}_\perp = -\sigma G_{abcd} \pi^{ab} \pi^{cd} - g^{1/2} 3R,$$  \hspace{1cm} (2.94)

$$\mathcal{H}_a = -2g_{ac} \pi^{cd} \big|_d.$$  \hspace{1cm} (2.95)

Therefore, in terms of the canonical variables the action (2.82) reads

$$S_{HE}[g_{ab}, \pi^{ab}, N^{(\mu)}] = \int_{t_0}^{t_1} dt \left\{ \int d^3x \pi^{ab} \dot{g}_{ab} - H_0 - \int d^3x N^{(\mu)} \mathcal{H}_\mu \right\},$$  \hspace{1cm} (2.96)

which makes it obvious that the functions $N^{(\mu)}$ are the Lagrange multipliers.

The stationary condition of the action (2.96) with respect to variations of $N^{(\mu)}$ leads to the equations

$$\mathcal{H}_\perp = 0,$$  \hspace{1cm} (2.97)

$$\mathcal{H}_a = 0,$$  \hspace{1cm} (2.98)

which are the constraints since they do not contain the time derivatives of the phase variables.

By direct evaluation of the Poisson brackets, one can show that these constraints satisfy the involution relations (2.29):

$$\begin{cases} 
\{ \mathcal{H}_\perp (x), \mathcal{H}_\perp (x') \} = -\sigma (\mathcal{H}^\alpha (x) \partial_\alpha \delta(x, x') - \mathcal{H}^\alpha (x') \partial_\alpha \delta(x', x)), \\
\{ \mathcal{H}_a (x), \mathcal{H}_\perp (x') \} = \mathcal{H}_\perp (x) \partial_0 \delta(x, x'), \\
\{ \mathcal{H}_a (x), \mathcal{H}_b (x') \} = \mathcal{H}_b (x) \partial_0 \delta(x, x') - \mathcal{H}_a (x') \partial_0 \delta(x', x).
\end{cases}$$  \hspace{1cm} (2.99)

Accordingly, we found that (2.97), (2.98) are the first-kind constraints.
The canonical equations for the phase variables of the gravitational field are obtained by the variation of the action (2.96) with respect to \( g_{ab} \) and \( \pi^{ab} \). The term \( H_0 \) in the total Hamiltonian of the system (2.92) is a surface integral, which does not explicitly contribute to the dynamical equations under the variation. As a result, the equations of motion in terms of the Poisson brackets take the following form:

\[
\dot{g}_{ab}(x) = \int d^3x' \left\{ g_{ab}(x), H_{(\mu)}(x') \right\} N^{(\mu)}(x'), \\
\dot{\pi}_{ab}(x) = \int d^3x' \left\{ \pi^{ab}(x), H_{(\mu)}(x') \right\} N^{(\mu)}(x').
\]

(2.100)

(2.101)

Computing the Poisson bracket in (2.100), we obtain the equation (2.66), where the extrinsic curvature \( K_{ab} \) is expressed in terms of the momenta \( \pi^{ab} \) via (2.91). Similarly, one can show that the equation (2.101) is equivalent to the following set of Einstein’s equations projected on the normal basis: \( NG_{ab} = 0 \).

The constraints (2.97), (2.98) are also the components of Einstein’s equations

\[
\mathcal{H} = 2\sigma g^{1/2} G_{\perp\perp}, \\
\mathcal{H}_a = 2g^{1/2} G_{a\perp}.
\]

(2.102)

(2.103)

Therefore, the system of equations (2.100), (2.101) is completely equivalent to Einstein’s equations \( G_{\mu\nu} = 0 \).

One can read from (2.99) the structural functions of the involution relations (2.29):

\[
U_{\perp x, \perp x'}^y = \delta(x, y) \partial_a \delta(x, x'), \\
U_{\perp x, \perp x'}^{\perp y} = \partial_a \delta(x, x') \delta_b^a \delta(y, x) \delta_a^c \delta(y, x'), \\
U_{\perp x, \perp x'}^{\perp \perp} = 0, \\
U_{\perp x, \perp x'}^{\perp y} = -\sigma(g^{ab}(x) \partial_b \delta(x, x') \delta(y, x) - g^{ab}(x') \partial_b \delta(x', x) \delta(y, x')).
\]

(2.104)

Since the phase variables in the interior points of the spacetime domain \( V \) commute in the sense of the Poisson brackets with the quantities on the boundary \( \partial V \), the structural functions \( Y_{(\nu)}^{(\mu)} \) vanish.

Using the results obtained in (2.30), (2.31), one can see that the canonical action (2.96) is invariant with respect to transformations:

\[
\delta^F \pi^{ab} = \int d^3x' \left\{ \pi^{ab}, H_{(\mu)}(x') \right\} F^{(\mu)}(x'),
\]

(2.105)

\[
\delta^F g_{ab} = \int d^3x' \left\{ g_{ab}, H_{(\mu)}(x') \right\} F^{(\mu)}(x'),
\]

(2.106)

\[
\delta^F N = \mathcal{F}_\perp - N^a \mathcal{F}_{\perp\perp} + N_{\perp\perp} \mathcal{F}^a,
\]

(2.107)

\[
\delta^F N^a = \mathcal{F}^a - (\mathcal{F}_\perp\perp N^b - N^a \mathcal{F}_{\perp\perp}) + \sigma(N \mathcal{F}_{\perp\perp\perp} - N^a \mathcal{F}_{\perp\perp}),
\]

(2.108)
with the infinitesimal parameters \( \mathcal{F}^{(\mu)} = (\mathcal{F}^\perp, \mathcal{F}^a) \), \( \mathcal{F}^\perp \equiv \mathcal{F}_\perp \). Let us show that the transformations (2.106)-(2.108), evaluated in the phase space for the extremal values of the momenta (2.88) [i.e., for the momenta which are the solution of the equations (2.100) in terms of \( g_{ab} \) and \( \dot{g}_{ab} \)], are equivalent to the transformation of the general coordinate group (2.52) with the parameters

\[
\delta \mathcal{F} g_{ab} = e_\alpha^a e_\beta^b \Delta \mathcal{F} g_{\alpha\beta} = 2f_{a;b},
\]

and hence using (2.71), (2.91) in combination with the fact that \( f_\perp = \mathcal{F}_\perp \) and \( f_a = g_{ab} \mathcal{F}^b \), we derive (2.106). Similarly, we have

\[
\delta \mathcal{F} N = \sigma \Delta \mathcal{F} n_\alpha \dot{e}_\alpha.
\]

The variation of the normal vector, caused by the variation of the four-dimensional metric, is obtained as a result of the variation of the equations (2.56), (2.57):

\[
\Delta \mathcal{F} n_\alpha = -\frac{1}{2} \sigma n_\alpha (n_\mu n_\nu \Delta \mathcal{F} g^{\mu\nu}),
\]

from which

\[
\Delta \mathcal{F} n_\alpha = \sigma n_\alpha f_{\perp;\perp}.
\]

Substituting (2.112) into (2.111) and taking into account (2.67), we obtain (2.107), since \( \delta \mathcal{F} \mathcal{F}_\perp = \dot{\mathcal{F}}_\perp - \mathcal{F}_\perp|_a a^a \). In the same way, the last relation (2.108) is derived.

Thus, we have demonstrated that the transformations (2.106)-(2.108) on the surface of the extremal momenta realize the general coordinate group (2.51) in the configuration space. Since this transformation of coordinates

\[
x^\alpha \rightarrow x^\alpha + f^\alpha(x)
\]

does not change, in view of (2.110), the form of the functions \( e^\alpha(x,t) \), they can be treated as the simultaneous transformation of the coordinate system (2.113) and the change of the hypersurface system

\[
e^\alpha(x,t) \rightarrow e^\alpha(x,t) - f^\alpha(x)|_{x = e(x,t)},
\]

so that the total change of the embedding functions \( e^\alpha(x,t) \) vanishes.

Consequently, the transformations (2.105)-(2.108) realize the group of the hypersurface deformations with the parameters \( -f^\alpha = -e^\alpha_{(\mu)} \mathcal{F}^{(\mu)} \). Eq. (2.109)
shows that $H(x) = H_\perp(x)$ is the generator of the hypersurface deformations at a point $x$ in the direction which is normal to the hypersurface. This generator is called a superhamiltonian. In a similar way, the constraints $H_a(x)$, which are called supermomentum components, are the generators of the tangential deformations or the transformations of the three-dimensional coordinate chart on the hypersurface. Therefore, in particular,

$$\int d^3x' \left\{ g_{ab}(x), H_c(x') \right\} N^c(x') = L_N g_{ab}(x).$$

The Lie derivative of the momentum field is also given by a similar expression.

Let us note that, in contrast to the general coordinate transformations that constitute the group, the transformations realizing them in the phase space do not constitute a group since the structural functions (2.104) are not constant and depend on the phase space variables, i.e., these transformations constitute the canonical “pseudogroup”. The difference of its structure from the group structure is that the coefficients of the transition from the parameters $f^\mu$ to $F^{(\mu)}$ in (2.109) are not constant and depend on the metric. One can show that the structural functions of the involution relations follow from the structural constants of the general coordinate group, and vice versa, the involution relations (2.99) lead to the commutation of the general coordinate transformations using the Lie brackets of the vector fields.

**Geometricdynamics with the matter sources**

The system of the gravitational and the matter fields is described by the total action

$$S = S_{HE} + S_m.$$  \hspace{1cm} (2.114)

As in Chapter 1, we assume that the interaction of the matter field $\Psi$ with the gravitation is minimal. Therefore, the action of an unspecified matter field $\Psi$ has the following form:

$$S_m = \int d^4x \sqrt{|g|}^{1/2} L_m(\Psi, \Psi_\mu).$$  \hspace{1cm} (2.115)

The construction of the canonical formalism for the system (2.114) starts with recasting of the matter field action (2.115) into the $(3 + 1)$-decomposed form. As before, we need to calculate the projections of the tensor field $\Psi$ on the normal basis. The complete set of these projections is denoted $\varphi$: these will be the phase coordinates in the theory. The momenta $p$ conjugated to them are determined in a standard way,

$$p = N g^{1/2} \frac{\partial L_m}{\partial \dot{\varphi}},$$

where we took into account the rule (2.80) of transition to the new spacetime coordinates. This procedure recasts the action (2.115) in terms of the canonical
variables into a form which is similar to (2.83):
\[ S_m [\varphi, p] = \int dt \int d^3x \left\{ p\dot{\varphi} - N(\mu) \frac{\partial}{\partial \mu} H^m_{(\mu)} \right\}, \] (2.116)
where
\[ \frac{\partial}{\partial \mu} H^m_{(\mu)} = \frac{\partial}{\partial \mu} H^m_{(\mu)} (\varphi, p, g_{ab}, K) \] (2.117)
are the superhamiltonian and the supermomenta of the matter field on the curved spacetime, which are the functions of the phase variables of the three-dimensional metric and extrinsic curvature \( K_{ab} \). It is important that \( \frac{\partial}{\partial \mu} H^m_{(\mu)} \) does not depend on the lapse and shift functions, so that the field Hamiltonian in (2.116) turns out to be a linear function of the latter.

The construction of momenta \( p^{ab} \) conjugated to the 3-metric of the gravitational field for the system (2.114) leads to the result
\[ p^{ab} = \pi^{ab} + P^{ab}, \] (2.118)
\[ P^{ab} = \frac{\partial}{\partial \dot{g}_{ab}} \left( N g^{1/2} L_m \right), \] (2.119)
where the purely gravitational momentum \( \pi^{ab} \) is determined by (2.88). The quantity \( P^{ab} \) (2.119) is an additional contribution to the gravitational field momentum, due to the presence of time derivatives of the metric in the matter field Lagrangian. It is clear that the derivatives of the metric enter \( L_m \) in terms of the Riemannian connection in the covariant derivatives \( \Psi_{\mu} \). Since the variation of the Christoffel symbols with respect to the metric is expressed in terms of the covariant derivatives of the metric variations \( \delta g_{\mu\nu} \), one can write the following variational identity:
\[ \frac{\partial L_m}{\partial \Psi_{\mu}} \delta g_{(\Psi_{\mu})} = \frac{1}{2} P^{\mu\nu\sigma} (\delta g_{\mu\nu})_{;\sigma}, \] (2.120)
which also serves as a definition for the tensor \( P^{\mu\nu\sigma} \) (see Chapter 1, Sec. 1.3.). Substituting (2.120) into (2.119) and using (2.68), we arrive at
\[ P^{ab} = \frac{1}{2} g^{1/2} P^{ab \perp}. \] (2.121)

For the theories of matter that does not interact with the derivatives of the metric, i.e., when \( \Psi_{\mu} = \partial_{\mu} \Psi \) in the action (2.115) (which is the case, for example, for the scalar or the electromagnetic fields) one finds \( P^{\mu\nu\sigma} = 0 \), and the momentum of the gravitational field \( p^{ab} \) identically coincides with the purely gravitational momentum
\[ p^{ab} = \pi^{ab}. \] (2.122)

It is convenient to rewrite the superhamiltonian of the matter \( H^m \) by isolating the term of a special structure as follows:
\[ H^m_{(\mu)} = H^m + 2K_{ab} P^{ab}. \] (2.123)
Now let us plug into (2.114) the expressions (2.96), (2.116) and (2.123), and in the latter equation we make use of (2.87). Integrating the term \(-2N(a|b)P^{ab}g^{1/2}\) by parts and dropping an inessential surface integral, we obtain the canonical action of the gravitating system:

\[
S \left[ g_{ab}, p^{ab}, \varphi, p, N(\mu) \right] = \int_{t_0}^{t_1} dt \left\{ \int d^3x \left( p^{ab} \dot{g}_{ab} + p \dot{\varphi} - N^{(\mu)} H(\mu) \right) - H_0 \right\}.
\]

Here the superhamiltonian \(H_{\perp}\) and the supermomenta \(H_a\) of the total system are given by:

\[
\begin{align*}
H_{\perp} &= \mathcal{H}^{m} + \mathcal{H} \big|_{\pi^{ab} = p^{ab} - P^{ab}}, \\
H_a &= \mathcal{H}^{m} - 2g_{ab}p^{bc}_{\mid c}. 
\end{align*}
\]

The extrinsic curvature \(K_{ab}\) should be expressed in these formulas in terms of the phase variables from the solution of the equation (2.91), taking the following form:

\[
K_{ab} = \sigma G_{abcd} \left[ p^{cd} - P^{cd} (K_{ef}) \right].
\]

One can show that the superhamiltonian (2.125) and the supermomenta (2.126) satisfy the involution relations

\[
\{ H(\mu), H(\nu) \} = U(\alpha) (\mu \mid \nu) H(\alpha),
\]

where the Poisson brackets are defined for the complete set of the phase variables \(g_{ab}, \pi^{ab}, \varphi\) and \(p\), and with the same structural functions as for the pure gravitation\(^3\). Consequently, \(H(\mu)\) are the first-kind constraints and they are generators of the pseudogroup of the space-like hypersurface deformations, considered in the previous section, that now acts in the total space of the variables \(g_{ab}, p^{ab}, \varphi, p, \Phi = (g_{ab}, p^{ab}, \varphi, p)\),

\[
\delta F \Phi = \int d^3x' \left\{ \Phi, H(\mu) (x') \right\} F^{(\mu)}(x').
\]

Therefore, the gravitating system described by the action (2.124) is degenerate, due to its invariance with respect to the transformations (2.128) realizing the general coordinate transformations in the configuration space.

The first-kind constraints of the theory

\[
H(\mu) \left( g_{ab}, p^{ab}, \varphi, p \right) = 0
\]

\(^3\)This conclusion is based on the independence of the total action (2.124) on the choice of a slicing (2.55) in the interior domain of space-time volume.
are equivalent to the four Einstein field equations with the tensor sources

\[ H_{\perp} = 2\sigma g^{1/2} \left( G_{\perp\perp} - \frac{1}{2} T_{\perp\perp} \right), \]
\[ H_a = 2g^{1/2} \left( G_{\perp a} - \frac{1}{2} H_{\perp a} \right), \]

where \( H_{\perp\perp} \) and \( H_{\perp a} \) are the projections of the energy-momentum tensor \( H_{\mu\nu} \) on normal basis vectors. The equations which are obtained by the variation of (2.124) with respect to \( g_{ab} \) and \( \pi^{ab} \) are equivalent to the remaining six Einstein’s equations \( G_{ab} - \frac{1}{2} T_{ab} = 0. \)

Let us note that in a particular case of the interaction without derivatives, when \( P^{ab} = 0 \) and the relation (2.122) is valid, the structure of the formulas (2.124)-(2.126) is essentially simplified. The superhamiltonian \( H_\perp \) of the total system is decomposed into a sum of the purely gravitational superhamiltonian plus the superhamiltonian of the matter in an external gravitational field.

The non-dynamical character of the torsion field in ECT leads to an insignificant modification of the superhamiltonian \( H_\perp \) of the total system, which does not destroy the structure of the general action (2.124) and the involution relations (2.127). Therefore, the canonical formalism of the gravitating fields in the Riemann-Cartan space has the same structure as in the Riemann space.
3

Dynamics of gravity theories of Hilbert-Einstein type

3.1. The problem of the “frozen” formalism

The dynamical contents of the field theory is determined by the sector of its physical degrees of freedom cleared of the ambiguity caused by the invariance under the action of the local group.

The gravity theory of the Hilbert-Einstein type in the space of a general affine connection can be effectively formulated in the Riemannian space. The torsion fields in the theories of this type are non-dynamical and they are algebraically excluded in terms of the components of the spin tensor of matter. Hence, they cannot affect the dynamical content of the theory. In this context, without the loss of generality, we will consider the dynamics of Einstein’s GR with arbitrary matter sources.

The canonical action of the gravitating system (2.124) is a particular case of the general class of degenerate systems described by the action (2.27). The lapse and shift functions $N^{(\mu)}$ are the Lagrange multipliers. Variation of (2.124) with respect to the phase variables $\Phi$ and the Lagrange multipliers yields the equations

$$\dot{\Phi} = \delta^N \Phi, \quad (3.1)$$
$$H_{(\mu)}(\Phi) = 0, \quad (3.2)$$

where $\delta^N \Phi$ is the canonical transformation (2.128) defined by the parameters $F^{(\mu)} = N^{(\mu)}$. The Lagrange multipliers are not determined from the equations of motion and remain arbitrary.
As a result, we find ourselves in a strange situation: the right-hand side of the canonical equations (3.1) turns out to be a pure canonical transformation (2.30), (2.30a) that leaves the action invariant, i.e., the dynamics of the phase variables reduces to canonical transformations with arbitrary parameters \(N^{(\mu)}\) realizing the local invariance group in the phase space. It appears that there are no dynamical degrees of freedom in the gravitating system, since its evolution reduces to transformations from the invariance group of the action.

The reason is that in the total Hamiltonian of the system

\[
H = H_0 + \int d^3x N^{(\mu)}H_{(\mu)}
\]

the quantity \(H_0\) does not explicitly contribute to the field equations (3.1), (3.2), since it is the surface integral (2.85). In fact, a contribution does exist, but it is implicit in its nature. The Hamiltonian \(H_0\) ensures the correctness of the variational problem for the action, leading to the proper boundary value problem for the second-order field equations [14].

Such a situation producing an illusion of the absence of dynamics in the gravity theory is called a problem of the “frozen” formalism. In a more explicit form, this problem arises in cosmological closed models, where the surface integral \(H_0\) is absent, \(H_0 \equiv 0\), and the complete Hamiltonian (3.3) vanishes on the surface of the constraints (3.2) in the phase space. The problem of the “frozen” formalism arises more seriously when we try to construct the quantum gravity theory, since in a naive approach to the quantization the vanishing of the Hamiltonian (3.3) of the system means the absence of evolution of the vector of the quantum field state

\[
|\Psi(t)\rangle = \exp (iHt)|\Psi(0)\rangle = |\Psi(0)\rangle.
\]

For the development of the canonical formalism of degenerate systems, analyzed in the previous chapter for the general theory, one needs to address the arising problem of the “frozen” formalism.

### 3.2. Arnowitt-Deser-Misner procedure of selection of physical degrees of freedom

Let us assume that the number of pairs of the phase variables of the matter field \(\varphi\) is equal to \(N\). Thus, the general dimension of the phase space of a gravitating system is equal to \(2n = 12 + 2N\), since the total number of the gravitational field momenta \(p^{ab}\) plus the components of the 3-metric \(g_{ab}\) is equal to 12. In accordance with the general procedure of the previous chapter, the number of the dynamical degrees of a degenerate system with \(m = 4\) constraints of the first kind (3.2) is equal to \(n - m = 2 + N\). Therefore, the pure gravitational field accounts for the two dynamical degrees of freedom. Let us find these variables.

Consider the canonical transformation from the initial set of the phase variables \(\Phi = (g_{ab}, p^{ab}, \varphi, p)\) to the new set of \((12 + 2N)\) variables:
3.2 Arnowitt-Deser-Misner procedure

\[(g_{ab}, p^b, \varphi, p) \rightarrow (X^\mu, \Pi_\mu; g^A, p_A),\]  

(3.4)

where \((X^\mu, \Pi_\mu)\) are the four pairs of the canonically conjugated phase variables, and \((g^A, p_A)\) are the remaining \((N+2)\) pairs \((A = 1, 2, \ldots, N+2)\).

This transformation may be absolutely arbitrary. The only restriction imposed on it is as follows. Expressing the old phase variables in terms of the new ones

\[\Phi = \Phi [X^\mu, \Pi_\mu; g^A, p_A],\]  

(3.5)

we obtain the geometrodynamical constraints \((2.125), (2.126)\) in terms of the new phase variables:

\[H_{(\mu)} = H_{(\mu)} (x) [X^\alpha, \Pi_\alpha; g^A, p_A].\]  

(3.6)

One can define the matrix

\[J_{(\nu)}(x, y) \equiv \frac{\delta H_{(\nu)}(y)}{\delta \Pi_\mu(x)} = \{X^\mu(x), H_{(\nu)}(y)\}.\]  

(3.7)

The transformation \((3.4)\) must satisfy the condition of invertibility of the functional matrix \((3.7)\), \(\det (J_{(\nu)}(x, y)) \neq 0\), so that the inverse matrix exists and reads

\[J_{(\nu)}^{-1}(y, z), \int d^3y J_{(\nu)}(x, y)J_{(\nu)}^{-1}(y, z) = \delta^{(\nu)}_{(\lambda)} \delta(x-z),\]  

and then one can solve the constraint equations \(H_{(\mu)} (x) [X^\alpha, \Pi_\alpha; g^A, p_A] = 0\) with respect to \(\Pi_\alpha\),

\[\Pi_\mu(x) = -P_\mu(x) [X^\alpha; g^A, p_A].\]  

(3.8)

Since the transformation \((3.4)\) is canonical, the action \((2.124)\) in terms of new variables takes the following form:

\[S = \int_{t_0}^{t_1} dt \int d^3x \left\{ p_A g^A + \Pi_\mu \dot{X}^\mu - H_0 - N^{(\nu)}H_{(\mu)}(x)[X, II; g^A, p_A] \right\}, \]  

(3.9)

where the surface integral \(H_0\) must also be written using \((3.5)\) in terms of the new phase variables.

The equations obtained from the variational principle for the action \((3.9)\) are equivalent to the complete set of the canonical equations. Let us write the equation from the variation with respect to \(\Pi_\mu\):

\[\dot{X}^\mu(x) = \int d^3y N^{(\nu)}(y)J_{(\nu)}^{(\mu)}(x, y).\]  

(3.10)

1Since nothing is assumed about the locality of the transformation \((3.4)\), \(\Phi\) are the functionals of the new variables as the functions of the three-dimensional coordinates \(x\).
In accordance with the general procedure of Chapter 2 we introduce the gauge conditions:
\[ \chi^\mu(x) = X^\mu(x) - f^\mu(t, x) = 0, \] (3.11)
where \( f^\mu(t, x) \) are some functions of the coordinates and the time parameter \( t \).

Then the parametrization (2.38) of the complete set of the phase variables in terms of the physical degrees of freedom \((q^*, p^*) = (g^A, p_A)\) via
\[ X^\mu = f^\mu(t, x), \quad \Pi_\mu = -P_\mu[f^\alpha; g^A, p_A], \] (3.12)
guarantees that the complete set of the constraints and gauge conditions (2.39) or (3.11) is satisfied. The action (3.9) in terms of the physical degrees of freedom is obtained after substituting (3.12) into (3.9):
\[ S[g^A, p_A] = \int_{t_0}^{t_1} dt \left\{ \int d^3x p_A \dot{g}^A - H_{\text{phys}}(t)[g^A, p_A] \right\}, \] (3.13)
where the total physical Hamiltonian is equal to
\[ H_{\text{phys}}(t)[g^A, p_A] = H_0 [g^A, p_A] + \int d^3x P_\mu(x) [f^\alpha; g^A, p_A] \dot{f}^\mu(t, x). \] (3.14)

When varying the action (3.13) with respect to \((g^A, p_A)\), we obtain the closed system of equations for the physical degrees of freedom, however, the lapse and shift functions, which are the Lagrange multipliers, remain undetermined. Without the knowledge of them, the dynamics of the gravitating system is not fully determined. If we require the preservation of the gauge conditions (3.11) in time, we obtain the equation (3.10) for \( N(\nu)(x) \), where the variables \( \dot{X}^\mu(x) \) are replaced with \( \dot{f}^\mu(t, x) \), and hence
\[ N(\nu)(x) = \int d^3y J^{-1(\nu)}(x, y) f^\nu(t, y). \] (3.15)

It is easy to see that these equations are equivalent to the equations (2.35) and with \( H_0 = 0 \) nonvanishing values of \( N(\nu)(x) \) are entirely due to explicit dependence of the gauge conditions (3.11) on time.

Therefore, the equations (3.13)-(3.15) solve the problem of selection of the dynamically independent degrees of freedom, and one can construct a nontrivial physical Hamiltonian. Let us consider the arbitrariness of the choice of the functions \( f^\mu(t, x) \) in the additional conditions (3.11).

The lapse and shift functions must satisfy the following conditions:
\[ N > 0, \quad N^2 - N^a N_a > 0. \] (3.16)

The first condition determines the positive direction of the time counted by the parameter \( t \), and the second condition means that the coordinate curve \( t \) is time-like, i.e., it lies within the light cone. Strictly speaking, the second condition is not mandatory, however, one should impose this condition if one
requires that the coordinate system \((t, x)\) could be realized using the physical bodies.

In addition to the arbitrariness related with functions \(f^\mu(t, x)\), there is a large arbitrariness in the choice of the new set of phase variables \((3.4)\) connected with an old set by a canonical transformation.

Due to the complicated differential character of the gravitational constraints, there are no spatially-local canonical transformations \((3.4)\) leading to the local function for the Hamiltonian \((3.14)\) density [50]. It means that in general case the ADM method produces a non-local action for the physical variables and the non-local expressions for the lapse and shift functions \((3.15)\). As a rule, the functional matrix \((3.7)\) is a differential operator, and hence \(J^{-1(\nu\lambda)}\) is an integral operation requiring to fix the boundary conditions at the spatial boundary.

### 3.3. Asymptotically flat and closed worlds

It was established above that the construction of a nontrivial physical Hamiltonian is performed simultaneously with the selection of the dynamically independent variables of the gravitating system. In addition, the lapse and shift functions are fixed uniquely, i.e., the physical time and coordinates are defined. Without referring to the ADM procedure, it is impossible to construct a regular Hamiltonian, since the naive Hamiltonian \((3.3)\) cannot generate the dynamics of the theory and produces the problem of the “frozen” formalism. That is quite clear, since the basic notions of any theory are not the energy and momentum, but the time and space where the dynamics of the physical degrees of freedom takes place. The energy and momentum are in a certain sense conjugated quantities to the time and coordinates, they are the generators of translations in the spacetime.

To construct the dynamics of the gravitating system, it is necessary to distinguish two important subclasses of problems with essentially different physical properties: the asymptotically flat insular worlds and the closed cosmological systems.

#### The energy of asymptotically flat worlds

Asymptotically flat worlds are characterized by the presence of a region in them, where the timelike and spacelike Killing vector fields\(^2\) exist. They cannot be determined globally, since the space on the whole is essentially inhomogeneous, however, they exist asymptotically.

An open asymptotically flat space should be considered as the region of the spacetime with the boundary, all points of which are directed to the spatial infinity. In the asymptotically flat space, the “lateral cylindrical” boundary

\(^2\)The worlds, where only timelike asymptotic Killing vector fields exist, can be analyzed similarly to the asymptotically flat worlds.
\( \partial V \) (see Fig. 2.1) is obtained in the limit of \( r \to \infty \) for the radius of the two-dimensional sphere representing \( \partial e(t) \). Besides, the vertical line (parallel to the time axis) generating the cylindrical boundary \( \partial V \) is set orthogonal to the hypersurfaces \( e(t) \), that is on the whole, the boundary \( \partial V \) is orthogonal to \( e(t) \) for any value of \( t \). Therefore, the normal \( n_a \) to \( \partial V \) lies in the plane of the hypersurface \( e(t) \) and is fixed by its projections \( n_a \) on \( e(t) \). Moreover, the vector \( n_a \) coincides with the normal to the two-dimensional sphere \( \partial e(t) \) in the three-dimensional space \( e(t) \). The element of the measure \( d\tilde{\sigma} \) turns out to be equal to

\[
d\tilde{\sigma} = Nd\sigma. \quad (3.17)
\]

Computing the extrinsic curvature of the boundary \( \partial V \), one finds the expression for its trace

\[
K = -\frac{1}{N} (n^a N)_a. \quad (3.18)
\]

Substituting (3.17) and (3.18) into (2.85) and taking into account that the lapse function of the flat space is equal to one, we have

\[
H_0 = \int_{\partial e(t)} d\sigma \left\{ -2 \left[ n_a \right] N + 2n_a (p^{ab} N_b - p N^a) \right\}, \quad (3.19)
\]

where \( \left[ n_a \right] \) is the difference of the values of \( n_a \) for the curved and the flat spaces.

The boundary conditions for the metric in an asymptotically flat space are fixed as

\[
g_{ab} = \delta_{ab} + \frac{M}{8\pi r} n_a n_b + \mathcal{O}(r^{-2}), \quad n^a = \frac{x^a}{r},
\]

\[
N = 1 - \frac{M}{16\pi r} + \mathcal{O}(r^{-2}),
\]

\[
N_a = \mathcal{O}(r^{-2}),
\]

where \( M \) is the total mass of an isolated distribution of matter.

Since \( ds = r^2 d\Omega^2 + \mathcal{O}(r^{-2}) \), where \( d\Omega^2 \) is the element of a solid angle, in the limit of \( r \to \infty \) only the terms of order not smaller than \( \mathcal{O}(r^{-2}) \) contribute to \( H_0 \). Therefore, when calculating \( \left[ n^a \right] \), it is sufficient to restrict oneself just to the first term of the power series of the metric expansion in its deviation from the flat metric \( h_{ab} = g_{ab} - \delta_{ab} \), which yields

\[
-2 \left[ n^a \right] = n^a \left( h_{ab} - h_{b|a} \right) + \mathcal{O}(r^{-3}).
\]

The momenta of the gravitational field are equal to

\[
p^{ab} = \frac{g^{1/2}}{2N} \left( g^{ab} g^{cd} - g^{ac} g^{bd} \right) (g^{cd} - 2N_{(c|d)}) = \mathcal{O}(r^{-3})
\]

and they do not contribute to (3.19).

Thus, the Hamiltonian in asymptotically flat worlds takes the following form:
3.3. Asymptotically flat and closed worlds

\[ H_0 = \int d\sigma_a \left( h_{ab}^I - h_{ab} \right). \]  

(3.21)

This expression is covariant and does not depend on a choice of the spatial coordinates. On the other hand, in Cartesian coordinates, it reduces to the well-known Arnowitt-Deser-Misner energy [41]

\[ E = \int d\sigma_a \left( \partial_b h_{ab}^I - \partial_a h_{bb}^I \right). \]  

(3.22a)

However, the expression (3.22a) is non-covariant, and for the computation in an arbitrary coordinate system, even an asymptotically Cartesian, it may lead, as shown in [49], to an arbitrary value of the energy \( H_0 \).

The formula (3.21) yields the unambiguous result \( H_0 = M \), which confirms the identification of the parameter \( M \) in the metric asymptotics (3.20) with the total mass of an isolated distribution of matter. However, one has to remember that the parameter \( M \) in the external Schwarzschild solution is not a pure mass of matter, but it also includes the gravitational mass defect or the own energy of the gravitational field.

Thus, in the framework of the general relativity theory on the basis of a correct gravitational action function (2.82), one can obtain a covariant definition of the energy of asymptotically flat worlds (in contrast to the opinion expressed in [49]). In a similar way, one can covariantize the expressions [43] for the total momentum and the angular moment of an asymptotically flat gravitating system.

The result (3.22a) for the total energy of the gravitating system in an asymptotically flat world can be derived from the Landau-Lifshitz energy-momentum pseudotensor [27].

When using in the ADM procedure the additional conditions (3.12) with the functions \( f^\mu(t, x) \) that do not depend on time,

\[ f^\mu = f^\mu(x), \quad \dot{f}^\mu(x) = 0, \]  

(3.23)

in view of (3.14) one obtains the physical Hamiltonian, which numerically coincides with (3.21), but is a functional of the independent degrees of freedom of the gravitating system

\[ H_{\text{phys}} (t) [g^A, p_A] = \oint d\sigma^a (\partial^b g_{ab} - \partial_a g_{bb}) \bigg|_{g_{ab} = g_{ab} [g^A, p_A]} . \]  

(3.24)

This statement represents the contents of the Regge-Teitelboim theorem [51] about the physical Hamiltonian in the asymptotically flat worlds.

Since the physical Hamiltonian is determined by the spatial asymptotics of the three-dimensional metric satisfying (3.20), we can illustrate the ADM methods on an example of a linearized gravitational field describing small deviations of the metric from the flat one:

\[ g_{ab} = \delta_{ab} + h_{ab}, \quad h_{ab} \ll 1. \]
Let us assume that the canonical transformation in (3.4) acts only on the phase variables of gravitational field \((g_{ab}, p^{ab})\). We transform them using the decomposition of an arbitrary symmetrical tensor \(f_{ab}\) into the longitudinal, transverse and transverse-traceless parts:

\[
f_{ab} = \partial_a f_b + \partial_b f_a + f_{ab}^T + f_{ab}^{TT}.
\]

Here the transverse part \(f_{ab}^T\) is expressed in terms of the trace \(f_T^a = f_{aa}^T \) using the projection operator

\[
f_{ab}^T = \frac{1}{2} \left[ \delta_{ab} - \frac{\partial_a \partial_b}{\Delta} \right] f_T,
\]

where \(1/\Delta\) is Green’s function of the Laplace operator \(\Delta = \partial_a \partial^a\) in the open flat space. The trace of the transverse part \(f_T^a\) satisfies the equation

\[
\Delta f_T^a = \Delta f_{aa}^T - \partial_a \partial_b f_{ab}.
\]

Using (3.25)-(3.27), we transform the following expression:

\[
\int dt \int d^3 x p^{ab} g_{ab} = \int dt \int d^3 x \left\{ p_{TT}^{ab} g_{ab}^{TT} - g_{ab}^T p_{TT}^{ab} - 2 \partial_b p^{ab} g_{a} \right\},
\]

where \(g_{ab}, g_{ab}^T, g_{ab}^{TT}\) are the components of the decomposition (3.25) of the three-dimensional metric \(g_{ab}\) (a similar notation is used for \(p^{ab}\)). In the second term of this expression, the integration by parts with respect to time \(t\) was done. Taking into account that according to (3.26) \(g_{ab}^T\) is expressed in terms of \(g_{ab}^T\), we obtain

\[
\int dt \int d^3 x p^{ab} g_{ab} = \int dt \int d^3 x \left\{ p_{TT}^{ab} g_{ab}^{TT} - g_{ab}^T p_{TT}^{ab} - 2 \partial_b p^{ab} g_{a} \right\},
\]

from which it is obvious that instead of \((g_{ab}, p^{ab})\) one can introduce the following canonically conjugated variables:

\[
X^a = g_a, \quad \Pi_a = -2 \partial_b p^{ab}, \quad X^0 = \frac{1}{2} \partial_T, \quad \Pi_0 = g_T, \quad g^A = g_{ab}^{TT}, \quad p_A = p_{TT}^{ab}, \quad A = 1, 2.
\]

Separating the linear part with respect to the fields in the gravitational constraints (2.125) and (2.126), we find

\[
H_\perp = \Delta g_{aa} - \partial_a \partial_b g_{ab} + h_\perp (h_{ab}, p^{ab}; \varphi, p),
H_a = -2 \partial_b p^{ba} + h_a (h_{ab}, p^{ab}; \varphi, p),
\]

where the terms \(h_\perp\) and \(h_a\) are of higher order with respect to the fields. In terms of the new phase variables (3.29)-(3.31), the constraints are recast into:

\[
H_\perp = \Delta \Pi_0 + h_0 [X, \Pi; g^A, p_A; \varphi, p],
H_a = \Pi_a + h_a [X, \Pi; g^A, p_A; \varphi, p],
\]
where in terms of the new variables, \( t_0 \) and \( t_a \) become functionals of the second and higher order in their arguments.

Eq. (3.32) shows that the matrix (3.7) in the linear approximation reads

\[
J_{(\mu)}(\mathbf{x}, \mathbf{y}) = \begin{bmatrix}
\Delta \delta(\mathbf{y} - \mathbf{x}) & 0 \\
0 & \delta_a^b \delta(\mathbf{y} - \mathbf{x})
\end{bmatrix}.
\]

(3.33)

Let us impose the additional conditions (3.11) satisfying (3.23):

\[
X^0 = 0, \quad X^a = \frac{1}{2} x^a,
\]

(3.34)

then the Hamiltonian (3.14) takes the form (3.24). We use this form as a volume integral [taking (3.27) into account]

\[
H_0 = - \int d^3 x \, \Delta g^T.
\]

(3.35)

In the linear approximation, it is easy to obtain the dependence of this quantity on the physical degrees of freedom \((g^T_{ab}, p^T_{ab}, \varphi, p)\). When solving the constraint (3.32) in this approximation, we have

\[
- \Delta g^T(\mathbf{x}) = h_0 [X^\mu, \Pi_{\mu}; g^A, p_A; \varphi, p] \bigg|_{X^0 = 0, X^a = \frac{1}{2} x^a, N_\mu = 0}.
\]

Taking into account (2.125) and (2.94), we can find \( t_0 \), from which the physical Hamiltonian of the gravitating system in the linear approximation is equal to

\[
H_{\text{phys}}(t)[g^T_{ab}, p^T_{ab}, \varphi, p] = \int d^3 x \left\{ (p^T_{ab})^2 + \frac{1}{2} (\partial_a g^T_{bc})^2 + \mathcal{H}_m(\varphi, p) \right\}.
\]

(3.36)

In this approach, the purely gravitational degrees of freedom are identified with the two transverse-traceless components of the metric and the conjugated momenta. These are the two well-known transverse polarizations of the gravitational wave.

Let us now determine the lapse and shift functions. In view of (3.33) and the conditions (3.23) which are satisfied by the additional conditions (3.34), the equations for \( N^\mu \) take the following form (in the linear approximation):

\[
\Delta N(\mathbf{x}) = 0, \quad N^a(\mathbf{x}) = 0.
\]

(3.37)

The equation (3.37) should be supplemented with the boundary condition, which is induced by the asymptotically flat type of the metric

\[
N \bigg|_{r \to \infty} = 1,
\]

consequently

\[
N(\mathbf{x}) = 1
\]

(3.38)

with an accuracy up to the terms linear in the fields.
This result demonstrates how important are the boundary conditions in the cosmological problems with the open models. Indeed, the entire dynamics of the gravitating system and its energy are determined by the spatial asymptotics. From the physical point of view, one can explain this as follows. The time and space coordinates which are realized and measured using the physical instruments, can be uniquely determined only on the spatial asymptotics, since there are time-like and space-like Killing vector fields. Thus, the Hamiltonian (3.21) is a generator of the global translation in time, determined asymptotically.

Dynamics of closed cosmological models

In contrast to the asymptotically flat worlds which were recently comprehensively studied, the closed cosmological models and their dynamics remain a stumbling block in the modern gravity theory.

In closed cosmological systems, the surface integral $H_0$ is absent, and therefore the physical Hamiltonian is determined only by the second term in (3.14) and it does not vanish only if $f^\mu(t,x) \neq 0$. The conditions (3.16) for the lapse and shift functions also require that not all $f^\mu(t,x)$ vanish.

Therefore, in order to have a nontrivial dynamics in the closed cosmological models one needs to introduce the canonical gauges that explicitly depend on time. Such a time-dependence of the gauge generates the dynamics of all the physical variables of the theory. Alternatively, one can introduce the evolution of the closed cosmology by fixing the Lagrange multipliers – the lapse and shift functions, avoiding the use of explicitly time-dependent functions. However, this is at most another way to construct the canonical gauge with a parametric time dependence.

We can demonstrate this by the simplest example of the closed Friedman Universe described by the metric

$$ds^2 = -N^2(t)\,dt^2 + R^2(t)\gamma_{ik}\,dx^i\,dx^k,$$

(3.39)

where the lapse function $N(t)$ and the cosmological radius $R(t)$ depend only on time, and $\gamma_{ik}$ is the metric of the three-dimensional hypersphere of a unit radius in some coordinates $x'. Since the model is homogeneous, its Lagrange function $L_g$ in terms of the finite number of degrees of freedom is obtained by the integration of (2.84) over the three-dimensional volume of the Universe:

$$L_g = \frac{1}{2l_0} \left\{ -\frac{l_0^2}{N} \frac{a\dot{a}}{N} + Na \right\},$$

(3.40)

where $a$ is the dimensionless radius of the Universe in the units of the Planck length $l_0 = \sqrt{2G\hbar}/3\pi c^3$, $R = l_0 \, a$.

The superhamiltonian corresponding to (3.40) is equal to

$$\int d^3x \, H(x) = \frac{-\varepsilon_0}{2} \left\{ \frac{P^2}{a} + a \right\},$$

(3.42)
3.3. Asymptotically flat and closed worlds

where \( p \) is the momentum conjugated to the variable \( a \):

\[
p = \frac{\partial L_g}{\partial \dot{a}} = -\frac{l_0}{N} a \dot{a},
\]

(3.43)

and \( \varepsilon_0 = l_0^{-1} \) is the Planck energy (\( \varepsilon_0 = \hbar c/l_0 \) in the usual units).

We assume that the Universe is homogeneously filled with a noninteracting dust without pressure, so that the integral over the three-dimensional volume of the Universe from the matter superHamiltonian \( H_\varphi \) in (2.125) is equal to its conserved mass-energy

\[
\int d^3x H_\varphi(x) = \varepsilon, \quad \varepsilon = \text{const}.
\]

Then the geometrodynamical constraint takes the following form:

\[
H_\perp = 0,
\]

(3.44)

\[
H_\perp = \int d^3x H_\perp(x) = \frac{\varepsilon_0}{2} \left( \frac{p^2}{a^2} - 1 \right) + \varepsilon.
\]

(3.45)

Because of the homogeneity of the Friedman model, other constraints are not important since they are satisfied identically due to the invariance of the system with respect to the spatial translations.

Therefore, a further analysis is equivalent to the ADM procedure with one constraint in which the index \( \mu \) takes only one value.

The equations (3.1) for the complete set of the phase variables \( \Phi = (a,p) \) reduce to

\[
\dot{p} = -\frac{N \varepsilon_0}{2} \left( \frac{p^2}{a^2} - 1 \right),
\]

and (3.43). The constraint equation (3.44) is the first integral of these equations, which corresponds to the condition of preservation of the constraint in time.

The evolution of the Friedman Universe is usually found by choosing the coordinate condition

\[
N = R,
\]

(3.46)

that brings the metric (3.39) to the conformally static form. Combining (3.43) and (3.46) we have \( p = -\dot{a} \), as a result, the constraint (3.44) becomes an equation for \( a(t) \), the solution of which gives

\[
a(t) = \frac{\varepsilon}{\varepsilon_0} (1 - \cos t),
\]

(3.47)

where the moment of time \( t = 0 \) corresponds to the cosmological singularity \( a = 0 \).

Let us show that the gauge (3.46), containing the Lagrange multiplier, corresponds to a canonical gauge that explicitly depends on time. For this, we choose the new canonical variables in accordance with the ADM procedure (3.4) \( \Pi_0 = \Pi, \chi^0 = T \):

\[
\Pi_0, \chi^0 = \text{const}.
\]

Therefore, the evolution of the Friedman Universe is usually found by choosing the coordinate condition

\[
N = R,
\]

(3.46)

that brings the metric (3.39) to the conformally static form. Combining (3.43) and (3.46) we have \( p = -\dot{a} \), as a result, the constraint (3.44) becomes an equation for \( a(t) \), the solution of which gives

\[
a(t) = \frac{\varepsilon}{\varepsilon_0} (1 - \cos t),
\]

(3.47)

where the moment of time \( t = 0 \) corresponds to the cosmological singularity \( a = 0 \).

Let us show that the gauge (3.46), containing the Lagrange multiplier, corresponds to a canonical gauge that explicitly depends on time. For this, we choose the new canonical variables in accordance with the ADM procedure (3.4) \( \Pi_0 = \Pi, \chi^0 = T \):
\[ a = \sqrt{-2\Pi} \cos T, \quad \text{(3.48)} \]
\[ p = \sqrt{-2\Pi} \sin T. \quad \text{(3.49)} \]

The constraint (3.45) in new variables is recast into the form

\[ H_\perp = -\frac{\varepsilon_0}{2} \frac{\sqrt{-2\Pi}}{\cos T} + \varepsilon, \quad \text{(3.50)} \]

and then a solution (3.8) of this constraint with respect to \( \Pi \) reads

\[ \Pi = -P, \quad P \equiv 2 \left( \frac{\varepsilon}{\varepsilon_0} \right)^2 \cos^2 T. \quad \text{(3.51)} \]

Imposing now the additional condition (3.11)

\[ T = \frac{t - \pi}{2}, \quad \text{(3.52)} \]

we find the physical Hamiltonian of the system, see (3.14)

\[ H_{\text{phys}}(t) = \frac{\varepsilon}{\varepsilon_0} \sin^2(t/2). \]

The dynamically independent degrees of freedom of this Hamiltonian are the variables of the dust-like matter, with the energy \( \varepsilon \) being the function of them. Substituting (3.51) into (3.48), we obtain (3.47) again, that is, the canonical gauge (3.52) which can be rewritten in the original phase variables as

\[ a + p \tan(t/2) = 0, \quad \text{(3.53)} \]

is equivalent to the coordinate condition (3.46). One can verify this using (3.15):

\[ N = \left( \frac{\partial H_\perp}{\partial \Pi} \right)^{-1} \dot{T} = l_0 a. \quad \text{(3.54)} \]

Therefore, the choice of a one-parameter family of surfaces in the phase space (3.46) or (3.53) determines the family of the space-like hypersurfaces in the four-dimensional spacetime, and the dynamics of the system reduces to the motion along this family with the growth of the time parameter \( t \). The only difference of fixing the gauge (3.46) from (3.53) is that in (3.53) the family of hypersurfaces is determined not globally, but differentially, by fixing not a one-parameter slicing itself, but rather its derivative with respect to the parameter \( t \). All this is directly extended to the general case of a non-homogeneous gravitating closed cosmological system, but the choice of a system of space-like hypersurfaces and their coordinate parametrization will be determined by the family of the four canonical gauges \( \chi^\mu(t, \Phi) = 0 \), explicitly depending on the time parameter \( t \).

Moreover, each choice of the gauges, i.e., of the functions \( f^\mu(t, x) \) in (3.14) will produce its own Hamiltonian. An ambiguity of determination of the energy in the closed world, in contrast to an asymptotically flat one, is caused by the fact that in closed Universes there are no a preferred observer and a preferred system of the space-like hypersurfaces, which in open systems are associated with the spacetime flat asymptotics.
4

Torsion effects on the structure and evolution of gravitating systems

4.1. Matter fields in the Einstein-Cartan theory

As we demonstrated in Sec. 1.3., the spin-connection interaction does not have an explicit dynamical nature in the Einstein-Cartan theory (ECT). Nevertheless, it modifies the structure of the matter sources of the gravitational field, and thereby the spin of matter affects the dynamics of the metric. Furthermore, it is worthwhile to stress that the ETC is not competing with GR since the \((\Gamma-S)\)-interaction has a microscopic nature and is manifest on a scale of the order of the mean distance between particles. The macroscopic gravitational theory is obtained as a result of an appropriate averaging process. The value of the term quadratic in spin \(T_{\mu\nu}^{\text{eff}}\) becomes comparable with \(T_{\mu\nu}\) at densities \(\rho \geq \rho_c = \frac{m^2 \hbar^2}{E_G}\) of the spinning matter built of particles with the mass \(m\) [7]. For the mass of a nucleon, the critical density \(\rho_c \approx 10^{57}\) kg/m\(^3\) is much smaller than the Planck density \(\rho_0 \sim 10^{97}\) kg/m\(^3\) at which the quantum-gravitational effects start dominating. Consequently, the torsion can be essential already at the level of the classical theory of the gravitational interactions.

Let us begin our study of the torsion effects in ECT with the analysis of the structure of the matter currents.

Weyssenhoff spinning fluid

Before we present a consistent variational theory, it is instructive to show that one can include a fluid with spin into the dynamical scheme of ECT on the
basis of covariant conservation laws of the energy-momentum and spin (1.45) and (1.47). Our presentation follows [191].

The ideal fluid with spin (or a spinning fluid) is a continuous medium, each element of which is characterized by the energy, momentum and an intrinsic angular momentum. The phenomenological model of such a medium was developed by Weyssenhoff and Raabe [54] in the flat Minkowski spacetime. Accordingly, the general-relativistic ideal fluid with spin will also be called a Weyssenhoff fluid.

In the phenomenological approach, elements of a medium are described by the (average) 4-velocity \( u^\mu \), the 4-vector \( P_\mu \) of the energy-momentum density, and by the density of the spin angular momentum \( S_{\mu\nu} = S_{[\mu\nu]} \). The spatial components of the latter constitute the 3-vector \( \mathbf{S} = \{ S_{23}, S_{31}, S_{12} \} \) which in the rest frame of the fluid is equal to the 3-dimensional density of intrinsic angular momentum. The remaining components also constitute the 3-vector \( \mathbf{q} = \{ S_{10}, S_{20}, S_{30} \} \). We assume that this vector vanishes in the rest frame of fluid, i.e.,

\[
S_{\mu\nu} u^\nu = 0. \tag{4.1}
\]

This assumption is called a Frenkel supplementary condition.

The 4-velocity is a timelike vector which is normalized by \( u^\mu u_\mu = -1 \). The projection of the 4-momentum on the rest frame of the fluid yields the energy density \( \varepsilon = -u^\mu P_\mu \).

As a first step, we consider the spinning dust on non-interacting elements with the vanishing pressure. The phenomenological postulate for the canonical energy-momentum tensor and the canonical spin tensor of a dust reads as follows:

\[
\begin{align*}
& t^\lambda_\mu = u^\lambda P_\mu, \\
& S^\lambda_\mu\nu = u^\lambda S_{\mu\nu}.
\end{align*} \tag{4.2}
\]

Making use of the covariant conservation law of the angular momentum (1.47), we find a relation between the densities of the energy-momentum and spin

\[
u_\mu P^\nu - u_\nu P^\mu = 2c(\nabla_\sigma - 2Q_\sigma)(u^\sigma S_{\mu\nu}). \tag{4.3}
\]

Contracting this equation with the 4-velocity, we find the 4-vector of momentum explicitly

\[
P_\mu = \varepsilon u_\mu + 2u^\lambda c\dot{S}_\mu^\lambda. \tag{4.4}
\]

We introduced a convenient notation for the derivative \( \dot{\varphi}_A := (\nabla_\sigma - 2Q_\sigma)(u^\sigma \varphi_A) \) for an arbitrary tensor density \( \varphi_A \). Substituting (4.4) back into (4.3), we recast the latter into the equation of motion of spin

\[
\dot{S}_{\mu\nu} = u_\mu u^\lambda \dot{S}_{\nu\lambda} - u_\nu u^\lambda \dot{S}_{\mu\lambda}. \tag{4.5}
\]

Finally, making use of (4.4) in (4.2), we obtain the canonical energy-momentum tensor of the Weyssenhoff spinning dust

\[
t_{\mu\nu} = \varepsilon u_\mu u_\nu + 2u_\mu u^\lambda c\dot{S}_{\nu\lambda}.
\]
The first term on the right-hand side represents the usual ideal fluid with the
dust equation of state, and the last term is the spin contribution.

The phenomenological theory can be upgraded to a complete dynamical the-
ory of a spinning fluid. In order to construct the Lagrangian, we need to choose
the appropriate variables which describe such a continuous medium. In the
classical continuum mechanics, the ordinary matter consists of the structure-
less elements - the material points without physical properties other than mass.
In contrast, the spinning fluid represents an example of a matter with mi-
structure, the elements of which have additional degrees of freedom such as
spin.

Following Cosserat brothers [25], to describe the dynamics of a medium with
microstructure we attach a triad of vectors with each matter element. The field
of such material triads is denoted $b^\mu_i$, $i = 1, 2, 3$, and together with the 4-velocity
it comprises a local frame $b^\mu_a$, $a = 0, 1, 2, 3$,

$$b^\mu_a = \{b^\mu_0 = u^\mu, b^\mu_1, b^\mu_2, b^\mu_3\}$$
defined at all the spacetime points of the domain occupied by the matter. The
vectors of the triad are chosen orthogonal to each other and to $u^\mu$ and have the
unit length:

$$g^\mu_\nu b^\mu_a b^\nu_b = \eta_{ab} = \text{diag}(-1, +1, +1, +1).$$

Such an orthonormal tetrad is called a material frame and it is different from
the spacetime tetrad $h^\mu_a$ which we discussed in Sec. 1.1.. We can choose $h^\mu_a$
arbitrarily and can always change this tetrad using a local Lorentz rotation. In
contrast, the material tetrad $b^\mu_a$ cannot be chosen arbitrarily, it is rigidly fixed
and moves together with the fluid.

We are now in a position to formulate the variational principle for the spinning
fluid. We begin by postulating (as in the usual case of an ideal fluid) that the
dynamics of the medium is such that the number of particles is constant, and
the entropy $s$ and the identity of particles are conserved along the flow lines of
the fluid. Mathematically, this means that we impose the constraints

$$\partial_\mu (\sqrt{g} \rho u^\mu) = 0, \quad u^\mu \partial_\mu s = 0, \quad u^\mu \partial_\mu X = 0.$$ 

Here $\rho$ is the particle number density, and the Lin variable $X$ is introduced to
identify matter elements.

We will describe the physical properties of the fluid by the particle density $\rho$, the internal energy density $\varepsilon$, and the specific spin density $\mu^{ij} = -\mu^{ji}$ (i.e.,
the spin density per matter element), $i, j = 1, 2, 3$.

With all these prerequisites, the Lagrangian of the spinning fluid reads

$$\mathcal{L}_W = -\varepsilon(\rho, s, \mu^{ij}) + \frac{1}{2} \rho \mu^{ij} g_{\mu \nu} b^\mu_i b^\nu_j + \rho u^\mu \partial_\mu \lambda_1 + \lambda_2 u^\mu \partial_\mu s + \lambda_3 u^\mu \partial_\mu X + \lambda^{ab} (g_{\mu \nu} b^\mu_a b^\nu_b - \eta_{ab}).$$

The first line represents the physically essential part related to the kinetic and
potential energy of the system, and the second line takes into account all the
The constraints which are imposed on the system by means of the Lagrange multipliers $\lambda_1, \lambda_2, \lambda_3$ and $\lambda^{ab}$.

The thermodynamical properties of the spinning fluid are described by the usual Gibbs law corrected by the spin energy contribution:

$$Tds = d\left(\frac{\varepsilon}{\rho}\right) + pd\left(\frac{1}{\rho}\right) - \frac{1}{2}\omega_{ij}d\mu^{ij}.$$  

Here $T$ is the temperature, $p$ is the pressure, and $\omega_{ij}$ is the thermodynamical variable conjugated to the specific spin density $\mu_{ij}$.

The thermodynamic properties of the spinning fluid are described by the usual Gibbs law corrected by the spin energy contribution:

$$Tds = d\left(\frac{\varepsilon}{\rho}\right) + pd\left(\frac{1}{\rho}\right) - \frac{1}{2}\omega_{ij}d\mu^{ij}.$$  

The equations of motion are derived from the variation of the matter action $\frac{1}{c} \int d^4x \sqrt{g} L_W$ with respect to the total set of (fundamental and auxiliary) variables $\Phi^A = \{\rho, s, X, \mu^{ij}, b^0_a, \lambda_1, \lambda_2, \lambda_3, \lambda^{ab}\}$. Variation with respect to the Lagrange multipliers, we obtain the set of constraints, whereas the variation with respect to $\rho, s, X$, and $\mu^{ij}$ yields, respectively:

$$-\varepsilon - p + \frac{1}{2}\rho\mu^{ij} g_{\mu\nu} b^\mu_i \eta^\lambda_j b^\nu_j + \rho u^{\lambda}_{\mu} \partial_{\mu} \lambda_1 = 0,$$

$$\rho T + \lambda_2 = 0, \quad \lambda_3 = 0, \quad \omega_{ij} = g_{\mu\nu} u^\alpha b^\mu_i \nabla_{b^\nu_j}.$$  

Here we had to make use of the Gibbs law to evaluate the derivatives of the internal energy density $\varepsilon$ with respect to its arguments.

Finally, from the variation of the material tetrad, $u^\mu$ and $b^\mu_i$, we find

$$\frac{1}{2}\rho\mu^{ij} g_{\mu\nu} b^\mu_i \eta^\lambda_j b^\nu_j + \rho \partial_{\mu} \lambda_1 + \lambda_2 \partial_{\mu} s + \lambda_3 \partial_{\mu} X + 2\lambda^0_a g_{\mu\nu} b^\mu_a = 0,$$

$$g_{\mu\nu} \left(\rho\mu^{ij} u^\lambda_i \partial_{\lambda} b^\nu_j + \frac{1}{2} \eta^\lambda_i b^\nu_j \partial_{\lambda} \mu^{ij} + 2\lambda^0_a b^\nu_j\right) = 0.$$  

Contracting these equations with $u^\mu$ and $b^\mu_k$, we find the Lagrange multipliers

$$2\lambda^{00} = \varepsilon + p, \quad 2\lambda^{0i} = \rho \mu^{ij} u^\lambda_i u^\alpha \nabla_{b^\nu_j}, \quad 2\lambda^{ik} = -\rho \mu^{(ij|k)} b^\nu_j u^\alpha \nabla_{b^\nu_j},$$

and obtain the equations of motion of spin:

$$u^\alpha \partial_{\alpha} \mu^{ij} + \mu^{ik} \omega^l_j + \mu^{kj} \omega^l_i = 0.$$  

The Latin 3-dimensional indices that refer to the material tetrad are moved with the help of the Euclidean metric $\delta^{ij}$ and $\delta_{ij}$. We also introduced the inverse material triad by $b^\mu_i = g_{\mu\nu} \eta^{ij} b^\nu_j = g_{\mu\nu} \delta^{ij} b^\nu_j$. One can verify the relations

$$b^\mu_i b^\mu_i = \delta^i_j, \quad b^\mu_i b^\mu_i = \delta_{ij} + u^\mu u_\nu.$$  

Finally, we apply these results to derive the canonical tensors of energy-momentum and spin. Recalling the definitions of Sec. 1.3., we find from the Weyssenhoff Lagrangian $L_W$:

$$t^\lambda_{\mu} = p \delta^\lambda_{\mu} + (p + \varepsilon) u^\lambda u_\mu + 2 u^\lambda u^\nu c \mathcal{S}_{\mu\nu}, \quad S^\lambda_{\mu\nu} = u^\lambda \mathcal{S}_{\mu\nu},$$

(4.6)
where we have the spin density explicitly

$$ S^{\mu\nu} = \frac{1}{2} \rho \mu^{ij} b_i^\mu b_j^\nu. $$

By construction, the spin tensor satisfies the Frenkel condition (4.1). The canonical sources (4.6) provide a consistent generalization of the phenomenological result (4.2) to the case of fluid with nontrivial pressure $p$.

This completes the construction of the variational theory of the Weyssenhoff spinning fluid. It is applicable to any gravitational theory.

Let us now specialize to the Einstein-Cartan gravity with the Lagrangian (1.34) and the field equations (1.48), (1.49). Combining last equation with the Frenkel condition (4.1), we find the torsion $Q_{\lambda\mu\nu} = \kappa c_{\lambda} S_{\mu\nu}$ (the torsion trace vanishes, $Q_\mu = 0$, since $S_\mu = u^\nu S_{\mu\nu} = 0$). Plugging this into (1.51), we find the effective Lagrangian of the spinning fluid

$$ L^{\text{eff}}_W = -\varepsilon(\rho, s, \mu^{ij}) + \frac{1}{2} \rho \mu^{ij} g_{\mu\nu} b_i^\mu b_j^\nu + \frac{\kappa c^2}{8} \rho^2 \mu^{ij} \mu_{ij} + \rho u^\mu \partial_\mu \lambda_1 + \lambda_2 u^\mu \partial_\mu s + \lambda_3 u^\mu \partial_\mu X + \lambda_3^{ab}(g_{\mu\nu} b_a^\mu b_b^\nu - \eta_{ab}). $$

The last term in the first line describes the typical spin-spin contact interaction. One can also write this as

$$ \frac{\kappa c^2}{8} \rho^2 \mu^{ij} \mu_{ij} = \kappa c^2 S^2, \quad S^2 := \frac{1}{2} S^{\mu\nu} S_{\mu\nu}, $$

but this form produces a misleading impression that the spacetime metric tensor is involved. In reality, the effective spin-spin interaction term depends only on the particle density $\rho$ and on the specific spin density $\mu^{ij}$. This is important to keep in one’s mind when performing the variation of the effective fluid action with respect to the metric.

Taking into account that the variation of the Christoffel symbols reads

$$ \delta \{^\lambda_{\mu\nu}\} = \frac{1}{2} g^{\lambda\alpha}(\nabla_\mu \delta g_{\nu\alpha} + \nabla_\nu \delta g_{\mu\alpha} - \nabla_\alpha \delta g_{\mu\nu}), $$

it is straightforward to find the effective metrical energy-momentum tensor

$$ T^{\text{eff}}_{\mu\nu} = p^{\text{eff}} g_{\mu\nu} + (\varepsilon^{\text{eff}} + p^{\text{eff}}) u_\mu u_\nu + 2c \left( u^\alpha u^\beta - g^{\alpha\beta} \right) \nabla_\alpha \left( u_{(\mu} S_{\nu)} \right), $$

$$ \varepsilon^{\text{eff}} = \varepsilon - \kappa c^2 S^2, \quad p^{\text{eff}} = p - \kappa c^2 S^2. $$

Summarizing, the structure of the sources of the gravitational field in the Riemann-Cartan spacetime with torsion for the Weyssenhoff spinning fluid is given by (4.6). For the case of the Einstein-Cartan theory, the dynamics of the gravitational field is described by Einstein’s equation (1.52) with the effective energy-momentum tensor (4.7) that depends on the spin density square and the derivatives of the spin tensor.

The phenomenological way to introduce the spin in the fluid model was discussed in the numerous papers, and a number of variational principle approaches for the spinning fluid were developed, see [12], for example.
Spinor fields in ECT

Of a particular interest is the study of the self-gravitating fields with spin which due to the contact character of spin-spin interaction of the Einstein-Cartan model gives rise to a nonlinear extension of GR expressed by the effective equation (1.52). The spin-spin interaction caused by space-time torsion is manifest as a nonlinearity of these fields, as we demonstrate below.

We consider here the most important physical example of the spinor Dirac field. As the Lagrangian of the Dirac field $\Psi$ in the space $U^4$, we take

$$L_D(\Psi, \bar{\Psi}, \omega_\mu, g_{\mu\nu}) = \frac{\hbar c}{2} \left( \bar{\Psi} \gamma^a D_a \Psi - D_a \bar{\Psi} \gamma^a \Psi \right) + mc^2 \bar{\Psi} \Psi. \quad (4.8)$$

Here $m$ is the rest mass of the Dirac fermion $\Psi$, $\gamma^a$ are the flat Dirac matrices in Weinberg’s representation [192], and $D_a = h^a_\mu D_\mu$ with

$$D_\mu \Psi = \partial_\mu \Psi + \omega_\mu \Psi.$$  

The expression for the covariant derivatives of spinors is defined by formulas (A2.7)-(A2.8) of Appendix A2.

As independent dynamical variables, one can take $\{\bar{\Psi}, \Psi, g_{\mu\nu}, \Gamma^\lambda_{\mu\nu}\}$, on which the metricity condition (1.8) is imposed as a constraint. If the latter is solved explicitly, the connection $\Gamma^\lambda_{\mu\nu}$ takes the form (1.12), and the spinor connection (A2.6) can be recast into

$$\omega_\mu = -\frac{1}{4} \gamma^\nu \nabla_\mu \gamma_\nu + \frac{1}{4} \gamma^\nu T^\lambda_{\mu\nu} \gamma_\lambda. \quad (4.9)$$

However, since the tetrad field is fundamentally involved in the definition of the spinor connection and the covariant derivative, it is more convenient to switch to the tetrad formalism and to take the equivalent set of independent dynamical variables $\{\bar{\Psi}, \Psi, h^a_\mu, \Gamma^a_{\mu\nu}\}$. The metricity condition (1.8) in the tetrad formalism reads

$$\nabla_\mu \eta_{ab} = -\Gamma_{a\mu} - \Gamma_{b\mu} = 0.$$  

Consequently, we conclude that the metric-compatible local Lorentz connection is skew-symmetric: $\Gamma_{a\mu} = \Gamma_{b\mu}$. The same property then shares the tetrad curvature tensor: $R_{a\mu\nu} = R_{b\mu\nu}$.

The field equations of the theory are derived from the variational problem for the total Lagrangian (1.34) in the tetrad disguise

$$S = \int d^4x \ h \left\{ \frac{1}{2\kappa c} R^a_{\mu\nu} h^\mu_b h^\nu_b + \frac{1}{c} L_D \right\}. \quad (4.10)$$

Here we denoted the determinant of the tetrad $h := \det h^a_\mu$; obviously we have $h = \sqrt{g}$. In the tetrad formalism we do not need to add the metricity condition with the Lagrange multipliers, it is sufficient just to take the skew-symmetric local Lorentz connection $\Gamma^a_{b\mu}$. 


4.1. Matter fields in the Einstein-Cartan theory

Variation of the action (4.10) with respect to the independent variables yields:

$$\frac{\delta S}{\delta h_{\mu}^a} = \frac{1}{\kappa c} (R^a_{\mu} - \frac{1}{2} h_{\mu}^a) - \frac{L_D}{c} h_{\mu}^a - \frac{\hbar}{2} (D_{\mu} \bar{\Psi} \gamma^a \Psi - \bar{\Psi} \gamma^a D_{\mu} \Psi) = 0,$$  \hfill (4.11)

$$\frac{\delta S}{\delta \Gamma^{ab\mu}} = - \frac{2}{\kappa c} \mathcal{D}_\nu (hh_{[a}^\nu h_{b]}^\mu) + \frac{\hbar}{8} \{ \gamma^\mu \gamma_{[a} \gamma_{b]} + \gamma_{[a} \gamma_{b]} \gamma^\mu \} \Psi = 0,$$  \hfill (4.12)

$$\frac{\delta S}{\delta \Psi} = h \gamma^a (D_a - Q_a) \Psi + mc \Psi = 0,$$  \hfill (4.13)

$$\frac{\delta S}{\delta \bar{\Psi}} = - h (D_a - Q_a) \bar{\Psi} + mc \bar{\Psi} = 0.$$  \hfill (4.14)

Here the covariant derivative $\mathcal{D}_\mu$ acts only on the tetrad indices. For example, $\mathcal{D}_\mu v^\nu = \partial \mu v^\nu - \Gamma^b_{\mu b} v^\nu$ for an arbitrary tensor object that carries both (world and tetrad) types of indices. Directly form the definition of the torsion, one can verify the useful geometrical identity

$$2 \frac{\hbar}{h} \mathcal{D}_\nu (hh_{[a}^\nu h_{b]}^\mu) \equiv Q^\mu_{ab} + h_{a}^\mu Q_b - h_{b}^\mu Q_a.$$  \hfill (4.15)

Making use of the definitions of Sec. 1.3., we find the canonical energy-momentum and spin tensors

$$t^\lambda_{\mu} = \mathcal{L}_D \delta^\lambda_{\mu} + \frac{\hbar c}{2} (D_{\mu} \bar{\Psi} \gamma^\lambda \Psi - \bar{\Psi} \gamma^\lambda D_{\mu} \Psi),$$

$$S^\lambda_{\mu\nu} = \frac{\hbar}{8} \{ \gamma^\lambda \gamma_{[\mu} \gamma_{\nu]} + \gamma_{[\mu} \gamma_{\nu]} \gamma^\lambda \} \Psi = \frac{i \hbar}{4} \varepsilon_{\mu\nu}^\lambda \bar{\gamma}_\sigma \gamma_5 \Psi.$$  \hfill (4.16)

Since the spin tensor of the Dirac field is completely antisymmetric, the torsion trace vanishes $Q_\mu = 0$ and hence the Dirac equation (4.13) in $U_4$ is simplified,

$$h \gamma^a D_a \Psi + mc \Psi = 0,$$  \hfill (4.13a)

and a similar result is found for the conjugated equation (4.14).

With the help of (4.15), the Palatini equation (4.12) determines the torsion

$$Q^\lambda_{\mu\nu} = \kappa c S^\lambda_{\mu\nu} = \frac{i \hbar c}{4} \varepsilon_{\mu\nu}^\lambda \bar{\gamma}_\sigma \gamma_5 \Psi.$$  \hfill (4.16)

This means that in the irreducible decomposition (1.13) only the pseudotrace of torsion is nontrivial, $\bar{Q}_\mu = \frac{i \hbar c}{4} \bar{\gamma}_\mu \gamma_5 \Psi$.

It is straightforward to see that the fermion Lagrangian (4.8) vanishes “on-shell”, $\mathcal{L}_D = 0$, when the spinor fields that satisfy the Dirac equation (4.13a). As a result, Einstein’s equation (4.11) reads (converting the tetrad indices into the world ones):

$$G_{\mu\nu}(\Gamma) = \kappa t_{\mu\nu} = \frac{\kappa \hbar c}{2} (D_\nu \bar{\Psi} \gamma_\mu \Psi - \bar{\Psi} \gamma_\mu D_\nu \Psi).$$  \hfill (4.17)

The system (4.13), (4.13a), (4.16), and (4.17) represents the complete set of the dynamical equations for the Dirac field $\Psi, \bar{\Psi}$ interacting with the gravitational field $g_{\mu\nu}, \Gamma^\lambda_{\mu\nu}$. 
Using (4.9) and (4.16), we recast the Dirac equation (4.13a) into

\[ \gamma^a \{ D_a \Psi + \frac{l_0^2}{\hbar} (\overline{\Psi} \gamma_a \gamma_5 \Psi) \gamma^a \gamma_5 \Psi + \frac{mc}{\hbar} \Psi = 0, \]  

(4.18)

where \( l_0^2 = \frac{3}{8} \kappa \hbar c \), i.e., \( l_0 \approx 4.95 \times 10^{-35} \) m. We thus conclude that the torsion induces a \( \Psi^4 \)-nonlinearity in the Dirac equation (4.18).

Combining the general formula for the Riemann-Cartan connection (1.12) with (4.16), we can decompose the Einstein tensor into the Riemannian and post-Riemannian parts:

\[ G_{\mu\nu}(\Gamma) = G_{\mu\nu} + \left( \frac{\kappa \hbar c}{4} \right)^2 \left[ g_{\mu\nu} (\overline{\Psi} \gamma_a \gamma_5 \Psi) (\overline{\Psi} \gamma_a \gamma_5 \Psi) + 2 (\overline{\Psi} \gamma_\mu \gamma_5 \Psi) (\overline{\Psi} \gamma_\nu \gamma_5 \Psi) \right]. \]

(4.19)

We need only the symmetric part, since the skew-symmetric part of (4.17) is satisfied due to the conservation law of the total angular momentum. The validity of the latter can be directly verified using the field equation (4.13) and the expressions for the canonical energy-momentum and spin.

As we described in Sec. 1.3., the dynamics of the system is determined by the effective Lagrangian

\[ \mathcal{L}_D = \frac{\hbar c}{2} \left\{ \overline{\Psi} \gamma^a D_a \Psi - (\overline{\Psi} \gamma^a D_a \Psi + \frac{l_0^2}{\hbar} (\overline{\Psi} \gamma_a \gamma_5 \Psi) (\overline{\Psi} \gamma_a \gamma_5 \Psi)) \right\} + mc^2 \overline{\Psi} \Psi. \]

(4.20)

The gravitational field equation (4.17) then reduces to the Einstein equation (1.52) with the effective energy-momentum tensor

\[ T_{\mu\nu}^{\text{eff}} = \frac{\hbar c}{2} \left\{ \overline{\Psi} \gamma^a D_a \Psi - (\overline{\Psi} \gamma^a D_a \Psi + \frac{l_0^2}{\hbar} (\overline{\Psi} \gamma_a \gamma_5 \Psi) (\overline{\Psi} \gamma_a \gamma_5 \Psi)) \right\}. \]

(4.21)

Now let us consider a nonlinear spinor field of the Ivanenko-Heisenberg [56] type as a source of the gravitational field in the framework of the Einstein-Cartan theory. Its generally covariant Lagrangian reads

\[ \mathcal{L}_{IH} = \frac{\hbar c}{2} \left\{ \overline{\Psi} \gamma^a D_a \Psi - (\overline{\Psi} \gamma^a D_a \Psi + \lambda_\mu^2 (\overline{\Psi} \gamma_\mu \gamma_5 \Psi) (\overline{\Psi} \gamma_\mu \gamma_5 \Psi)) \right\}, \]

(4.22)

where \( \lambda_\mu \) is the coupling constant with the dimension of length.

The nonlinear spinor Ivanenko-Heisenberg equation in the Riemann-Cartan spacetime \( U_4 \) follows from the Lagrangian (4.22):

\[ \gamma^\mu D_\mu \Psi + \lambda_\mu^2 (\overline{\Psi} \gamma_\mu \gamma_5 \Psi) \gamma^\mu \gamma_5 \Psi = 0. \]

The canonical tensor of spin coincides with that of the Dirac field, and the torsion is again determined by the Palatini equation (4.16). Substituting the torsion, we recast the nonlinear spinor equation into the form similar to (4.18):

\[ \gamma^\mu \{ D_\mu \Psi + \frac{l_0^2 + \lambda_\mu^2}{\hbar} (\overline{\Psi} \gamma_\mu \gamma_5 \Psi) \gamma^\mu \gamma_5 \Psi = 0. \]

(4.23)
If we fix the coupling constants so that \( \lambda^2 = l_0^2 \) and choose the upper sign in the last term of (4.23), the latter becomes the field equation of a Dirac neutrino in the Riemannian spacetime

\[
\gamma^\mu D_\mu \Psi = 0. \tag{4.24}
\]

The metrical energy-momentum tensor coincides with the symmetric part of the canonical energy-momentum tensor, and for the nonlinear spinor field with the coupling constant \( \lambda^2 = l_0^2 \) we derive

\[
T_{\mu\nu} = \frac{\hbar c}{2} \left\{ D_\mu \overline{\Psi} \gamma_\nu \Psi - \overline{\Psi} D_\nu \gamma^\mu \Psi + g_{\mu\nu} l_0^2 (\overline{\Psi} \gamma_a \gamma_5 \Psi)(\overline{\Psi} \gamma^a \gamma_5 \Psi) \right\}. \tag{4.25}
\]

Decomposing this into the Riemannian and post-Riemannian parts and combining with (4.19), we arrive at the Einstein equation (1.52) with the effective energy-momentum tensor

\[
T^{\text{eff}}_{\mu\nu} = \frac{\hbar c}{2} \left\{ D_\mu \overline{\Psi} \gamma_\nu \Psi - \overline{\Psi} D_\nu \gamma^\mu \Psi \right\}. \tag{4.26}
\]

The whole dynamics of the coupled spinor and gravitational fields is described by the corresponding effective Lagrangian

\[
\mathcal{L}^{\text{eff}}_{IH} = \frac{\hbar c}{2} \left\{ \overline{\Psi} \gamma_\alpha D_\alpha \Psi - D_\alpha \overline{\Psi} \gamma^\alpha \Psi \right\}.
\]

We can formulate the conclusions that follow from our observations in the form of the two equivalence theorems as follows.

The Einstein-Cartan theory of the neutrino (massless Dirac spin-\( \frac{1}{2} \) fermion) field in the Riemann-Cartan spacetime

\[
G_{(\mu\nu)}(\Gamma) = \frac{\kappa \hbar c}{2} \left\{ D_\mu \overline{\Psi} \gamma_\nu \Psi - \overline{\Psi} D_\nu \gamma^\mu \Psi \right\}, \quad \gamma^\alpha D_\alpha \Psi = 0,
\]

is equivalent to the Einstein theory of the nonlinear spinor (of the Ivanenko-Heisenberg type) field in the Riemannian spacetime:

\[
G_{(\mu\nu)} = \frac{\kappa \hbar c}{2} \left\{ D_\mu \overline{\Psi} \gamma_\nu \Psi - \overline{\Psi} D_\nu \gamma^\mu \Psi - g_{\mu\nu} l_0^2 (\overline{\Psi} \gamma_a \gamma_5 \Psi)(\overline{\Psi} \gamma^a \gamma_5 \Psi) \right\}, \quad \gamma^\alpha D_\alpha \Psi + l_0^2 (\overline{\Psi} \gamma_a \gamma_5 \Psi) \gamma^a \gamma_5 \Psi + \frac{mc}{\hbar} \Psi = 0.
\]

In a similar way, under the condition \( \lambda^2 = l_0^2 \), the Einstein-Cartan theory of the nonlinear spinor field in the Riemann-Cartan spacetime

\[
G_{(\mu\nu)}(\Gamma) = \frac{\kappa \hbar c}{2} \left\{ D_\mu \overline{\Psi} \gamma_\nu \Psi - \overline{\Psi} D_\nu \gamma^\mu \Psi + g_{\mu\nu} l_0^2 (\overline{\Psi} \gamma_a \gamma_5 \Psi)(\overline{\Psi} \gamma^a \gamma_5 \Psi) \right\}, \quad \gamma^\alpha D_\alpha \Psi - l_0^2 (\overline{\Psi} \gamma_a \gamma_5 \Psi) \gamma^a \gamma_5 \Psi + \frac{mc}{\hbar} \Psi = 0
\]
is equivalent to the Einstein theory of the neutrino Dirac spinor field in the Riemannian spacetime:

\[ G_{(\mu\nu)} = \frac{\kappa \hbar c}{2} \left\{ D_{(\mu} \Psi \gamma_{\nu)} - \Psi \gamma_{(\mu} D_{\nu)} \right\}, \]

\[ \gamma^\alpha D_\alpha \Psi = 0, \]

These results actually reveal the geometrical nature of nonlinearity of the Ivanenko-Heisenberg [58] equations. It must be stressed that our conclusions follow from the analysis of the complete system of equations describing the self-gravitating spinor fields.

However, it is worthwhile to mention that the equality of the coupling constants \( l_0 \) and \( \lambda_p \) in the Ivanenko-Heisenberg equation is physically unsubstantiated unless it becomes a natural part of a fundamental theory, or the specific physical conditions arise which lead to such an equality. In particular, one can show that the aforementioned equivalence may take place at the initial stages of the cosmological evolution [60] by means of the spontaneous symmetry breaking mechanism of the vacuum state of a self-interacting scalar field with conformal coupling [71, 72] in an external metric of an open Friedman model [61].

Another possibility to adjust the coupling constants \( l_0 \) and \( \lambda_p \), is to extend the Einstein-Cartan model (as it is actually required in the gauge theory for a non-semisimple group) by including into the gravitational Lagrangian, along with the Hilbert term with Einstein’s coupling constant \( \kappa \), the term quadratic in the torsion tensor with an additional dimensionless coupling constant \( \chi \). A particular example of such an extended theory gives the Lagrangian \( \frac{1}{2 \kappa c} (R(\Gamma) + \chi Q^\mu \lambda^\nu Q_{\mu \lambda^\nu}) \).

Vector fields in ECT

The gravitational interaction of vector fields is an interesting issue in the framework of the Einstein-Cartan theory. If we start with a theory of a massless vector field \( A_\mu \) in the flat spacetime, and formally apply the minimal coupling principle by replacing the ordinary derivatives by the covariant ones, then in \( U_4 \) we obtain the Lagrangian

\[ \mathcal{L}_A = - \frac{1}{16\pi} F_{\alpha\beta}(\Gamma) F^{\alpha\beta}(\Gamma), \]

where the generalized vector field strength tensor reads

\[ F_{\alpha\beta}(\Gamma) = \nabla_\alpha A_\beta - \nabla_\beta A_\alpha = F_{\alpha\beta} + 2Q^{\lambda}_{\alpha\beta} A_\lambda. \]

\(^1\)The dynamics of the two-component neutrino in the \( U_4 \) Riemann-Cartan spacetime identically coincides with the dynamics of the neutrino in the Riemann spacetime. Within the so-called strong gravitation of A. Salam [59] one has \( l_0 = \lambda_p \).
Here \( F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha \) is the usual Maxwell tensor. The field equations for the coupled gravitational and vector fields are obtained from the action \((1.34)\) by the variation with respect to the fundamental variables \(g_{\mu\nu}, \Gamma^\lambda_{\mu\nu}, A_\mu\), and the Lagrange multipliers \(\Lambda^{\alpha\beta\gamma}\). The resulting system of the Einstein-Cartan equations \((1.48)\) and \((1.49)\) describes the dynamics of the gravitational fields \((g_{\mu\nu}, Q^\lambda_{\mu\nu})\) created by its sources: the canonical tensors of the energy-momentum and the spin of the vector field

\[
t_{(\mu\nu)} = \frac{1}{4\pi} \left\{ F_{\mu\alpha}(\Gamma) F^{\alpha}_{\nu}(\Gamma) - \frac{1}{4} g_{\mu\nu} F_{\alpha\beta}(\Gamma) F^{\alpha\beta}(\Gamma) + (\nabla_\lambda - 2Q_\lambda) A_{(\mu} F_{\nu)}^{\lambda}(\Gamma) \right\},
\]

\[
S^\lambda_{\mu\nu} = - \frac{1}{4\pi c} A_{[\mu} F_{\nu]}^{\lambda}(\Gamma).
\]

In addition, the dynamics of the vector field \(A_\mu\) is described by the equation

\[
(\nabla_\alpha - 2Q_\alpha) F^{\alpha\beta}(\Gamma) = 0. \tag{4.27}
\]

Let us solve the Palatini equation \((1.49)\) with respect to the torsion tensor. This is a nontrivial task since the torsion appears on both sides (explicitly on the left-hand side, and implicitly in \(F_{\mu\nu}(\Gamma)\) on the right-hand side). After some algebra we find

\[
Q^\lambda_{\mu\nu} = - \frac{2G}{c^4} \left\{ A_{[\mu} F_{\nu]}^{\lambda} + \left( \frac{G}{c^2} A^{\lambda} A_{\mu} - \frac{1}{2} \delta^\lambda_{[\mu} A_{\nu]} A^\sigma \right) \right\}. \tag{4.28}
\]

Here we denoted \(A^2 = A_\mu A^\mu\), and let us remind that \(F_{\mu\nu}\) (no dependence on \(\Gamma\)) is the usual Maxwell tensor. For the trace we have a simple result

\[
Q_\mu = - \frac{G}{2c^4} F_{\mu\nu}(\Gamma) A^\nu = - \frac{G}{2c^4} \frac{F_{\mu\nu} A^\nu}{1 - \frac{G}{c^3} A^2}.
\]

Note that \(Q^\mu A_\mu = 0\).

Substituting \((4.28)\) into \((4.27)\), we obtain the generalized “Maxwell equation”, the nonlinearity of which is caused by space-time torsion induced by the vector field spin \([62]\)

\[
\frac{1}{\sqrt{g}} \partial_\alpha \left( \sqrt{g} F^{\alpha\beta}(\Gamma) \right) = \frac{4\pi}{c} J^\beta, \tag{4.29}
\]

where the effective 4-current reads

\[
J^\beta = \frac{G}{4\pi c^3} F^{\alpha\beta}(\Gamma) F_{\alpha\gamma}(\Gamma) A^\gamma.
\]

This current satisfies the continuity relation \(\partial_\beta(\sqrt{g} J^\beta) = 0\) by virtue of the field equation \((4.29)\). Making use of the torsion \((4.28)\), we have the explicit form of the generalized vector field strength

\[
F_{\mu\nu}(\Gamma) = F_{\mu\nu} - \frac{2G}{c^4} A_{[\mu} F_{\nu]}^{\lambda} A^\lambda \frac{1}{1 - \frac{G}{c^3} A^2}. \tag{4.30}
\]
As we explained in Sec. 1.3., the system of the vector field equation (4.27), (4.29) plus the gravitational equation (1.46) can be rewritten as the system of the Einstein equation of GR and of the non-linear vector field equation described by the total Lagrangian \( \frac{1}{8\pi} R + \mathcal{L}_{\text{eff}}^A \) with the effective Lagrangian

\[
\mathcal{L}_{\text{eff}}^A = -\frac{1}{16\pi} \left\{ F_{\mu\nu}F^{\mu\nu} + \frac{2G}{\pi} F_{\mu\alpha}F^{\mu\beta} A^\alpha A_\beta - \frac{2}{\sqrt{g}} \frac{\partial_\mu}{1 - \frac{G}{c^4} A^2} \left( \sqrt{g}F^{\mu\nu} A_\nu \right) \right\}.
\]

The torsion was eliminated with the help of (4.28), so that \( R \) is the scalar of the Riemannian curvature, and \( F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \) is the usual Maxwell tensor. The last term is a total divergence and it does not contribute to the field equations. One can directly verify that the nonlinear vector field equation (4.29) is obtained from the effective Lagrangian above by means of the variation with respect to the vector field \( A_\mu \), so that \( J^\beta = -c\partial \mathcal{L}_{\text{eff}}^A / \partial A_\beta \).

When introducing the interaction of a massless vector field with the spacetime torsion, the gauge invariance of the massless vector field theory (4.29) is violated\(^2\), which essentially changes the dynamical content of the theory. Indeed, the usual Maxwell equations describe the massless spin-1 particle (photon) with the two degrees of freedom. But now the vector field \( A_\mu \) has 3 degrees of freedom and thus cannot be interpreted as the electromagnetic field. Nevertheless, the resulting nonlinear theory still does have the current conservation law, namely \( \partial_\beta (\sqrt{g} J^\beta) = 0 \), i.e., the theory under consideration is partially gauge-invariant in the sense of Glashow and Gell-Mann [64]. Partial gauge-invariant nature of the vector field in the presence of the gravitation is well consistent with the interaction hierarchy, where the weaker interaction violates the symmetries typical for the stronger interaction. In the framework of such a hierarchy, the requirement of the local gauge invariance, according to which each conservation law corresponds to some gauge field, turns out to be valid only approximately, when the weaker interactions are neglected.

Let us study the “electrostatic” solutions of the equation (4.29) in the Minkowski spacetime with \( g_{\mu\nu} = \text{diag}(-1, +1, +1, +1) \). Denoting the Cartesian coordinates \( x^\mu = \{ t, x^i \} \), with \( i = 1, 2, 3 \), we use the standard ansatz for the 4-potential: \( A_\mu = \{ \varphi(x^i), 0, 0, 0 \} \). The generalized “electromagnetic” field strength is defined in terms of the tensor \( F_{\alpha\beta}(\Gamma) = (E, B) \), which is a covariant (in terms of the spacetime \( U_4 \)) generalization of Maxwell’s tensor \( F_{\alpha\beta} \). Then (4.29) reduces to the Gauss type equation [62]

\[
\nabla \cdot E = 4\pi \rho,
\]

where from (4.30) we find

\[
\rho = \frac{G}{4\pi c^4} E^2 \varphi, \quad E = \frac{-\nabla \varphi}{1 + \frac{G}{c^4} \varphi^2}.
\]

\(^2\)Something similar happened for the covariant generalization of spinor, neutrino (massless) equations, when their conformal invariance was violated.
The equation (4.31) for the function $\varphi$ explicitly reads
\[
\left(1 + \frac{G}{c^4}\varphi^2\right)\Delta \varphi - \frac{G}{c^4}\varphi(\nabla \varphi)^2 = 0.
\]
Although the resulting nonlinear equation looks quite nontrivial, remarkably there exists a substitution $\varphi = \frac{c^2}{\sqrt{G}} \sinh \chi$, that yields a linear Laplace equation of Maxwell' electrostatics $\Delta \chi = 0$.

It is straightforward to obtain a spherically symmetric solution of (4.31) for $\varphi(x^i) = \varphi(r)$, with $r = \sqrt{x_i x^i}$:
\[
\varphi = \frac{c^2}{\sqrt{G}} \sinh \left(\frac{q \sqrt{G}}{c^2 r}\right).
\]

Here $q$ is an integration constant with a dimension of a charge. This solution obviously has a Coulomb asymptotic behaviour at infinity, for $r \to \infty$. However, the usual Coulomb singularity at $r = 0$ is absent\(^3\). As a result, the corresponding “electric” field and the “charge” density are also regularized:
\[
E = q \frac{\cosh(q \sqrt{G}/c^2 r)}{c^2 r}, \quad \rho = \frac{\sqrt{G} \sinh(q \sqrt{G}/c^2 r)}{4\pi c^2 \cosh^2(q \sqrt{G}/c^2 r)} \frac{q^2}{r^4}.
\]

The qualitative behaviour of the solution, see Fig. 4.1, is as follows\(^4\): Both functions display the Coulomb asymptotics at infinity, whereas they both vanish at the origin, $E(0) = 0, \rho(0) = 0$. The integral over the 3-space of the charge density is finite and easily evaluated:
\[
\int_V d^3 x \rho = q,
\]
so that we can indeed identify the integration constant $q$ with the total “charge”. The nonsingular generalized “electric” field reaches its largest value at $r_{E_{\text{max}}} = 0.14 \times 10^{-24} q$, whereas the charge density is maximal at $r_{\rho_{\text{max}}} = 0.07 \times 10^{-24} q$.

To make an estimate, for an elementary charge of an electron $q = 4.8 \times 10^{-10}$ (in Gaussian absolute electrostatic units), we find $r_{E_{\text{max}}} = 0.7 \times 10^{-34}$ cm and $r_{\rho_{\text{max}}} = 0.35 \times 10^{-34}$ cm, respectively.

One can evaluate the energy of the regular “electromagnetic” field configuration corresponding to this solution. It is straightforward to write down the energy density for the spherically symmetric solution making use of the canonical energy-momentum tensor:
\[
\mathcal{H} = t_{00} = \frac{1}{8\pi} E^2 - \varphi \rho + \frac{1}{4\pi} \nabla(\varphi E).
\]

\(^3\)Note, that such form of the potential is obtained by Urbakh within proposed nonlinear electrodynamics [55].
\(^4\)Taking into account the static spherically symmetric gravitational metric field $g_{\mu\nu}$ does not essentially change the behaviour of $E(r)$ and $\rho(r)$ [62, 67].
Using the Gauss theorem we see that the last term does not contribute to the total energy since both the potential \( \varphi \) and the “electrostatic” field \( E \) vanish at the spatial infinity. The first term on the right-hand side is analogous to the energy density of linear (Maxwell’s) electromagnetic field, whereas the second may be interpreted as the energy density of the self-interaction due to the nonlinear character of the resulting effective theory. The latter term is singular at the origin and it makes a divergent contribution to the integrated total mass \( \int d^3x \mathcal{H} \). Subtracting this divergent self-interaction term, we find the total mass of “electrostatic” field configuration:

\[
m_{em} = \frac{W}{c^2} = \frac{1}{8\pi c^2} \int d^3x \, E^2 = \frac{q}{\sqrt{G}}.
\]

For the electron charge, \( m_{em} = \sqrt{\frac{\alpha \hbar c}{G}} = 1.86 \times 10^{-9} \) kg, which is close to the mass of the classical Markov maximon [66] (with the fine structure constant \( \alpha \approx 1/137 \)).

We conclude our discussion of the vector fields by a brief outline of the massive case. The mass term in the Lagrangian does not change the Palatini equation, and the effective nonlinear field equation of the massive vector field is obtained similarly to (4.29) by eliminating the torsion using (1.49) and (4.28):

\[
\frac{1}{\sqrt{g}} \partial_\alpha (\sqrt{g} F^\alpha \beta (\Gamma)) - m^2 A_\beta = \frac{4\pi}{c} J^\beta.
\] (4.32)

Using the static ansatz \( A_\mu = \{ \varphi(x^t), 0, 0, 0 \} \), and subsequently making a substitution \( \varphi = \frac{c^2}{\sqrt{G}} \sinh(\eta/2) \), we recast (4.32) into

\[
\Delta \eta = m_A^2 \sinh \eta.
\] (4.32a)
Then for the spherically symmetric solution we find \( \eta \sim \exp\left(-\frac{m_A r}{r}\right) \), i.e., the solution asymptotically (with \( r \to \infty \)) approaches the Yukawa potential.

Another interesting class of exact solutions can be obtained for the ansatz \( A_\mu = \{0, 0, A_2(x, t), 0\} \) [63]. Plugging this into the nonlinear equation (4.32), and making a substitution \( A_2 = \frac{e^2}{\sqrt{G}} \sin \phi \), we derive the well-known sin-Gordon equation

\[ \Box \phi - m_A^2 \sin \phi = 0, \]

where the 2-dimensional d’Alembert operator \( \Box = -\frac{1}{c^2} \frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x^2} \). This is a completely integrable system which admits soliton solutions.

Therefore, we established the geometrical nature of nonlinearity of sine-Gordon in the same way that the geometrical nature of nonlinearity of spinor fields of Ivanenko-Heisenberg was established.

### 4.2. Conformal invariance and spacetime torsion

In Sec. 1.3., when varying the Lagrangian \( L_{ECT} \) independently with respect to the metric \( g_{\mu \nu} \) and the linear connection \( \Gamma^\lambda_{\mu \nu} \) without imposing restrictions on these dynamical variables, we obtained the semi-metric theory with the connection (1.32) which is defined by the condition (1.31). This manifests a close relation between the torsion vector \( Q_\mu \) and the nonmetricity vector \( K_\mu \) that appears in the Weyl theory [5]:

\[ \nabla_\alpha g_{\mu \nu} = K_\alpha g_{\mu \nu}. \]

Weyl constructed his theory of gravity on the basis of conformal transformations of the interval \( ds^2 \). Weyl’s theory is invariant under the transformation of the metric of the following form:

\[ g_{\mu \nu} \longrightarrow g'_{\mu \nu} = e^{2\sigma} g_{\mu \nu}. \quad (4.33) \]

These are the so-called Weyl transformations, parametrized by an arbitrary scalar function \( \sigma = \sigma(x^\mu) \). Despite the mathematical elegance of this theory, a quite serious objection was found that strongly depreciated Weyl’s gravity model: the length standards and the clocks do not have an invariant meaning in this theory, and in particular, they “depend on the history”. Having failed to overcome this problem, the Weyl theory remained a beautiful geometrical model without direct applications to physics.

However, taking into account the significance of the conformal symmetry in physics, it is interesting to look for a possibility to include Weyl’s transformations (4.33) into a gravity theory.\(^5\)

\(^5\)This is once again demonstrated in another Dirac’s attempt to introduce the dilations by assuming the existence of different types of standards [68].
Here we outline a possible natural solution of the problem of the length standards by constructing the conformally invariant gravity theory of the Hilbert-Einstein type in the Riemann-Cartan spacetime \( U_4 \), [69].

The tetrad formalism proves to be most useful for our purposes, and we take the tetrad and the local Lorentz connection fields, \( h^a_{\mu} \) and \( \Gamma^a_{\mu} \), as the basic independent dynamic variables. The crucial point is as follows: we postulate that the conformal Weyl transformations \((4.33)\) of the metric are induced by

\[
\begin{align*}
    h^a_{\mu} &\rightarrow h'^a_{\mu} = e^\sigma h^a_{\mu}, \\
    \Gamma^a_{\mu} &\rightarrow \Gamma'^a_{\mu} = \Gamma^a_{\mu}.
\end{align*}
\]

(4.34)

Under these transformations, the tetrad is scaled whereas the Lorentz connection remains invariant, hence the Riemann-Cartan curvature is not changed too: \( R'^{ab}_{\mu\nu} = R^{ab}_{\mu\nu} \). However, the world connection is transformed in accordance with \((1.20)\) as

\[
\Gamma^{\alpha}_{\beta\mu} \rightarrow \Gamma'^{\alpha}_{\beta\mu} = \Gamma^{\alpha}_{\beta\mu} + \delta^{\alpha}_{\beta} \partial_{\mu} \sigma.
\]

(4.35)

These are the famous Einstein’s \( \lambda \)-transformations [70] which leave the curvature \((1.1)\) tensor \( R^a_{\beta\mu\nu}(\Gamma) \) (constructed from the world connection \( \Gamma^a_{\beta\mu} \)) invariant. It is important that the Riemann-Cartan geometrical structure is preserved by these transformations. This is evident from the fact that the metricity condition \((1.8)\) is not violated:

\[
\nabla_\alpha g_{\mu\nu} = 0 \rightarrow \nabla'_\alpha g'_{\mu\nu} = 0.
\]

In accordance with \((4.35)\), the torsion tensor \((1.3)\) is transformed as

\[
Q'^{\alpha}_{\mu\nu} \rightarrow Q'^{\alpha}_{\mu\nu} = Q^{\alpha}_{\mu\nu} + \delta^{\alpha}_{\mu} \partial_{\nu} \sigma.
\]

(4.36)

A remarkable fact is that \((4.36)\) affects the trace of torsion only

\[
Q_{\mu} \rightarrow Q'_{\mu} = Q_{\mu} - \frac{3}{2} \partial_{\mu} \sigma.
\]

The rest of the irreducible parts of the torsion \((1.13)\) are conformally invariant.

Now we will demonstrate that the group of local conformal transformations introduced above \((4.34)\) in the spacetime \( U_4 \) provides a natural basis for the description of the scale symmetry of both matter and gravitation.

At first, we consider the massless fermion field described by the Dirac action

\[
\frac{1}{2} \int d^4x \sqrt{|g|} L_D
\]

with the Lagrangian \((4.8)\) in which we put the rest mass equal zero \( m = 0 \). The fermion action is explicitly invariant with respect to scale transformations \((4.34)\) provided the Dirac field is scaled in accordance with its canonical dimension,

\[
\Psi \rightarrow \Psi' = e^{-\frac{3}{2} \sigma} \Psi.
\]

Let us note that the extra terms proportional to \( \partial_{\mu} \sigma \) occurring in the transformation \((4.34)\) are mutually cancelled, so there is no need to introduce additional
4.2. Conformal invariance and spacetime torsion

Weyl fields. It is worthwhile to note that such an intrinsic conformal invariance is actually provided by the torsion. Indeed, the variation of (4.8) with respect to $\Psi$ leads to the Dirac equation (4.13) in $U_4$ (with $m = 0$): $\gamma^\mu (D_\mu - Q_\mu) \Psi = 0$. This equation is explicitly conformally invariant due to the fact that the torsion trace $Q_\mu$ acts as the Weyl field.

Now we will show the possibility to construct a conformally invariant theory of gravity in $U_4$. Let us study the Hilbert-Einstein model with the action

$$S_{\text{conf}} = \int d^4x \sqrt{g} \varphi^2 R(\Gamma).$$

(4.37)

To provide the conformal invariance, we introduced an additional field $\varphi$ with the scaling behaviour under (4.34)

$$\varphi \rightarrow \varphi' = e^{-\sigma} \varphi,$$

in accordance with its canonical dimension. For the curvature scalar we obviously have $R'(\Gamma) = e^{-2\sigma} R(\Gamma)$, whereas $\sqrt{g'} = e^{4\sigma} \sqrt{g}$. The role of the scalar field $\varphi$ is (as in the similar cases in the Riemann geometry) to introduce a natural scale in the theory. Comparing (4.37) with the action (1.34) of the Einstein-Cartan model, we notice that the scalar field square $\varphi^2$ replaces $\frac{1}{16\kappa}$ and thus plays the role of a Brans-Dicke-type variable coupling function.

The field equations of the conformally invariant theory of gravity in $U_4$ can be obtained by variation of (4.37) either with respect to the metric and the torsion $g_{\mu\nu}, Q_{\lambda\mu\nu}$, or in the tetrad formalism with respect to the tetrad and the local Lorentz connection $h^a_\mu, \Gamma^a_{\mu b}$, $\Gamma^a_{\mu b}$. Both methods yield the same result. In the absence of the matter, we obtain the field equations

$$G_{\mu\nu}(\Gamma) = 0,$$

(4.38)

$$Q^\lambda_{\lambda\mu\nu} = \frac{1}{\varphi} \delta^\lambda_{[\mu} \partial_{\nu]} \varphi.$$

(4.39)

In addition, we have to vary the action with respect to the scalar field $\varphi$. However, the corresponding field equation, $R(\Gamma) = 0$, is redundant in view of (4.38). The equation (4.39) shows that the torsion is expressed in terms of its trace $Q_\mu$, for which the scalar field $\varphi$ serves as a “potential”. Substituting (4.39) into (4.38), we obtain

$$G_{\mu\nu} = -\frac{6}{\varphi^2} T_{\mu\nu},$$

(4.40)

where

$$T_{\mu\nu} = \partial_\mu \varphi \partial_\nu \varphi - \frac{1}{2} g_{\mu\nu} \partial_\alpha \varphi \partial^\alpha \varphi - \frac{1}{6} \left( \nabla_\mu \nabla_\nu - g_{\mu\nu} \Box \right) \varphi^2$$

is the so-called “improved” energy-momentum tensor of Callan-Coleman-Jackiw [71, 72] for the scalar field $\varphi$. Here $\Box := \nabla_\mu \nabla^\mu$ is the covariant d’Alembertian

---

6The Lagrangian of such a theory may also include all possible terms quadratic in the curvature $R^\alpha_{\beta\mu\nu}(\Gamma) R^\beta_{\alpha\gamma\delta}(\Gamma), R_{\mu\nu}(\Gamma) R^\mu_{\nu}(\Gamma), R^2(\Gamma), \text{etc.}$, which are all conformally invariant.
with respect to the Christoffel connection of the Riemannian space. The contraction of (4.40) yields the explicitly conformally invariant equation
\[ \Box \varphi - \frac{R}{6} \varphi = 0. \] (4.41)

Thus, we once again observe how the torsion provides the intrinsic conformal invariance of the gravitational theory. The “wrong” sign of the energy-momentum tensor \( T_{\mu\nu} \) of the scalar field in (4.40) manifests a non-physical nature of \( \varphi \) which is actually the Goldstone type field \([73, 74]\). Using the conformal freedom, we can choose the gauge \( \varphi = \text{const} \) (which means the choice of a certain length scale), and thus we arrive at the standard ECT.

Summarizing, we have demonstrated a natural way to introduce the Weyl transformations (4.33) and to construct the conformally invariant theory of gravity of the Hilbert-Einstein type in the Riemann-Cartan spacetime with torsion. Distinctive features of such an approach are the postulate of the tetrad scaling (4.34) as a source of conformal transformations (4.33) and the special role of the torsion trace \( Q_\mu \) as an effective Weyl field.

One of the most important advantages is the absence of the problem of the length scales, since \( D_\mu \eta_{ab} = 0 \) and, correspondingly, \( \nabla_\alpha g_{\mu\nu} = 0 \) (which are both explicitly conformally invariant relations). Consequently, length standards are conserved and there is no need in Dirac’s propositions of the two types of clocks \([68]\), etc.

The analysis of the possible relation of the torsion with conformal transformations clarifies the source of the vacuum production process of scalar particles described by the Klein-Gordon equation with the conformal coupling in the homogeneous and isotropic spacetimes (see Sec. 4.4.). Indeed, we find \( Q_\mu = 0 \) due to the Palatini equation (4.16) which results in the violation of the conformal invariance of the Dirac equation.

### 4.3. Pre-Friedman stage of Universe’s evolution and spacetime torsion

In this section, we turn to the study of the cosmological effects of ECT.

The theory of an expanding Universe – the Friedman cosmology based on Einstein’s equations (with the initial data) – is the most spectacular achievement of GR. Numerous important physical effects predicted on its basis were verified experimentally \([27]\), among them the relict microwave radiation, the cosmological red shift, etc. However, the problem of an initial state of the Universe remains unsolved within this framework. The presence of singularities in the general solution of Einstein’s equations \([86]\) is one of the central problems of the theory of gravitational interactions. One can expect that a further development of the fundamental principles of Einstein’s GR, that would allow to take into account the effects of the gravitational and matter fields which
are insignificant under the normal conditions (quantum and/or microstructural properties, e.g.), may provide possible ways to solve these problems [87].

In connection with this, one of the most important achievements of the Einstein-Cartan theory is the prediction of the avoidance (for an appropriate choice of the torsion tensor and of its spinning sources) of singularities in the cosmological models of the Universe. Such an opportunity is opened up when the torsion violates the Penrose-Hawking energy condition [86], which in the ECT takes the following form [7, 88]:

\[ T_{\mu\nu}^{\text{eff}} t^\mu t^\nu \geq \frac{1}{2} T_{\mu\nu}^{\text{eff}} g_{\mu\nu}, \]  

(4.42)

where \( t^\mu \) is an arbitrary time-like vector.

Such a violation is not unexpected. The Einstein-Cartan theory, as we will demonstrate in Chapters 5, 6, represents a special model of the gauge theory of gravity for the Poincaré group \( \mathcal{P}_{10} \) (a semidirect product of the Lorentz group \( SO(1,3) \) times the translation group \( T_4 \)). The non-compactness of the Lorentz group results in the indefiniteness of the energy sign, which leads to a possibility of violation of the condition (4.42).

It is important to note that in the regions occupied by the spinning matter, even in the case of chaotically oriented spins such that the mean value \( \langle S^\lambda_{\mu\nu} \rangle = 0 \) vanishes, the macroscopic spin effects are different from zero. This is due to the fact that in the equation (1.52) that has a microscopic meaning before the averaging procedure, there are terms quadratic in spin. Let us model the source of the gravitational field by the ideal spinning Weyssenhoff fluid (see Sec. 4.1.), consisting of fermions with the particle density \( \rho \), then \( S^2 = \frac{1}{2} S_{\mu\nu} S^{\mu\nu} = \hbar^2 \rho^2 \). Averaging yields \( \langle S^2 \rangle = \langle \frac{1}{2} S_{\mu\nu} S^{\mu\nu} \rangle = \hbar^2 \langle \rho^2 \rangle / 8 \), where the mean particle density \( \langle \rho \rangle \) is related to the energy density \( \varepsilon \) via \( \varepsilon \propto A_k \langle \rho \rangle^{1+k} \) for the equation of state \( p = k\varepsilon \) with \( 0 \leq k \leq 1 \), and \( A_k \) is some constant depending on the equation of state [120]. Then

\[ \langle S^2 \rangle = \frac{\hbar^2}{8 A_k^{1+k}} \varepsilon^{\frac{1}{1+k}}. \]  

(4.43)

If \( \langle S \rangle = 0 \), one can understand \( \langle S^2 \rangle \) as a squared dispersion of spin density distribution about the zero mean value,

\[ \langle \Delta S \rangle^2 = \langle (S) - S \rangle^2 = \langle S \rangle^2 - 2\langle S \rangle^2 + S^2 = \langle S^2 \rangle. \]

The existence of a nonzero \( \langle S^2 \rangle \) here is similar to the existence of a nonzero average kinetic molecular energy \( \langle p^2 / 2m \rangle \neq 0 \), when the average momentum vanishes, \( \langle p \rangle = 0 \), for the classical gas in equilibrium.

Let us consider the homogeneous isotropic cosmological model with the metric

\[ ds^2 = -c^2 dt^2 + R^2(t) \left[ \frac{dr^2}{1 - \mathcal{K} r^2} + r^2 d\Omega^2 \right], \quad \mathcal{K} = \pm 1, \ 0, \]  

(4.44)
with the Weyssenhoff spinning fluid as a matter source. The parameter $K = +1$ ($K = -1$) corresponds to the closed (open) Universe, and $K = 0$ to the quasi-Euclidean world [120]. Einstein’s equations (1.52) with the effective energy-momentum tensor (4.7) reduce to

$$\frac{\dot{R}^2(t)}{R^2(t)} + \frac{Kc^2}{R^2(t)} = \frac{8\pi G}{3c^2} \left( \varepsilon - \frac{8\pi G S^2}{c^2} \right), \quad (4.45a)$$

$$-2\frac{\dot{R}(t)}{R(t)} \frac{\dot{R}^2(t)}{R^2(t)} - \frac{Kc^2}{R^2(t)} = \frac{8\pi G}{c^2} \left( \frac{p - 8\pi G S^2}{c^2} \right). \quad (4.45b)$$

The covariant conservation law of the effective energy-momentum tensor $T^\mu_\mu = 0$ yields

$$\frac{\dot{\varepsilon}}{\varepsilon + p} = 3 \frac{\dot{R}}{R}. \quad (4.46a)$$

Therefore, for the case of the dust ($p = 0$), we have

$$\varepsilon = \frac{\varepsilon_0}{R^3}, \quad S^2 = \frac{S_0^2}{R^6}. \quad (4.46b)$$

where $\varepsilon_0$ and $S_0$ are the integration constants which describe the energy density and the particle density of the matter in the Universe at the moment of time when $R = 1$. When obtaining (4.46b), we took (4.43) into account.

Substituting (4.46a) into the Friedman equation (4.45a), we find

$$\dot{R}^2 = -Kc^2 + \frac{8\pi G \varepsilon_0}{3c^2 R} - \frac{64\pi^2 G^2 S_0^2}{3c^4 R^4}. \quad (4.47)$$

As compared to GR, we have an additional term proportional to $R^{-4}$ which arises due to the (Γ-S)-interaction. It can be interpreted as a centrifugal potential energy, and it is easy to see that this term leads to the absence of a cosmological singularity. In particular, for the quasi-Euclidean Universe with $K = 0$, the explicit solution of (4.47) reads

$$R(t) = \sqrt[3]{R_0^3_{\text{min}} + 6\pi G \varepsilon_0 t^2 / c^2}, \quad R_0^3_{\text{min}} = \frac{8\pi G S_0^2}{c^2 \varepsilon_0}. \quad (4.48)$$

To make an estimate, if we assume that the Universe is filled with $N \approx 10^{80}$ nucleons (hence the total mass of matter $\approx 10^{53}$ kg), we find for the minimal value of the cosmological scale factor $R_{\text{min}} \approx 10^{-28}$. The corresponding minimal radius of the Universe $\approx 10^{-2}$ m, and the maximal density $\approx 10^{57}$ kg/m$^3$ [97].

Another estimate can be obtained if we replace the nucleon as an elementary unit of matter with a hypothetical “fundamental atom” – the Planckon [121]. The latter is a black hole with the mass of $m_{\text{pl}} \approx 10^{-8}$ kg and the spin $\frac{1}{2}\hbar$.

---

7On the right-hand side of the equations, the averaged quantities are understood.
4.3. Pre-Friedman stage of Universe’s evolution and spacetime torsion

Since the Compton wave length of particles $\lambda = \hbar/mc$ cannot be greater than the gravitational radius of a configuration, the Planckon is a smallest possible black hole [122]. Then, assuming that at an initial moment of time the Universe (as the interior of a collapsing star) consisted of completely spin-polarized Fermi fluid formed by $10^{61}$ Planckons (to keep the total mass $\approx 10^{53}$ kg), we find for the minimal radius of the Universe $\approx 10^{-15}$ m.

Since the spin is a vector and thus it obviously defines a preferred direction in space, thereby breaking the spatial isotropy, it is perhaps more realistic to consider anisotropic spatially homogeneous models. Let us study the Bianchi I model with the metric

$$\text{ds}^2 = -c^2 dt^2 + a^2(t)(dx^2 + dy^2) + b^2(t)dz^2,$$

where the two scale factors $a = a(t)$ and $b = b(t)$ describe a possible difference of the cosmological dynamics along a preferred axis ($z$-coordinate) and in the orthogonal plane.

The effective Einstein equations for this metric read

$$\frac{\dot{a}^2}{a^2} + 2\frac{\dot{a}\dot{b}}{ab} = \kappa\varepsilon\text{eff},$$

$$-\frac{\ddot{a}}{a} - \frac{\ddot{b}}{b} = \kappa p\text{eff},$$

$$\frac{\ddot{a}}{a} + \frac{\ddot{b}}{b} - \frac{\dot{a}\dot{b}}{ab} = 0.$$

Introducing the new variables

$$R^3 := a^2 b, \quad \sigma := \frac{\dot{a}}{a} - \frac{\dot{b}}{b},$$

we recast this system into

$$3\frac{\dot{R}^2(t)}{R^2(t)} = \frac{8\pi G}{c^2} \left( \varepsilon - \frac{8\pi G}{c^2} S^2 \right) + \frac{1}{3}\sigma^2,$$

$$-2\frac{\dot{R}(t)}{R(t)} \frac{\dot{R}^2(t)}{R^2(t)} - \frac{Kc^2}{R^2(t)} = \frac{8\pi G}{c^2} \left( p - \frac{8\pi G}{c^2} S^2 \right) + \frac{1}{3}\sigma^2,$$

$$\dot{\sigma} + 3\frac{\dot{R}(t)}{R(t)} \sigma = 0.$$

The last equation is easily integrated and yields the anisotropy expansion function (the difference of the Hubble functions):

$$\sigma = \frac{\sigma_0}{R^3}.$$

The equation of motion of spin (4.5) is also straightforwardly integrated to give

$$S = \frac{S_0}{a^2 b} = \frac{S_0}{R^3}.$$

Substituting all this into the Friedman equation, for the Weyssenhoff dust ($p = 0$) as a source of gravitational field, we find a modified version of (4.47)

$$\dot{R}^2 = \frac{8\pi G\varepsilon_0}{3c^2 R} \frac{\mu_0}{R^4},$$
where we introduced a combination of the integration constants
\[ \mu_0 = \frac{64\pi^2 G^2 S_0^2}{3c^4} - \frac{\sigma_0^2}{9} \]

The qualitative behaviour of a cosmological solution depends crucially on the balance between the spin and anisotropy: in simple terms, on the value of the constant \( \mu_0 \). For \( \mu_0 = 0 \) one finds the usual singular Friedman solution for the dust. For \( \mu_0 < 0 \), we also have a singular solution. However when \( \mu_0 > 0 \), the solution does not have cosmological singularities [123]:

\[
\begin{align*}
a &= \sqrt{\frac{3c^2 \mu_0}{8\pi G \varepsilon_0}} + \frac{6\pi G \varepsilon_0}{c^2} r^2 \times \exp\left( \frac{2\sigma_0}{9\sqrt{\mu_0}} \arctan \left( \frac{4\pi G \varepsilon_0 t}{c^2 \sqrt{\mu_0}} \right) \right), \\
b &= \sqrt{\frac{3c^2 \mu_0}{8\pi G \varepsilon_0}} + \frac{6\pi G \varepsilon_0}{c^2} r^2 \times \exp\left( -\frac{4\sigma_0}{9\sqrt{\mu_0}} \arctan \left( \frac{4\pi G \varepsilon_0 t}{c^2 \sqrt{\mu_0}} \right) \right).
\end{align*}
\]

Besides the elimination of singularities in the cosmological solutions, the spin terms produce another interesting effect by stabilizing the matter distribution, so that for the closed Universes (only in this case) there exist static solutions of the Friedman equations (4.45). For the dust, such a Universe has the size \( \approx \lambda_f \) of a Compton wave length of a fermion with the mass \( m_f \), and the matter density is about \( \rho_{cr} = m_f^2 c^4 / G \hbar^2 \) (\( \approx 10^{57} \) kg/m\(^3\) for a nucleon), [125]. Such a static solution

\[ ds^2 = -c^2 dt^2 + R_0^2 \left[ \frac{dr^2}{1 - r^2} + r^2 d\Omega^2 \right] \quad (4.49) \]

coincides with the metric of Einstein’s Universe [120]. However, in contrast to GR, here the matter density does not vanish, and the cosmological \( \Lambda \)-term is absent.

The dynamics of the local small perturbations is an important issue in cosmology. Within the framework of GR, this problem was solved for the first time by E.M. Lifshitz [124]. One can show that on the background of the static solution (4.49), the density perturbation \( \delta \varepsilon \) develops in the two exponential modes

\[ \delta \varepsilon = \frac{c^2 \omega^2}{120 \epsilon G} \left[ C_1 \exp \left( \frac{\sqrt{5}}{3} \omega t \right) + C_2 \exp \left( -\frac{\sqrt{5}}{3} \omega t \right) \right], \]

where \( \omega = 2\pi c / \lambda \), and \( \lambda \) is a perturbation wave length; \( C_1 \) and \( C_2 \) are the integration constants. The development of the exponentially growing and decaying modes is typical for any exact static solutions.

Another important type of perturbations is the occurrence of a global homogeneous isotropic motion of matter, that does not violate the spatial symmetry of the model.

The static solution (4.49) obtained in the framework of the Einstein-Cartan theory differs from the Einstein solution of GR with the \( \Lambda \)-term in that it turns out to be stable with respect to this type of perturbations for the equation
4.3. Pre-Friedman stage of Universe’s evolution and spacetime torsion

Of state $p = k \varepsilon$, with $0 \leq k < 1$. Indeed, small perturbations $\delta R$ satisfy the equation

$$\ddot{\delta R} + \omega^2(k)\delta R = 0, \quad \omega^2(k) = \frac{4\pi G \varepsilon_0}{c^2} (-3k^2 + 2k + 1), \quad (4.50)$$

where $\varepsilon_0$ is the constant energy density of the solution, and $\omega^2(k) > 0$ for $0 \leq k < 1$.

These properties allow one to view such a static solution as the initial and the final stages of the observable non-stationary picture of the Universe [125]. In such a model, the evolution of the Universe consists of the following stages (Fig. 4.2): the time interval $t_0t_1$ describes the initial static state, when small perturbations are quickly damped; at the moment $t_1'$ a large fluctuation occurs which is sufficient to form for the Universe observed today; the next interval $t_1't_2$ covers the initial stage of expansion with the dominating torsion effects; the interval $t_2t_3$ is a stage of the “standard” cosmology when the torsion effects are negligible and the solution is close to the Friedman one; finally, during the interval $t_3t$ the Universe returns back to the static state. If the energy of the initial perturbation does not manage to be spent on dissipative phenomena during the expansion and contraction phases, the Universe contracts to a smaller size than the original one $R_0$ and returns to the initial scale after several oscillations (the branch I), otherwise the damping is smooth (the branch II).

No matter how large the initial perturbation is, the expansion necessarily is followed by a contraction (the closed Universe), and the contraction does not reach a singularity due to the short-range spin term ($\sim S^2$) in the Friedman equations.

The issue of a transition from the static stage ($t_0t_1$) to the dynamical one ($t_1't_2$) is not completely clear in this model. In order to answer this question, we consider in detail how the small homogeneous isotropic perturbations behave. Up to now, we neglected the dissipative phenomena (i.e., the increase of the entropy) and, as a result, we obtained the small harmonic oscillations (4.50) of the scale factor around its equilibrium value $R_0$. However, it is easy to see that the second law of thermodynamics forbids the exact repetition of the previous
cycles. Moreover, the increase of the entropy $s$, $ds/dt > 0$ (both at the expansion stage, and at the contraction stage) turns out to be a natural mechanism that drives the initially static Universe into an observable strongly non-stationary mode. Indeed, the condition that the scale factor reaches an extremal value $R_{\text{extr}}$ is given by the Friedman equation (4.45a):

$$R_{\text{extr}}^2 \left\{ \varepsilon (R_{\text{extr}}, s) - \frac{8\pi G}{c^2} S^2 (R_{\text{extr}}) \right\} = \frac{3c^4 \mathcal{K}}{8\pi G}.$$  

(4.51)

Evaluating the logarithmic derivative of (4.51), after some calculations, we find

$$\frac{dR_{\text{max}}}{ds} = \pm \frac{4\pi G}{3c^2} \left| R \left( R_{\text{max}} \right) \right| \frac{\partial \varepsilon}{\partial s}, \quad \frac{\partial \varepsilon}{\partial s} > 0.$$  

(4.52)

The equation (4.52) demonstrates that the maximum of the scale factor $R_{\text{max}}$ in the repeating cycles increases due to the irreversible phenomena such as friction (for example, the viscosity of matter). This ultimately results in the natural transition from the static pre-Friedman stage ($t_0 t_1$) to the dynamical Friedman stage ($t'_1 t_2$).

Summarizing the analysis of the influence of spin on the cosmological evolution of the Universe in the framework of the Einstein-Cartan theory, we can make the following conclusion.

Already the simplest model of a spinning matter in the form of the Weyssenhoff fluid demonstrates that the torsion brings in a new essential bit to the understanding of the spacetime structure of the initial state of the Universe. However, as was noted above, the Weyssenhoff fluid represents a quasi-classical limit of a real quantum matter with spin. Thus, the consistent analysis of cosmology in the Einstein-Cartan model should be performed also for the Dirac spinor matter. Although the classical spinor fields do not eliminate\(^8\) singularities [126], the quantum treatment of spinors [127] essentially confirms the correctness of the results obtained for the Weyssenhoff fluid.

### 4.4. Production of scalar particles by cosmological torsion field

One of the most important issues of the modern relativistic astrophysics is the problem of the gravitational collapse, i.e., the study of the evolution of massive gravitational systems (both, isolated objects and the Universe on the whole) when all the internal resources of energy are exhausted. The analysis of such a process in the classical approximation is based on Einstein’s equations (the classical gravitation and the classical matter). The classical picture of the

---

\(^8\)Due to space-like nature of the pseudovector of the spin, $(\mathcal{F} \gamma^\mu \gamma_5 \Psi)^2 \geq 0$. 
gravitational collapse is based on these equations, for which all the key evolution scenarios are established and classified: the equilibrium configurations, the cases when the contraction stops under the gravitational radius, and actually the case of a collapse when the matter is uncontrollably contracting under the influence of gravitational forces. During the contraction, the matter can become extremely dense (reaching the nuclear density $10^{17}$ kg/m$^3$ and greater). Strictly speaking, at this stage it is necessary to use the Einstein equation with the source on the right-hand described by the averaged energy-momentum tensor of the quantum matter\footnote{The averaged energy-momentum tensor of matter is usually understood as the result of the quantum averaging over the pure states, for example, over the space of occupation numbers. But because of the fact that the system during the evolution can come to a thermodynamic equilibrium, it seems to be necessary to make also a statistical averaging over the equilibrium distribution \cite{75} (see Sec. 7.2.).}. At the final stages of the gravitational collapse at the density greater than $10^{99}$ kg/m$^3$ and the distance smaller than $10^{-35}$ m, it is necessary to replace the classical equations by their quantum counterparts.

However, due to the great complexity of the problem of construction of the full quantum version of Einstein’s equations (see Chapter 7), it is reasonable to consider the semi-consistent problem as a first approximation, when the gravity is treated classically and the matter - quantum-mechanically. The expediency of such approximation is obvious since already in this picture one can expect new features in the evolution of the gravitating configurations as compared to existing processes at a classical level.

The external classical non-stationary gravitational field may lead to the instability of vacuum of the quantized physical fields, i.e., to the production of the respective particles. As we show here, the torsion (within the framework of ECT) brings the new features to the particle generation process.

The classical dynamics of a massless scalar field $\phi$ will be described by the Klein-Gordon equation with the conformal coupling (4.41), with the Riemann-Cartan scalar curvature $R(\Gamma)$ in place of $R$. The action for the scalar field has the following form:

$$ S = \frac{1}{c} \int d^4x \sqrt{g} \mathcal{L}_\phi, \quad \mathcal{L}_\phi = -\frac{\hbar c^2}{2} \left( g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi + \frac{R(\Gamma)}{6} \phi^2 \right). $$

(4.53)

We rewrite the metric (4.44) of the quasi-Euclidean (with $K = 0$) homogeneous isotropic Universe as

$$ ds^2 = a^2(-d\eta^2 + \gamma_{ik} dx^i dx^k), $$

(4.54)

introducing the conformal time $c dt/R(t) = d\eta$ and denoting the new scale factor $a(\eta) = R(t(\eta))$. The three-dimensional Euclidean metric $\gamma_{ik}$ is not necessarily written in Cartesian coordinates. Furthermore, we assume that the evolution of the Universe is described by the cosmological solution (4.48).

We model the source of the external torsion field as a massless fermionic fluid described by the Dirac equation (4.18) (with $m_\Psi = 0$) with the spin directed
along the axis $z$. The torsion is then represented by the pseudotrace that has the form $Q_\mu = (0, 0, 0, Q)$, where $Q = \kappa c^2 \frac{\hbar}{\rho}$ and $\rho$ is the particle density of fermions. The conservation law of the number of particles in the metric (4.54) yields $Qa^2 = \frac{\hbar}{2} \kappa c \rho_0$ with the integration constant $\rho_0$ that gives the fermion particle density at the moment when $a = 1$.

Introducing a new variable for the scalar field $\chi = a \phi$, we get the Lagrange function

$$L_\chi = \frac{\hbar c}{2} \int d^3 x \sqrt{\gamma} \left\{ \dot{\chi}^2 + \chi \Delta \chi - \frac{Q_0^2}{a^4} \chi^2 \right\}. \tag{4.55}$$

Here the dot denotes the derivative with respect to the conformal time $\eta$, the $3$-metric $\gamma_{ik}$ determines the integration measure $\sqrt{\gamma} = (\text{det} \gamma_{ik})^{1/2}$ and the covariant Laplacian $\Delta = \gamma^{ik} \nabla_i \nabla_k$, and $Q_0 = \kappa c^2 \frac{\hbar}{2} \rho_0$ is the parameter with the dimension of $1/\text{length}$ which is proportional to the particle density of fermions that create the spacetime torsion.

Let us decompose the redefined scalar field

$$\chi(x, \eta) = \sum_n \chi_n(\eta) Z_n(x) \tag{4.56}$$

with respect to the complete orthonormal system of eigenfunctions $Z_n(x)$

$$\Delta Z_n(x) = -\omega_n^2 Z_n(x),$$

$$\int d^3 x \sqrt{\gamma} Z_n(x) Z_m(x) = \delta_{nm}, \quad \left\{ \right\} \tag{4.57}$$

of the $3$-dimensional Laplace operator, where the collective index $n$ labels the whole set of eigenvalues $\omega_n^2$. For the quasi-Euclidean model, the spectrum of the Laplacian is continuous, so the sum over the eigenmodes should be actually understood as an integral with an appropriate choice of the measure.

Substituting (4.56) and using (4.57), we rewrite the Lagrange function (4.55) in terms of an infinite set of discrete degrees of freedom of the scalar field $\chi_n(\eta)$:

$$L_\chi = \frac{\hbar c}{2} \sum_n \left\{ \dot{\chi}_n^2 - \omega_n^2 \chi_n^2 - \frac{Q_0^2}{a^4} \chi_n^2 \right\}. \tag{4.58}$$

The corresponding equation of motion

$$\ddot{\chi}_n(\eta) + \Omega_n^2(\eta) \chi_n(\eta) = 0, \quad \Omega_n^2(\eta) = \omega_n^2 + \frac{Q_0^2}{a^4(\eta)} \tag{4.58}$$

represents a discrete non-Riemannian version of (4.41).

Introducing as usual the canonical momenta\(^{10}\), $p_n = \partial L_\chi / \partial (c \dot{\chi}_n) = \hbar \dot{\chi}_n$ we construct the Hamilton function

$$H_\chi = \sum_n p_n c \chi_n - L_\chi = \frac{\hbar c}{2} \sum_n \left\{ (p_n/\hbar)^2 + \omega_n^2 \chi_n^2 + \frac{Q_0^2}{a^4} \chi_n^2 \right\}. \tag{4.58}$$

\(^{10}\)In $\dot{\chi}_n = d\chi_n/d\eta$ the variable $\eta$ has the dimension of length, so we need a factor $c$ to obtain the velocity $c \dot{\chi}_n$ with the correct dimension.
The same Hamiltonian can be alternatively derived from the metrical energy-momentum tensor \[71, 72\]

\[
T_{\mu\nu} = \hbar c \left\{ \partial_\mu \varphi \partial_\nu \varphi - \frac{1}{2} g_{\mu\nu} g^{\alpha\beta} \partial_\alpha \varphi \partial_\beta \varphi + \frac{1}{6} \left( G_{\mu\nu} \varphi^2 - \frac{1}{2} \nabla_\mu \nabla_\nu \varphi^2 + g_{\mu\nu} \Box \varphi^2 \right) + \hat{Q}_\mu \hat{Q}_\nu \varphi^2 - \frac{1}{2} g_{\mu\nu} g^{\alpha\beta} \hat{Q}_\alpha \hat{Q}_\beta \varphi^2 \right\}. \tag{4.59}
\]

The Hamilton function is then recovered as the integral

\[
H_\chi = a^2 \int d^3 x \sqrt{-g} T_{00}. \tag{4.60}
\]

The scalar field is quantized in a standard way by treating \(\chi_n(\eta)\) and the canonically conjugated momenta \(p_n(\eta)\) as operators satisfying the equal-time commutation relations

\[
[\chi_m(\eta), \chi_n(\eta)] = 0, \quad [p_m(\eta), p_n(\eta)] = 0, \quad [\chi_n(\eta), p_m(\eta)] = i \hbar \delta_{nm}. \tag{4.61}
\]

To introduce the particle interpretation, we decompose the operator \(\chi_n(\eta)\) into the two complex-conjugate solutions \((u_n, u_n^*)\) of the equations (4.58):

\[
\chi_n(\eta) = a_n u_n(\eta) + a_n^* u_n^*(\eta). \tag{4.62}
\]

The basis functions \((u_n, u_n^*)\) are chosen so that they satisfy the normalization conditions

\[
u_n \dot{u}_n^* - \dot{u}_n u_n^* = i. \tag{4.63}
\]

The creation \(a_n^*\) and annihilation \(a_n\) operators satisfy

\[
a_n, a_m] = 0, \quad [a_n^*, a_m^*] = 0, \quad [a_n, a_m^*] = \delta_{nm}.
\]

These standard commutation relations follow from (4.61)-(4.63).

A detailed analysis of the particle interpretation of the system can be found in [61]. Inserting (4.62) into the Hamilton function, we find

\[
H_\chi = \frac{\hbar c}{2} \sum_n \Omega_n(\eta) \left\{ E_n(a_n a_n^* + a_n^* a_n) + F_n(a_n)^2 + F_n^*(a_n^*)^2 \right\}, \tag{4.64}
\]

where we denoted the dimensionless functions

\[
E_n(\eta) = \frac{\vert \dot{u}_n^* \vert^2}{\Omega_n(\eta)} + \Omega_n(\eta) \vert u_n \vert^2, \quad F_n(\eta) = \frac{\vert \dot{u}_n \vert^2}{\Omega_n(\eta)} + \Omega_n(\eta) \vert u_n \vert^2. \tag{4.65}
\]

Note that \((E_n)^2 - |F_n|^2 = 1\) in view of (4.63).

We choose the initial conditions \(\{u_n, u_n^*, \dot{u}_n, \dot{u}_n^*\}\) at an arbitrary moment of time \(\eta = \eta_0\) so that to satisfy (4.63) and to make the Hamilton function (4.64) diagonal by setting \(E_n(\eta_0) = 1\) and \(F_n(\eta_0) = 0\):

\[
u_n(\eta_0) = u_n^*(\eta_0) = \frac{1}{\sqrt{2 \Omega_n(\eta_0)}}, \quad \dot{u}_n(\eta_0) = -\dot{u}_n^*(\eta_0) = -i \sqrt{\frac{\Omega_n(\eta_0)}{2}}.
\]
The operators $a_n$ and $a_n^*$ are the creation and annihilation operators of scalar particles at the initial moment $\eta_0$. This interpretation is also preserved for any $\eta$ if $\Omega_n$ is constant, then the Hamiltonian remains diagonal at all times. However, for the time-dependent $\Omega_n(\eta)$ the Hamilton function (4.64) ceases to be diagonal in terms of $a_n$ and $a_n^*$ at any time $\eta > \eta_0$. This means that the physical vacuum is unstable and hence the notion of a particle is not constant but depends on time, which is manifest in the production of scalar particles.

In GR this effect is absent in the conformally-flat space-time (4.54) for the model under consideration due to the conformal invariance of the scalar field theory. In the ECT, the particle creation effect is nontrivial which is explained by the fact that the fermion torsion sources violate the conformal invariance of the field $\varphi$, as we noted in Sec. 4.2.

Thus, at an initial time $\eta = \eta_0$, we have the particle interpretation of $\varphi$. The field energy in $n$-mode is $\varepsilon_n = \frac{\hbar}{2c} \Omega_n(\eta_0) (a_n a_n^* + a_n^* a_n)$. In order to find the particle interpretation for an arbitrary time, we use the diagonalization method of the Hamilton function (4.64). In accordance with this method, it is necessary to introduce the new (time-dependent) creation and annihilation operators $A_n^*(\eta)$ and $A_n(\eta)$ by

$$a_n = \alpha_n(\eta) A_n(\eta) + \beta_n(\eta) A_n^*(\eta), \quad a_n^* = \alpha_n^*(\eta) A_n^*(\eta) + \beta_n^*(\eta) A_n(\eta), \quad (4.66a)$$

where the coefficients $\alpha_n(\eta)$ and $\beta_n(\eta)$ satisfy the conditions

$$|\alpha_n(\eta)|^2 - |\beta_n(\eta)|^2 = 1, \quad F_n \beta_n = (1 - E_n) \alpha_n. \quad (4.66b)$$

In terms of these operators, the Hamilton function (4.64) has diagonal form

$$H_\chi = \frac{\hbar c}{2} \sum_n \Omega_n(\eta) \{ A_n^*(\eta) A_n(\eta) + A_n(\eta) A_n^*(\eta) \}.$$ 

At any given time, the field particles are determined by the time-dependent creation $A_n^*(\eta)$ and annihilation $A_n(\eta)$ operators. By construction, we have $\alpha_n(\eta_0) = 1, \beta_n(\eta_0) = 0$ and hence the operators coincide $A_n(\eta_0) = a_n$, and $A_n^*(\eta_0) = a_n^*$ at the initial moment of time $\eta_0$.

The instantaneous physical vacuum $|0_\eta\rangle$ is defined as the quantum state that satisfies $A_n(\eta)|0_\eta\rangle = 0$. For an arbitrary operator $\Phi$ we introduce the normal ordering

$$N_\eta(\Phi) = \Phi - \langle 0_\eta | \Phi | 0_\eta \rangle.$$ 

This obviously depends on time: one subtracts the vacuum average with respect to the instantaneous vacuum, thus removing the (usually diverging) contribution of the vacuum excitations $a_\eta$. In order to find the physical effects due to the particle production, one needs to compute vacuum averages with respect to the initial vacuum:

$$\langle 0_{\eta_0} \mid N_\eta(\Phi) \mid 0_{\eta_0} \rangle = \langle 0_{\eta_0} \mid \Phi \mid 0_{\eta_0} \rangle - \langle 0_\eta \mid \Phi \mid 0_\eta \rangle.$$
Particularly important are the operators of number of particles $A_n^*(\eta) A_n(\eta)$ (in the $n$-th mode) and the operators of the energy-momentum $T_{i,k}(x)$. The latter is given explicitly in (4.59). Direct computation of the vacuum averages of the normal ordered operators yields:

\[
\begin{align*}
    n(\eta) & = \langle 0_{\eta_0}| N_\eta (A_n(\eta) A_n^*(\eta))|0_{\eta_0}\rangle = \sum_n |\beta_n|^2, \\
    \mathcal{E}(\eta) & = - \langle 0_{\eta_0}| \mathcal{N}_\eta (T_{1,0})|0_{\eta_0}\rangle = \frac{\hbar c}{a^2(\eta)} \sum_n \Omega_n(\eta)|\beta_n|^2, \\
    \mathcal{P}_1(\eta) & = \langle 0_{\eta_0}| \mathcal{N}_\eta (T_{1,1})|0_{\eta_0}\rangle = \frac{1}{3} \{ \mathcal{E}(\eta) - \gamma(\eta) \}, \\
    \mathcal{P}_2(\eta) & = \langle 0_{\eta_0}| \mathcal{N}_\eta (T_{2,2})|0_{\eta_0}\rangle = \frac{1}{3} \{ \mathcal{E}(\eta) - \gamma(\eta) \}, \\
    \mathcal{P}_3(\eta) & = \langle 0_{\eta_0}| \mathcal{N}_\eta (T_{3,3})|0_{\eta_0}\rangle = \frac{1}{3} \{ \mathcal{E}(\eta) + 2\gamma(\eta) \},
\end{align*}
\] (4.67)

where

\[
\gamma(\eta) = \frac{\hbar c Q_0^2}{a^8(\eta)} \sum_n \left\{ |u_n|^2 - \frac{1}{2\Omega_n(\eta)} \right\}.
\]

One can verify that for the off-diagonal components of the energy-momentum tensor, the vacuum averages vanish, $\langle 0_{\eta_0}| \mathcal{N}_\eta (T_{i,k})|0_{\eta_0}\rangle = 0$ with $i \neq k$. The total number of particles $n(\eta)$ created from vacuum by the torsion field, as well as the energy density $\mathcal{E}(\eta)$ and the anisotropic pressure $\mathcal{P}_1(\eta), \mathcal{P}_2(\eta), \mathcal{P}_3(\eta)$ of the created matter are determined by the solutions $u_n(\eta)$ of the evolution equation (4.58). From (4.67), one can obtain the equation of state for the created matter

\[
\mathcal{E} = \mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3.
\]

It is impossible to solve the oscillator equation (4.58) exactly for the time-dependent frequency $\Omega_n(\eta)$ with an arbitrary cosmological evolution $a = a(\eta)$. So, in order to obtain numerical estimates for $(n, \mathcal{E}, \mathcal{P}_1)$, one needs a perturbation scheme to construct approximate solutions of the equation (4.58) with the given initial conditions. As a first step, we recast the differential equation (4.58) into the equivalent integral equation of Volterra type:

\[
u_n(\eta) = e^{-i\Omega_n(\eta_0)(\eta - \eta_0)} + \int_{\eta_0}^{\eta} d\xi G(\eta - \xi) \Delta Q^2(\xi) u_n(\xi),
\]

where $\Delta Q^2(\eta) = Q^2(\eta_0) - Q^2(\eta) = Q_0^2 \left\{ 1/a^4(\eta_0) - 1/a^4(\eta) \right\}$, and $G(\eta - \xi)$ is the Green function that satisfies

\[
\ddot{G}(\eta - \xi) + \Omega^2_n(\eta_0)G(\eta - \xi) = \delta(\eta - \xi).
\]
The resulting integral equation is then solved iteratively, and the solution reads

\[
\begin{align*}
  u_n(\eta) &= u_n^{(0)}(\eta) + u_n^{(1)}(\eta) + \cdots + u_n^{(m)}(\eta) + \ldots, \\
  u_n^{(0)}(\eta) &= e^{-i\Omega_n(\eta_0)(\eta-\eta_0)} \sqrt{2\Omega_n(\eta_0)} , \\
  u_n^{(m)}(\eta) &= \int_{\eta_0}^\eta d\xi G(\eta-\xi) \Delta Q^2(\xi) u_n^{(m-1)}(\xi).
\end{align*}
\]

We take \( G(\eta-\xi) = \frac{\sin(\Omega_n(\eta_0)(\eta-\xi))}{\Omega_n(\eta_0)} \) as the explicit form of the Green function.

We choose the initial time \( \eta_0 \) as the moment when the scale factor \( a(\eta) \) takes the minimal value \( a(\eta_0) = a_{\text{min}} = R_{\text{min}} \) (see Sec. 4.3.). Limiting ourselves to the first two terms in (4.68), from (4.67), (4.68) we can obtain the numerical estimates for the particle number, the energy and the pressure of the created matter. In the lowest order we find

\[
\begin{align*}
  n(\eta) &= \sum_n \frac{(\Omega_n(\eta) - \Omega_n(\eta_0))^2}{2\Omega_n(\eta)\Omega_n(\eta_0)}, \\
  \mathcal{E}(\eta) &= \frac{\hbar c}{a^4(\eta)} \sum_n \frac{(\Omega_n(\eta) - \Omega_n(\eta_0))^2}{2\Omega_n(\eta_0)}, \\
  \gamma(\eta) &= \frac{\hbar c Q^2_0}{a^8(\eta)} \sum_n \frac{\Omega_n(\eta) - \Omega_n(\eta_0)}{2\Omega_n(\eta)\Omega_n(\eta_0)}.
\end{align*}
\]

The details of the computations can be found in [78]. Evaluation of the sums in (4.69) (more exactly, of the integrals) is a nontrivial task. Here we summarize the qualitative results.

1. The maximal number of created particles corresponds to the scale factor \( a_0 \approx \sqrt{2}a_{\text{min}} \) and the density of created particles is equal to \( \approx \frac{Q^3_0}{a_{\text{min}}^3} \), that is directly related to the density of fermions generating the spacetime torsion.

2. At the large distances (far from the Planck region \( a > a_0 \)), the intensity of the scalar particle production by the torsion field drops. Conversely, due to the high power of the inverse scale factor in the expressions for \( n(\eta) \), \( \mathcal{E}(\eta) \), at the small distances the process of scalar particle production by the torsion field dominates over the similar processes of the matter particle production in GR.

3. The account of the back reaction of the created particles on the classical evolution of the isotropic Universe should lead to its anisotropization in the early epochs due to an extra contribution \( \gamma(\eta) \) making the pressure anisotropic. The estimate of the maximal anisotropization degree yields the times close to \( a_0 \). At the large distances (late epochs with \( a \to \infty \)), \( \gamma(\eta) \) quickly decreases to zero. As a result, during Universe’s evolution, the isotropy is recovered.
5

Kinematics of gauge theories of gravity

5.1. Special aspects of gauge approach in gravitation

Mathematical beauty of the classical GR as a dynamical theory of a three-dimensional spatial geometry, as well as its agreement with the present-day experimental data in cosmology, astrophysics and planetary astronomy, allow one to view GR as a fair description of the macroscopic gravitational phenomena. However, the existence of essential difficulties in it (such as the quantization problem [92] and the singularity problem [86]) at small distances apparently manifests its limited validity in the microworld.

Recently numerous attempts were made to formulate the microscopic gravity theory within the framework of a consistent gauge-theoretic approach. This was essentially motivated by the successful construction of renormalizable unified gauge models of the weak and electromagnetic, as well as of the strong interactions. On the other hand, construction of the microscopic gravity theory appears to be a necessary step to establish the geometrical structure of supergravity [93] as the theory of the local supersymmetry.

The gravity theory, understood in a broad sense as the geometrodynamics\(^1\) of the spacetime, is much richer and more complex in the differential-geometrical aspects than the gauge theories of internal symmetries. As a result, there exists a variety of approaches and interpretations of the gravitational field as a gauge one, and there is no clear understanding of the answer to the key question: is gravity a gauge theory and in which sense? This question includes the three

\(^1\)That is the dynamics of the geometrical structure of spacetime.
aspects [97]: 1) What is the gauge group of the theory; what is the role of the general coordinate transformations? 2) What are the relevant gauge fields; what is their relation with the geometrical objects on the spacetime manifold; in particular, what is the status of the metric? 3) What is theory’s dynamics; is it determined by the requirements of the gauge invariance?

In view of these issues, it is reasonable to consider in detail the difficulties arising in the development of a consistent gauge approach to the gravity theory. To begin with, we have to clarify the terminology, because different researchers understand the meaning of expressions “gauge field” and “gauge theory” in different ways. We define the gauge theory as a physical theory in which the fundamental dynamical variable—the gauge field—is a connection on a principal bundle. Among such theories, we distinguish the consistent models that do not contain any additional postulates of the non-gauge type from the inconsistent theories containing such assumptions. Besides that, certain theories that are not covered by this definition, still have some features of the gauge theories, such as the presence of arbitrary functions in the description of the fields and the existence of constraints; we call them pseudo-gauge theories. There are many examples: the theory of massless fields of arbitrary spin; generally covariant models of material fields in a curved space; the simple supergravity; the general relativity theory [35]. Quantization of the gauge and pseudo-gauge theories takes into account their common feature of the invariance with respect to an infinite group of local transformations, and this similarity brings in a terminological confusion, which results in the unfortunate statements, such as “GR is the gauge theory of the group of general coordinate transformations”.

The issue of finding the gauge group for the gravity theory is one of the most complicated ones, since the gravity is related to spacetime symmetries, in contrast to the gauge fields of “internal” symmetries. However, being well-defined as global symmetries in the flat Minkowski space, the “localized” spacetime symmetries become general coordinate transformations, losing their original features. Moreover, there is a difficulty of interpretation of the “local” translations which are usually identified with the infinitesimal general coordinate transformations [12, 13]. The corresponding “gauge fields” – the tetrads – do not have a standard (Yang-Mills) transformation law, since they are not connections.

As it is clear now, all these and related inconsistencies arise from the confusion of the gauge structure with the coordinate invariance. However, they should be carefully distinguished, because the group of general coordinate transformations does not have a direct relation to the gauge theory of gravity [96]. If one views the general coordinate transformations simply as the maps between charts covering the spacetime manifold, they obviously do not carry any dynamical meaning. The invariance with respect to such passive transformations is a common property of all covariant physical theories, not only of gravity. More-

---

2 An example of the latter is an ad hoc postulate of the non-dynamical vanishing of the torsion of a linear connection (see below).
over, using modern coordinate-free methods for the description of the physical systems (the exterior calculus), one can provide the general covariance without any appeal to the gravity theory. There is also another point of view on general coordinate transformations as the group of local diffeomorphisms of the spacetime manifold. These are active transformations due to the motion of spacetime points. But also in this case, the study of the invariance properties of physical systems is not related to gravity since the corresponding mathematical formalism of the Lie derivatives along vector fields, defined by diffeomorphisms, does not demand the existence of a geometrical structure on the manifold.

Below we show that it more natural to describe the action of the spacetime symmetries not on the base—the spacetime itself—but rather in bundle of affine or linear frames over it. Then the gravitational field turns out to be a connection in this principal fiber bundle. There is a natural homomorphism from the bundle of frames to the tangent bundle, using which the gauge gravitational field induces geometrical structures on the spacetime: the connection and the metric. The issue of the status of the metric in this approach has an unexpected dual solution. Namely, the metric structure is explained with the help of a spontaneous symmetry breaking mechanism when the reduction of the inhomogeneous general affine group $GA(4, R)$ to linear group $GL(4, R)$ allows to interpret the tetrad field as a nonlinear realization of the connection, and on the other hand, the reduction of $GL(4, R)$ to the Lorentz group adds a Goldstone contribution to the metric. Roughly speaking, the metric turns out to be some hybrid of the two popular interpretations—the translational gauge field (in the nonlinear realization) and the Goldstone field taking values in the homogeneous space $GL(4, R)/SO(3, 1)$.

In the consistent gauge approach to the gravity theory, the largest difficulties arise when constructing its dynamics. This is explained by an essentially non unique definition of an action by the gauge invariance. In chapter 6, we will consider some models, for which the study of the physical properties can lead to the final establishment of the structure of the gauge gravity theory.

### 5.2. Bundle of frames and generalized affine connection

We construct the consistent gauge scheme for the gravitational field using the mathematical theory of the generalized affine and linear connections in the bundles of frames over the spacetime $M_4$. We understand the latter as the four-dimensional smooth ($C^\infty$) manifold. When discussing the geometrical aspects of the bundle spaces, we use the local coordinates writing the basic objects in components and defining their transformation laws. Certainly, we then lose the advantages of coordinate-free methods that make a stress on the global invariant meaning of various operations, however, the component formalism is more familiar to physicists. A brief summary of the basic mathematical definitions (manifold, forms, bundles, connection, etc.) is given in Appendix A1.
Bundle of linear frames

A linear frame at the point \( x \in M \) is the basis \( \{ e_a \} = \{ e_0, e_1, e_2, e_3 \} \) of the tangent space \( T_x M \). The set of frames \( G_x \) at this point is isomorphic to the general linear group \( GL(4, R) \), since any frame is obtained by a linear transformation from the standard basis \( E_1 = (1, 0, 0, 0), E_2 = (0, 1, 0, 0), E_3 = (0, 0, 1, 0), E_4 = (0, 0, 0, 1) \) in \( R^4 \cong T_x M \): \( e_a = L^b_a E_b \) with the \( 4 \times 4 \) matrix \( L \in GL(4, R) \). The union \( \bigcup_{x \in M} G_x = L(M) \), provided with a smooth structure, is a \( (4 + 4^2) \)-dimensional manifold, where the natural projection \( \pi : L(M) \to M \) maps the linear frame \( e_a \) at the point \( x \) to \( x \). The manifold \( L(M) \) is called the bundle of linear frames. This is the principal bundle with the structural group \( GL(4, R) \); the right action \( GL(4, R) \) on \( L(M) \) is evident: \( e_a \in G_x \to e'_a = e_a L^{-1} b \in G_x \) for \( L \in GL(4, R) \). The tangent bundle \( T(M) \) may be viewed as an associated with \( L(M) \) bundle with the standard fiber \( R^4 \).

The connection in \( L(M) \) is called the linear connection in \( M \). It introduces the parallel transport in \( L(M) \) and induces the parallel transport of the tangent vectors in the associated \( T(M) \) and defines the covariant differentiation of the vector fields. Choosing the local coordinates \( \{ x^i \} \) on \( M \) and a cross-section \( \sigma : M \to L(M) \), we find the 1-form on the base \( \sigma^* \omega = \omega^{ab} dx^a dx^b \) which is uniquely determined by the linear connection form \( \omega \) on \( L(M) \) with the values in the Lie algebra \( gl(4, R) \). Here \( E^a_b \) is the natural basis of the Lie algebra \( gl(4, R) \), i.e., the \( 4 \times 4 \) matrix with elements at the intersection of \( a \)-th column and \( b \)-th line equal to 1, and with the remaining elements equal zero. The coefficients \( \omega^{ab} \) of this form are called the coefficients of the local linear connection. They define the coordinate notation of the covariant derivative of the tangent vector \( \nabla_{dx^a} v^b = \partial_a v^b + \omega^{ab} v^b \), where the vector \( v = v^a e_a \) is decomposed with respect to the basis \( e_a(x) \) that is defined by the cross-section \( \sigma \). When the cross-section is changed, \( \sigma \to \sigma' \), the connection coefficients are transformed as follows:

\[
\omega^{ab}_{\mu} \longrightarrow \omega'^{ab}_{\mu} = L^c_{a \mu} \omega^{cb} d\mu L^{-1}_{\nu b} + L^c_{a c} \partial_a \mu L^{-1}_{\nu b} c, \tag{5.1}
\]

where the matrix of the linear transformation is determined from

\[
e_a'(x) = e_a(x) L^b_{a \nu} (x).
\]

In the bundle of linear frames \( L(M) \), one can introduce (absolutely independently from the connection) another basic object – the canonical 1-form \( \theta \). This \( R^4 \)-valued form is defined on \( L(M) \) by the equation

\[
d\pi(X) = \theta^a(X) e_a, \tag{5.2}
\]

where \( X \in T_u(L(M)), u = (x, e_a) \). The canonical 1-form is horizontal, i.e., it vanishes on vertical (tangent to the fiber) vectors

\[
\theta^a(X) = 0 \iff d\pi(X) = 0.
\]

This property is obvious. In addition, \( \theta \) is equivariant, i.e., for any \( L \in GL(4, R) \)

\[
R^*_L \theta^a = \theta^a \circ dR_L = L^{-1}_{\nu b} \theta^b.
\]
Indeed, let \( u = (x, e_a) \in L(M) \), and \( d\pi (X) = \theta^a(X)e_a \). Then the right action \( R_Lu = (x, e_a L^a_b) \), and its differential \( dR_L(X) \in T_{R_Lu}L(M) \), as a result,

\[
(R^a_L \theta^a) e_a = d\pi(dR_LX) = d\pi(X) = (\theta^a(X)L^{-1}e_a)(L^b_ce_b).
\]

In the local notation, the canonical form determines the coframe 1-form \( h^a : T_x(M) \to R^4 \), i.e., the basis of the cotangent space \( T_x^*M \) dual to \( e_a \). If the cross-section is chosen \( \sigma : M \to L(M) \) [i.e. the field of frames \( e_a(x) \)], then we have \( h^a(x) = (\sigma^*\theta^a)(x) \), where \( h^a(e_b) = \delta^a_b \).

With the connection \( \omega \) and the canonical form \( \theta \) defined on the bundle of linear frames, one can introduce a parallelization of \( L(M) \). Namely, one can:

1) map each element of the Lie algebra \( gl(4, R) \) to a fundamental\(^3\) vertical vector field; in particular, map the basis of the Lie algebra \( gl(4, R) : E^a_b \to E^a_b \), so that \( \omega(E^{a}_{ab}) = E^a_b \), and 2) map each vector from \( R^4 \) to a standard horizontal vector field, in particular map the basis \( E_a \to E_a^* \), so that \( \theta(E_a^*) = E_a \). Then \( 4 + 4^2 \) vector fields \( E_a^* \), \( E^{a}_{ab} \) constitute at each point \( u \in L(M) \) the basis of the tangent space \( T_u(L(M)) \).

The linear connection in \( L(M) \) is characterized by the curvature and the torsion. Let us define the curvature 2-form \( R \) of the linear connection \( \omega \) as covariant exterior differential\(^4\)

\[
R = D\omega = d\omega(h). \tag{5.3}
\]

The torsion 2-form \( Q \) of the linear connection \( \omega \) is defined as the covariant exterior differential of the canonical form \( \theta \):

\[
Q = D\theta = d\theta(h). \tag{5.4}
\]

The 2-forms \( R \) and \( Q \) satisfy the structure equations

\[
\begin{align*}
Q^a &= d\theta^a + \omega^a_b \wedge \theta^b, \\
R^a_b &= d\omega^a_b + \omega^a_c \wedge \omega^c_b,
\end{align*}
\tag{5.5}
\]

where we made the expansion with respect to the natural bases in \( R^4 \) and \( gl(4, R) : Q = Q^a E_a, \theta = \theta^a E_a, R = R^a_b E^b_a \).

Introducing the local coordinates \( x^\mu \) on the manifold \( M_4 \), we naturally define the corresponding cross-section in \( L(M) \) by attaching the coordinate basis \( \{ \partial_\mu \} \) at \( x \in M_4 \). Then the local coordinates in \( L(M) \) are \( (x^\mu, h^a_\mu) \), where \( e_a = h^a_\mu \partial_\mu \), and it is obvious that \( \det(h^a_\mu) \neq 0 \). Determining the inverse matrix \( h^a_\mu \) so that \( h^a_\mu h^\mu_b = \delta^a_b \), we obtain from the definitions (5.2)-(5.4) the local expressions for the canonical 1-form

\[
\theta^a = h^a_\mu dx^\mu \tag{5.6}
\]

\(^3\)For definitions, see Appendix A3 and references therein.

\(^4\)In (5.3) and (5.4), \( h \) denotes the projection to the horizontal subspace of \( TL(M) \)
and for the connection 1-form
\[ \omega^a_b = h^a_\alpha \Gamma^\alpha_{\beta\nu} h^\beta_b \, dx^\nu + h^a_\mu dh^\mu_b. \] (5.7)

Then using the structure equations (5.5), we arrive at the ordinary tensor expressions for the torsion (1.3) and the curvature (1.1), constructed from the components of the linear connection \( \Gamma^\alpha_{\beta\mu} \).

**Bundle of affine frames**

The tangent space \( T_x(M) \cong \mathbb{R}^4 \) may be considered as the affine space \( A^4 \). In this case, we call it the tangent affine space and denote \( A_x(M) \). Its basis, which is called the affine frame, includes the point \( z \in A_x(M) \) and linear frame \( e_a \).

The set of affine frames at the point \( x \in M_4 \) is isomorphic to general affine group \( GA(4, R) \), since an arbitrary affine frame \( (z, e_a) \) is obtained using the affine transformation \( A \in GA(4, R) \) from the natural frame \( (0, E_a) \) in \( A^4 \). The set of all affine frames on \( M \) is called the bundle of affine frames \( A(M) \). The projection \( \pi : A(M) \to M \) maps a frame at the point \( x \) to this point \( x \in M_4 \). The bundle \( A(M) \) is the principal bundle over \( M \) with the structural group \( GA(4, R) \) with an obvious definition of the group action.

A natural realization of affine transformations in \( A_x(M) \) is obtained, if we view the affine space \( A^4 \) as a hypersurface in \( \mathbb{R}^5 \) with a fixed fifth coordinate. Then an arbitrary element of \( A^4 \) is represented as a column \( \begin{pmatrix} z \\ 1 \end{pmatrix} \), and the action of the affine group \( GA(4, R) \) in the affine space \( A^4 \) is described by the \( 5 \times 5 \) matrix of the form \( g = \begin{pmatrix} L & b \\ 0 & 1 \end{pmatrix} \), where \( L \in GL(4, R) \) and \( b \in \mathbb{R}^4 \):

\[
\begin{pmatrix} z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} z' \\ 1 \end{pmatrix} = \begin{pmatrix} L & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z \\ 1 \end{pmatrix} = \begin{pmatrix} Lz + b \\ 1 \end{pmatrix}.
\]

The bundle of affine frames \( A(M) \) and the bundle of linear frames \( L(M) \) are closely related to each other and to the associated tangent bundles. In particular, the affine tangent bundle represents an important example of a soldered bundle, the properties of which are reviewed below.

We say that the bundle \( E(M, F, G, P) \) associated with the principal bundle \( P(M, G) \) is soldered to the base \( M \), if [3]:

1) the fiber \( F \) of the bundle \( G \) is a homogeneous factor-space \( F = G/H \), where \( H \subset G \) is a stationary subgroup (or stabilizer), and the dimension of \( F \) is equal to dimension of the base \( M \);

2) there is a global cross-section \( \sigma : M \to E \), and there is a natural isomorphism between the tangent bundle \( T(M) \) and the space of all vectors tangent to the fibers \( F \) at the points of the cross-section \( \sigma(M) \). In other words, the “soldering” of the bundle \( E \) to the base \( M \) means that the fibers \( F \) are tangent
5.2. Bundle of frames and generalized affine connection

To the manifold \( M \) (in the sense of coincidence of the tangent spaces) at the soldering point set by the cross-section \( \sigma \). Furthermore, the bundle \( \tilde{T}(M) \) of the tangent spaces \( T_{\sigma(M)}F \) may be considered as a vector bundle associated with the principal bundle \( \tilde{P}(M,H) \), obtained as a submanifold in \( P(M,G) \) when homeomorphisms \( F \to F_x \) are such that the “center” of \( F \) is mapped into \( \sigma(M) \) – the soldering point.

Let us consider the most important case when the fiber \( F \) of the soldered bundle \( E \) is a weakly reductive homogeneous space. This means that the Lie algebra \( G \) of the structural group \( G \) is decomposed into the sum of the subalgebra \( H \) (corresponding to the subgroup \( H \)) and the vector space \( V \)

\[
G = H \oplus V \tag{5.8}
\]

so that \([H, H] \subset H, [H, V] \subset V\), where \([\ , \] is the commutator in the Lie algebra. The tangent space \( T_{\sigma(M)}F \) may be identified with \( V \).

A \( V \)-valued 1-form \( \tilde{\theta} \) on the principal bundle \( \tilde{P}(M,H) \) is called a soldering form if it satisfies the following conditions:

1) \( \tilde{\theta}(\tilde{X}) = 0 \iff d\tilde{\pi}(\tilde{X}) = 0 \), where \( \tilde{X} \in \tilde{T}\tilde{P} \), i.e., it vanishes on vertical vectors;

2) under the right action \( R^*_h \tilde{\theta} = \tilde{\theta} \circ dR_h = h^{-1}\tilde{\theta}h \), where \( h \in H \).

One can show [3] that the existence of the soldering 1-form \( \tilde{\theta} \) is necessary and sufficient condition for the bundle \( E \) to be soldered to \( M \).

Now we consider in details the relation of the bundles of affine and linear frames. It is obvious that \( L(M) \) may be realized as a natural subbundle \( A(M) \). The corresponding embedding \( \gamma : L(M) \to A(M) \) is defined by the map

\[
\varphi : A(M) \to R^4, \tag{5.9}
\]

such that

\[
R^{*g^{-1}} \circ \varphi^a = A^a_b \varphi^b + b^a, \quad g = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} \in GA(4,R).
\]

Then the subbundle \( L(M) \) is determined as a kernel

\[
L(M) = \{ u \in A(M) \mid \varphi(u) = 0 \}.
\]

The canonical projection \( \beta : A(M) \to L(M) \) can be defined as

\[
\beta(u) = R_{\Phi(u)} \circ (u), \quad \Phi(u) = \begin{pmatrix} 1 & \varphi(u) \\ 0 & 1 \end{pmatrix}, \quad u \in A(M),
\]

where \( 1 \) is the unit in \( GL(4,R) \). It is obvious that \( \beta \circ \gamma = id_{L(M)} \). In local coordinates, the embedding of \( L(M) \) into \( A(M) \) maps the linear frame \((x, e_a)\) to the affine frame \((x, O_x, e_a)\), where \( O_x \) is the origin of the affine tangent
space $A_x$. Conversely, the projection $\beta: (x, z, e_a) \rightarrow (x, e_a)$. Using the above-mentioned realization of the general affine transformations, and consequently, of the affine frames as a $5 \times 5$ matrix of the following form:

$$u = \begin{pmatrix} L & z \\ 0 & 1 \end{pmatrix}, \quad L \in GL(4, R), \ z \in R^4,$$

we get explicitly the mapping $\varphi: A(M) \rightarrow R^4$. For $(x, u) \in A_x$, it is defined as $\varphi(u) = -L^{-1}z$.

The matrix inverse to $g = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}$ reads $g^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}b \\ 0 & 1 \end{pmatrix}$, and it is easy to show that under the right group action

$$R_{g^{-1}}u = ug^{-1} = \begin{pmatrix} L & z \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A^{-1} & -A^{-1}b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} LA^{-1} & -LA^{-1}b + z \\ 0 & 1 \end{pmatrix}$$

we have the correct transformation property

$$R_{g^{-1}}^* \varphi(u) \equiv \varphi(R_{g^{-1}}u) = -(LA^{-1})^{-1}(z - LA^{-1}b)$$

$$= -AL^{-1}z + b = A\varphi(u) + b.$$

Defined as the kernel $\varphi(u) = 0$, the subbundle $L(M)$ is formed by the set of matrices $\begin{pmatrix} L & 0 \\ 0 & 1 \end{pmatrix} \in GL(4, R)$, which is indeed the bundle of linear frames. The canonical projection is also quite transparent in the matrix representation:

$$\beta(u) = R_{\Phi(u)} \circ u = \begin{pmatrix} L & z \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \varphi(u) \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} L & L\varphi(u) + z \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} L & 0 \\ 0 & 1 \end{pmatrix}.$$
field of affine frames $u(x) = \begin{pmatrix} L(x) & z(x) \\ 0 & 1 \end{pmatrix}$ on $M_4$, and the generalized affine connection 1-form $\Omega$ induces the form $\Omega_\sigma = \sigma^* \Omega$ on the base space. Since the Lie algebra of the general affine group is decomposed into the (semidirect) sum $\mathfrak{ga}(4, \mathbb{R}) = \mathfrak{gl}(4, \mathbb{R}) + \mathbb{R}^4$, the local connection form $\Omega_\sigma$ also can be decomposed into the sum $\Omega_\sigma = \tilde{\omega} + \tilde{\phi}$.

It is more convenient to use the matrix representation for the Lie algebra $\mathfrak{ga}(4, \mathbb{R}) = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$, $a \in \mathfrak{gl}(4, \mathbb{R})$, $b \in \mathbb{R}^4$, and write the local connection form as the $5 \times 5$ matrix $\Omega_\sigma = \begin{pmatrix} \tilde{\omega} & \tilde{\phi} \\ 0 & 0 \end{pmatrix}$.

(5.10)

Now we will show that this generalized affine connection determines the linear connection and some tensor form in $M_4$. For that purpose, we consider the transformation of the connection (5.10) under the change of the cross-section $\sigma \to \sigma'$. Then we have

$u'(x) = R_{g(x)}u(x) = u(x)\begin{pmatrix} A(x) & b(x) \\ 0 & 1 \end{pmatrix}$

and $\Omega_{\sigma'} = g\Omega_\sigma g^{-1} + gdg^{-1}$. Explicitly,

$$\begin{pmatrix} \tilde{\omega}' & \tilde{\phi}' \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{\omega} & \tilde{\phi} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}^{-1} + \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} d \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} A\tilde{\omega}A^{-1} + AdA^{-1} & A\tilde{\phi} - db \ - (A\tilde{\omega}A^{-1} + AdA^{-1})b \\ 0 & 0 \end{pmatrix}.$$
\[ \bar{\theta} = \bar{\varphi} + d\varphi + \bar{\omega}\varphi. \]  \hspace{1cm} (5.11)

Indeed, under the change of the cross-section
\[ \bar{\varphi}' = A\bar{\varphi} - db - (A\bar{\omega}A^{-1} + AdA^{-1})b, \]
\[ \bar{\omega}' = A\bar{\omega}A^{-1} + AdA^{-1}, \]
\[ \varphi' = -L^{-1}z' = A\varphi + b, \]
from which we find
\[ \bar{\theta}' = \bar{\varphi}' + d\varphi' + \bar{\omega}'\varphi' = A\bar{\varphi} - db - (A\bar{\omega}A^{-1} + AdA^{-1})b \]
\[ + d(A\varphi + b) + (A\bar{\omega}A^{-1} + AdA^{-1})(A\varphi + b) = A\bar{\theta}. \]  \hspace{1cm} (5.12)

The converse is also true. Indeed, given a local linear connection 1-form \( \omega \)
and a tensor form \( \varphi_1 \) with the transformation law \( \varphi_1' = A\varphi_1, \ A \in GL(4, R) \),
this pair uniquely determines the generalized affine connection in \( M_4 \):
\[ \Omega_\sigma = \left( \begin{array}{ccc} \omega & \varphi_1 - d\varphi - \omega\varphi \\ 0 & 0 \end{array} \right). \]  \hspace{1cm} (5.13)

Obtained in local coordinates, these conclusions are valid also in the bundle
(without introducing the cross-sections). Namely, there is a one-to-one correspondence between the set of generalized affine connections and the set of pairs
(linear connection + tensor 1-form) [3].

Recall now that another important structure—the canonical form \( \theta \)—is introduced in \( L(M) \) independently of the linear connection \( \omega \).
In accordance with (5.13), the pair \( (\omega, \theta) \) determines a special case of the generalized affine connection \( \Omega \), which is then called the affine connection. However, the tensor form \( \theta \)
defined by the generalized affine connection using (5.11) does not coincide with
the canonical form. Moreover, the set of generalized affine connections may be divided into two essentially different classes. The first class are connections \( \Omega \)
that define non-degenerate tensor 1-forms \( \theta \), for which \( \theta(X) = 0 \) means
that the vector \( X \) is vertical, i.e., tangent to a fiber in \( L(M) \). The second class encompasses the connections \( \Omega \) that define degenerate \( \theta \). In the local coordinates
\( \theta = \theta^a_{\mu} E_\mu dx^a \), and the degeneracy means that \( \det \theta^a_{\mu} = 0 \).
One can show that \( \theta \) of the first class is a soldering form for the bundle of tangent affine spaces.

The affine tangent bundle \( E(M, A_4, GA(4, R), A(M)) \) is associated with the principal bundle \( A(M) \) of affine frames. The typical fibre in it is the affine space
\( F = R^4 = GA(4, R)/GL(4, R) \) with the dimension \( 4 = \dim M \), and this is a
weakly reductive homogeneous space, since \( G = ga(4, R), \ H = gl(4, R), \ V = R^4 \)
and \([H, V] \subseteq V, [V, V] \subseteq H\). The soldering is performed by the choice of
a vector field of the “origin” of tangent affine spaces. The fact that the non-degenerate 1-form \( \bar{\theta} \) is a soldering form, is almost trivial. Indeed, if \( X \in TL(M) \)
is vertical, \( d\sigma(X) = 0 \), then \( \bar{\theta}(X) = D\varphi(X) = d\varphi(hX) = 0 \); the converse is
true due to the non-degeneracy of \( \bar{\theta} \). The second property is also fulfilled:
\[ R^*_b \left( \begin{array}{cc} 0 & \bar{\theta} \\ 0 & 0 \end{array} \right) = \left( \begin{array}{cc} 0 & A^{-1}\bar{\theta} \\ 0 & 0 \end{array} \right) = \left( \begin{array}{cc} A^{-1} & 0 \\ 0 & 1 \end{array} \right) \left( \begin{array}{cc} 0 & \bar{\theta} \\ 0 & 0 \end{array} \right) \left( \begin{array}{cc} A & 0 \\ 0 & 1 \end{array} \right), \]
with
\[ h = \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \in GL(4, R). \]

The vector field \( \varphi^a(x) = -L^{-1}a_b(x)z^b(x) \) defines the origin of tangent affine spaces for the chosen cross-section \( \sigma \):
\[ M \rightarrow A(M), \quad u = u(x). \]

### 5.3. Spontaneous symmetry breaking and non-linear realizations

Now we consider the general formalism of nonlinear group realizations as the most appropriate method of description of the spontaneous symmetry breaking.

At the core the method of nonlinear realizations is the notion of auxiliary fields which are identified with the Goldstone fields \([115]\). The latter appear in a physical system when the initial symmetry group \( G \) of Lagrangian of the theory is spontaneously broken to some subgroup \( H \subset G \), with respect to which the physical vacuum is invariant. In this case, the ground state becomes degenerate, and the “manifold of vacua” is isomorphic to homogeneous factor space \( G/H \).

The parameters of \( G/H \) corresponding to the generators of the broken symmetry can be identified with the Goldstone fields which are considered as vector fields over \( M \) with the values in \( G/H \). From the geometrical point of view, the Goldstone fields are thus the cross-sections of the vector bundle with the fiber \( G/H \) that is associated with some principal bundle \( P(M, G) \). This refers to the local or the gauge symmetries of a physical system. According to the reduction theorem \([3]\), the existence of a cross-section of an associated \( (G/H) \)-bundle is necessary and sufficient for the reduction of the principal \( P(M, G) \) bundle to the subbundle \( \tilde{P}(M, H) \). Therefore, in the geometrical approach to the theory of gauge fields, one can speak of a spontaneous symmetry breaking \( G \rightarrow H \), when there is a reduction of the corresponding principal bundles. Here we consider this geometrical picture, that should be distinguished from the spontaneous symmetry breaking in the sense of a non-invariance of the quantum vacuum state, since the “vacuum” and everything else in this theory are the purely classical (geometrical) notions.

A nonlinear realization of the group \( G \) in the theory with the broken \( G \rightarrow H \) symmetry is constructed from the linear realization, and the key link in this construction are the Goldstone or auxiliary fields. Let the matter field \( \Psi(x) \) be described as a cross-section of a vector bundle \( E(M, V^m, G, P) \) associated with the principal bundle \( P(M, G) \), where the fiber \( V^m \) is the vector space of a representation \( \rho \) of the group \( G \) in \( GL(m, R) \).

The field \( \Psi \) is transformed with respect to the linear (irreducible) representation \( \rho \) of the group \( G 
\)
\[ \Psi \rightarrow \rho(g)\Psi, \quad g \in G. \]  
\( (5.14) \)
The connection in the principal bundle \( P(M, G) \) introduces the parallel transport in \( E \). Accordingly, choosing some cross-section, the covariant derivative of the field \( \Psi \) is defined, which has the following form in the local coordinates:

\[
D\Psi = d\Psi + \rho'(\omega)\Psi, \tag{5.15}
\]

where \( \omega \) is a \( G \)-valued connection 1-form, and \( \rho' \) is an appropriate representation of the Lie algebra \( G \).

When the symmetry \( G \) is spontaneously broken to the group \( H \), there is a reduction \( P(M, G) \to \tilde{P}(M, H) \), and hence there exists a cross-section of an associated \( G/H \)-bundle – the Goldstone field:

\[
\xi(x) \in G/H. \tag{5.16}
\]

Using the auxiliary fields (introducing the “origin” in the fibers of an associated bundle), one can construct the nonlinear realization of the group \( G \). For this purpose, we consider nonlinear fields defined by

\[
\Psi_H(x) \equiv \rho(\xi^{-1}(x))\Psi(x). \tag{5.17}
\]

We see that a nonlinear realization \( G \) is defined, if under the action of elements of the group \( g \in G \), the pair \( (\xi, \Psi_H) \) is transformed according to

\[
\begin{align*}
\xi & \rightarrow g \xi' \\
\Psi_H & \rightarrow g \rho(h')\Psi_H,
\end{align*} \tag{5.18a}
\]

\[
\begin{align*}
\rho(\xi') & = \rho(\xi' h) \Psi_H,
\end{align*} \tag{5.18b}
\]

where \( \xi' \in G/H \), \( h' \in H \), and \( \xi' = \xi'(g, \xi) \), \( h' = h'(g, \xi) \) are the nonlinear functions of their arguments that are uniquely determined from

\[
g\xi = g' = \xi' h', \tag{5.19}
\]

since any element of the group \( G \) is uniquely represented as a product of the element from the factor space \( G/H \) times an element of the subgroup \( H \). It is obvious that under the action of the subgroup, the representation (5.18) is linear. Indeed, let \( g = h \in H \), then \( \xi \rightarrow \xi' = Ad_h \xi = h\xi h^{-1} \), \( \Psi_H \rightarrow \rho(h)\Psi_H \).

The notion of a nonlinear gauge field is the crucial one in the theory with the spontaneously broken gauge symmetry. The nonlinear gauge field defines the covariant derivative of \( \Psi_H \) and it is constructed on the basis of the linear representation (5.14), (5.15) with the help of the transition formula (5.17). Let \( \omega \) be the original \( G \)-valued connection 1-form on \( M \). Under the change of the cross-section, it transforms as

\[
\omega \rightarrow \omega' = g\omega g^{-1} + gdg^{-1}, \quad g(x) \in G.
\]

In the method of nonlinear realizations, we define a nonlinear gauge field as

\[
A = \xi^{-1} \omega \xi + \xi^{-1} d\xi, \tag{5.20}
\]
where $\xi(x)$ is an auxiliary field (5.16). Under the action of $g$, the connection $A$ is transformed nonlinearly,

$$A \rightarrow A' = h' Ah'^{-1} + h' dh'^{-1},$$

(5.21)

where $h' = h'(\xi, g)$ is determined from (5.19). Therefore,

$$D^A \Psi_H = d\Psi_H + \rho'(A)\Psi_H$$

is called the covariant derivative of $\Psi_H$ in the nonlinear realization of $G$.

Usually, the homogeneous space $G/H$ is weakly reductive. The Lie algebra of the gauge group $G$ can be decomposed into the sum $G = H \oplus V$, where $H$ is the Lie algebra of the subgroup $H$, [116]. Let us consider the corresponding decomposition $A = \Gamma + B$ of the nonlinear gauge field $A$ into the $H$-valued 1-form $\Gamma$ and the $V$-valued form of $B$. From the definition (5.20), the following transformation law under the action of $g \in G$ is obvious:

$$\Gamma \rightarrow \Gamma' = h' \Gamma h'^{-1} + h' dh'^{-1},$$

(5.22)

$$B \rightarrow B' = h' B h'^{-1}.$$  

(5.23)

We thus find that $\Gamma$ transforms as the usual connection (associated with the connection in reduced principal bundle $\tilde{P}(M, H)$), although with the nonlinearly realized $h' = h'(\xi, g) \in H$. However, the $V$-component of the nonlinear connection $B$ transforms homogeneously. Taking this fact into account, one can treat the method of nonlinear realizations as an alternative to the Higgs mechanism, since due to the homogeneous law (5.23), the nonlinear gauge field allows for a non-zero mass term in the Lagrangian.

In conclusion, let us consider the important example of a nonlinear realization that clearly demonstrates the above-mentioned method: the spontaneous symmetry breaking in the Higgs mechanism.

Take the Lagrangian for the multiplet of scalar (Higgs) fields $\varphi$,

$$L = -\frac{1}{2} g^{\mu\nu} (D_\mu \varphi)(D_\nu \varphi) + V(\varphi),$$

(5.24)

where $D_\mu \varphi = \partial_\mu \varphi + \rho'(\omega_\mu) \varphi$. This Lagrangian is invariant with respect to the linear gauge transformations, $\varphi \rightarrow \rho(g) \varphi$, of the group $G \ni g$. When the symmetry is spontaneously broken $G \rightarrow H$, we proceed to the nonlinear realization according to (5.17) by defining the Goldstone fields $\xi(x)$. We introduce the nonlinear matter field $\varphi_H = \rho(\xi^{-1}) \varphi$, in terms of which (5.24) is recast into

$$L = -\frac{1}{2} g^{\mu\nu} (D^A_\mu \varphi_H)(D^A_\nu \varphi_H) + V(\varphi_H).$$

(5.25)

In view of the spontaneous symmetry breaking, the vacuum average value $\eta = \langle \varphi_H \rangle = \text{const} \neq 0$, and we make the usual shift to the physical fields $\chi = \varphi_H - \eta$. Furthermore, we have $D^A \varphi = D^A \chi + \rho'(B) \eta$, since by assumption,
the “vacuum” is invariant with respect to the subgroup \( H \subset G \), therefore its generators annihilate \( \eta \), whereas the Higgs field itself transforms as a nonlinear multiplet \( \chi \rightarrow \rho(h')\chi \) under the action of \( G \). Hence, the Lagrangian (5.25) manifests the Higgs phenomena in the nonlinear realization: the Goldstone degrees of freedom \( \xi \) were absorbed, so that the nonlinear field \( B_\mu \) became massive:

\[
g^{\mu\nu}(D^A_\mu \varphi_H)(D^A_\nu \varphi_H) = g^{\mu\nu}(D^F_\mu \chi)(D^F_\nu \chi) + 2\rho'(B^\mu)\eta D_\mu \chi + \eta^2 B^\mu B_\mu.
\]

The total Lagrangian (5.25) is still invariant with respect to the group \( G \), which is now realized nonlinearly, and this symmetry is “hidden”. The remaining symmetry with respect to the subgroup \( H \) is explicit and realized linearly.

5.4. Kinematics of the gauge gravity theory

The kinematics of the gauge theory is usually understood as the identification of the basic symmetry group and the description of the corresponding dynamical variables – the gauge fields, as well as the investigation of their connection with the geometrical structure of spacetime, and the related issues.

Nonlinear gauge theory of Poincaré group

The problem of a choice of the basic symmetry group in the gauge gravity theory is not purely theoretical, and it can be ultimately settled by experiment only. In a similar way, when constructing the unified theories of strong, weak and electromagnetic interactions, one can discuss only to a limited extent such questions as the number of quarks, the basic group \( SU(N), N = ? \), etc; all this belongs to the field of experiment, so at best, one can specify only some minimal model that consistently describes the experimental data available at the current technological level.

The Poincaré group is important in the relativistic particle theory, and one can assume that this spacetime symmetry is directly relevant to the formation of the geometrical spacetime structure. Thus, the minimal model of the gauge gravity theory should be most probably based on the Poincaré group.

We construct the gauge theory of the Poincaré group \( P_{10} \) similarly to the gauge theories of internal symmetries as a theory of connection in a corresponding principal bundle. However, the specific feature of the gravity theory (namely, the relation between the gauge transformations and the symmetries of spacetime) calls forth the attempts to explain the spacetime geometry directly in terms of the gauge fields. This means that, in contrast to the theories of internal symmetries, where the spacetime exists as an external background unrelated to the gauge fields, in gravity theory the geometrical spacetime structure (metric and connection, see Chapter 1) is fully determined by the corresponding gauge theory potentials. Accordingly, the principal bundle cannot be arbitrary in this case, but should be constructed as a natural spacetime object.
As the principal $P_{10}$-bundle of the gauge theory of Poincaré group, we consider the bundle of affine orthonormal frames $O(M)$ as a natural subbundle of the bundle of affine frames $A(M)$. The gauge potential is then obtained as a restriction of the generalized affine connection to $O(M)$.

The central issue is to explain how the gauge gravity field (namely, the $P_{10}$-connection) determines the geometrical structure of $M_4$, i.e., the metric and connection on the tangent bundle $TM$. One conclusion looks fairly obvious: the rotational (corresponding to the Lorentz subgroup $L_6$) part $\tilde{\omega}$ of the generalized affine connection $\Omega$, recall (5.10), may be identified with the Lorentz connection (when the cross-section $\sigma : M \to O(M)$ is chosen): $\Gamma^b_a \equiv \tilde{\omega}^b_a |_{P_{10}}$.

The situation is more complex with the translational part $\tilde{\phi}$ (corresponding to the subgroup of translations $T_4 = R^4$) of $\Omega$ in (5.10). It is clear that the 1-form $\tilde{\phi}$ cannot be identified with any tensor object like the tetrad or the metric field, as one could expect, taking into account that one of the sources of the gravitational field –the energy-momentum tensor– is obviously related to translations. However, we have shown that the reduction of $A(M)$ to the bundle of linear frames $L(M)$, defined by the 0-form $\phi$ (5.9), uniquely maps $\tilde{\phi}$ into the tensor 1-form of $\tilde{\theta}$ (5.11). Accordingly, we can try to explain the spacetime metric structure in terms of the field $\tilde{\theta}$.

The analysis of the generalized affine connection (see Sec. 5.2.), demonstrates that the gauge $P_{10}$-theory is essentially wider than its usual interpretation as a dynamical theory of the tetrad $h^a_\mu$ and the Lorentz connection $\Gamma^a_{b\mu}$ fields. Indeed, the $P_{10}$-theory of the second class, see Sec. 5.2., does not define the metric structure on $M_4$ at all, since in this case the form $\tilde{\theta}$ is degenerate ($\det \tilde{\theta}^a_\mu = 0$) and cannot be identified with the 1-form of coframe. More exactly, we can say that in this case the spacetime has the ultralocal degenerate Carroll geometry [117], where the metric is degenerate, $\det g_{\mu\nu} = 0$, everywhere in $M_4$. As it is shown in [118], such a state could be realized in the singularity, and within the perturbation theory developed in [118, 119] this corresponds to the zero metric signature with $\sigma = 0$ in (2.94), and (7.51). Physically, such region is characterized by distances much smaller than the Planck length $l_0 = 10^{-35}$ m.

In contrast to this unusual situation, the gauge $P_{10}$-theory of the first class introduces the metric on $TM$ in a consistent way. However, in this case the resulting structure is much wider than the Riemann-Cartan structure. Indeed, we have shown in Sec. 5.2. that the 1-form $\tilde{\theta}$ is the generalized soldering form of the affine tangent bundle, but it does not coincide with the canonical form $\theta$ (the coframe 1-form $\theta = h$), as it was noticed in [103]. The difference of $\tilde{\theta}$ from $\theta$ is in arbitrary general linear transformation $\tilde{\theta}^a = L^a_b \theta^b$, $L \in GL(4, R)$.

With an account of the explicit covariance with respect to the local Lorentz transformations (we consider the principal bundle of orthonormal frames), it reduces to the difference on the Goldstone transformation from the factor space $GL(4, R)/SO(3, 1)$. One can use this arbitrariness in the definition of tetrads to construct the conformal gravity theory of new type in the Riemann-Cartan spacetime, as we discussed in Sec. 4.1., see [69].
In the rest of the section, we present the rigorous foundation for these conclusions in the framework of the nonlinear realization of the gauge $\mathcal{P}_{10}$-theory.

Construction of the nonlinear gauge theory of the Poincaré group consists (schematically) of the two stages. The first one is formulated [99]-[105] as the soldering procedure, when the bundle of affine frames is reduced to the bundle of linear frames. In physical terms, this means the spontaneous breaking of the translational symmetry, which reflects the fact that the physical matter fields depend only on the points of the base $M_4$, but not on the points of the fiber $[99, 100]$. The (implicit) second stage is very important: the principal $\mathcal{P}_{10}$-bundle of the gauge gravity theory is itself a subbundle in $\mathcal{A}(M)$, in other words, it is obtained as the corresponding restriction or reduction of the structural group $G(4, R) \rightarrow \mathcal{P}_{10}$. In accordance with this, we notice that an arbitrary element $g \in GA(4, R)$ is uniquely represented in the $5 \times 5$ matrix disguise as

$$g = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \xi \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u & 0 \\ 0 & 1 \end{pmatrix} h, \quad \text{where} \quad h = \begin{pmatrix} L & 0 \\ 0 & 1 \end{pmatrix}.$$  

Here, the first factor is the Goldstone translational field – the element from $GA(4, R)/GL(4, R)$, and the second factor $u \in GL(4, R)/SO(3, 1)$ is the Goldstone field responsible for the reduction of $A(M)$ to $AO(M)$; finally, the matrix $h \in SO(3, 1)$ represents the exact (unbroken) Lorentz symmetry.

Now we proceed from the linear gauge field $\Omega$, the generalized affine connection given in (5.10), to the nonlinear gauge field via (5.20) with the help of the auxiliary fields

$$\begin{pmatrix} 1 & \xi \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} u & \xi \\ 0 & 1 \end{pmatrix}.$$  

A straightforward computation yields

$$A = \begin{pmatrix} u & \xi \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \tilde{\omega} & \tilde{\varphi} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u & \xi \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} u & \xi \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \tilde{\omega} & \tilde{\varphi} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u & \xi \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \tilde{\omega} & \tilde{\varphi} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u & \xi \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \tilde{\varphi} + d\xi + \tilde{\omega} \xi \\ 0 \end{pmatrix} + \begin{pmatrix} u & \xi \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \tilde{\omega} & \tilde{\varphi} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u & \xi \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \tilde{\varphi} + d\xi + \tilde{\omega} \xi \\ 0 \end{pmatrix}.$$  

As a result, we finally find that the rotational part of the gauge gravitational field $\Gamma := u^{-1} \omega u + u^{-1} du$ is realized linearly with respect to the Lorentz group $L_6$ and defines the local Lorentz connection on $M_4$, whereas the tetrad field (the canonical form or the 1-form of coframe) $\theta := u^{-1} \theta = u^{-1} (\tilde{\varphi} + d\xi + \tilde{\omega} \xi)$ is interpreted as the nonlinear gauge field with an additional contribution of the Goldstone fields $u \in GL(4, R)/SO(3, 1)$, transforming homogeneously under the action of translations.

In this way, the Riemann-Cartan (see Chapter 1) geometry $\mathcal{U}_4$ naturally arises on the spacetime manifold $M_4$, where the metric (tetrads) and the connection
are expressed in terms of the basic dynamical variables of the gauge $\mathcal{P}_{10}$-theory of gravity – the generalized affine connection $\Omega$, or the equivalent set of the nonlinear gauge fields $\Gamma$ and $\theta$ and the Goldstone variables $u$, $\xi$.

**Nonlinear gauge theory of de Sitter group**

The Poincaré group $\mathcal{P}_{10}$ is not semi-simple. In Chapter 6, we will demonstrate that this fact has the far-reaching consequences when constructing the dynamics of the gauge gravity theory, and actually yields unsatisfactory results. In an effort to improve the situation, we now consider the gauge gravity theory for the nearest semi-simple extension of $\mathcal{P}_{10}$ – the de Sitter group $S_{10}$, which reduces to the Poincaré group via the Wigner-Inonu “contraction”.

Technically, it will be more convenient to work in a vector bundle, and not in the principal one, although one should remember that the gauge field is originally defined as a connection in the principal bundle, and then it uniquely introduces the connection in the associated bundle, which for this reason is also called the gauge field.

De Sitter group $S_{10} = SO(1, 4)$ is defined as a group of motion of the de Sitter space $\Sigma^4$ which is a 4-dimensional pseudo-Riemannian space of constant negative curvature. The latter is usually represented as a hypersurface in $\mathbb{R}^{1,4}$, the 5-dimensional pseudo-Euclidean space with the metric $\eta_{AB} = \text{diag}(-1, 1, 1, 1, 1)$ (the indices $a, b, c, d, \ldots = 0, 1, 2, 3, 4$ label 5-dimensional components). The de Sitter space is defined by the equation

$$\eta_{AB} x^A x^B = \ell^2,$$

where $\{x^A\}$ are the global Cartesian coordinates in $\mathbb{R}^{1,4}$, and $\ell > 0$ is the curvature radius of the de Sitter space. The equivalent form of (5.26) reads

$$\eta_{ab} x^a x^b + (x^4)^2 = \ell^2,$$

in terms of the 4-dimensional Minkowski metric $\eta_{ab} = \text{diag}(-1, 1, 1, 1)$; the lower case indices $a, b, c, \ldots = 0, 1, 2, 3$.

The group $SO(1, 4)$ is ten-parametric, and its generators $M_{AB} = -M_{BA}$ satisfy commutation relations

$$[M_{AB}, M_{CD}] = \eta_{AD} M_{BC} - \eta_{BD} M_{AC} - \eta_{AC} M_{BD} + \eta_{BC} M_{AD}. \quad (5.27)$$

It is convenient to decompose the generators $M_{AB}$ into two groups: $J_{ab} = M_{ab}$ and $P_a = \ell^{-1} M_{4a}$. Then we rewrite (5.27) as

$$[J_{ab}, J_{cd}] = \eta_{ad} J_{bc} - \eta_{bd} J_{ac} - \eta_{ac} J_{bd} + \eta_{bc} J_{ad},$$

$$[P_a, P_b] = -\frac{1}{\ell^2} J_{ab}, \quad [P_a, J_{bc}] = \eta_{ab} P_c - \eta_{ac} P_b, \quad \left\{ (5.28) \right\}$$

which shows that the Lie algebra $\mathcal{G} = so(1, 4)$ is decomposed into the sum $\mathcal{G} = \mathcal{H} \oplus V$, where $\mathcal{H}$ is the subalgebra formed by six generators $J_{ab} = -J_{ba}$.
isomorphic to the algebra $\mathfrak{so}(1, 3)$ of the Lorentz group; $V$ is the 4-dimensional vector space spanned by $P_\alpha$.

The de Sitter space $\Sigma^4$ is homogeneous, since $SO(1, 4)$ acts transitively on $\Sigma^4$. Denoting $0 = \{x^a = 0, a = 0, 1, 2, 3\}$, we choose the point $(0, \ell) \in \Sigma^4$ as the center, its stabilizer is the Lorentz group $SO(1, 3) = L_6$ formed by the $5 \times 5$ matrices $S^A_B = \begin{pmatrix} L^a_b & 0 \\ 0 & 1 \end{pmatrix}$, with $L^a_b \in SO(1, 3)$. Accordingly, $\Sigma^4 = S_{10}/L_6$.

From (5.28) we see that this is the weakly reductive space, since $[\mathcal{H}, V] \subset V$.

Let $E(M, R^{1,4}, S_{10}, P)$ be a vector bundle associated with the principal bundle $P(M, S_{10})$. The action of the structural group $S_{10}$ in $E$ we realize by the $5 \times 5$ matrices with the following parametrization distinguishing the subgroup $L_6$ (we omit the indices to make the formulas more compact):

$$S = S_T S_L, \quad S \in SO(1, 4).$$

Here $S_L = \begin{pmatrix} L & 0 \\ 0 & 1 \end{pmatrix}$, $L \in SO(1, 3)$ describes the purely Lorentz rotations, and $S_T$ is the matrix of the de Sitter boost mapping the vector $(0, \ell) \in \Sigma^4$ into an arbitrary point of the de Sitter space $V^A = \begin{pmatrix} V^a \\ V^4 \end{pmatrix} \in \Sigma^4$, with $V^A = \ell t^A$, $\eta_{AB} V^A V^B = \ell^2$. Explicitly:

$$S_T = \begin{pmatrix} \delta^a_b - t^a t_b (1 + t^4)^{-1} & t^a \\ -t^b & t^4 \end{pmatrix}.$$  \hspace{1cm} (5.30)

The finite rotation $S_T$ can be obtained from the infinitesimal one by the exponential map $(S_T)^A_B = \exp \begin{pmatrix} 0 & \chi^a \\ -\chi_b & 0 \end{pmatrix}$, where $\chi^a$ (with $\chi_b = \eta_{ab} \chi^a$) are the four transformation parameters which determine $t^A$ in (5.30) by the formulas

$$t^a = (\sin \chi) \chi^{-1} \chi^a, \quad t^4 = \cos \chi, \quad \chi^2 = \eta_{ab} \chi^a \chi^b.$$  \hspace{1cm} (5.31)

The linear gauge field for the de Sitter group is identified with the connection in $E$. As an element of the Lie algebra of the de Sitter group $SO(1, 4)$, the connection 1-form has the following structure:

$$\Omega^A_B = \begin{pmatrix} \tilde{\omega}^a_b & \tilde{\theta}^a \\ -\tilde{\theta}^b & \tilde{\theta}_b \end{pmatrix}, \quad \tilde{\omega}_{ab} = -\tilde{\omega}_{ba}, \quad \tilde{\theta}_b = \eta_{ab} \tilde{\theta}^a.$$  \hspace{1cm} (5.31)
5.4. Kinematics of the gauge gravity theory

However, this decomposition is invariant only with respect to the Lorentz group, whereas under the action of the full $S_{10}$ transformations (5.29), the fields $\tilde{\omega}$ and $\tilde{\theta}$ are mixed. Therefore, one cannot identify $\tilde{\omega}$ with the connection in the tangent bundle, and $\tilde{\theta}$ with the fundamental form. To obtain the geometrical structure on the base $M_4$, it is necessary to take into account that the $S_{10}$-theory is not an “internal” one, and to perform a soldering of the bundle to the spacetime manifold. For this purpose, as we observed from the example of the Poincaré group above, we have to choose the cross-section of the $(G/H)$-bundle, i.e., the Goldstone field, and proceed to the nonlinear realization of the group $G$. In this case, $G = SO(1, 4)$, $H = SO(1, 3)$, $G/H = \Sigma^4$ and the Goldstone field is actually the 5-vector $t^A = (t^a, t^4)$, instead of which however we will (more rigorously) consider the de Sitter boost (5.30).

Following the general framework of Sec. 5.3., we construct the nonlinear gauge field using $\xi = S_T$:

$$A = \xi^{-1}\Omega\xi + \xi^{-1}d\xi = \begin{pmatrix} \Gamma^a_b & \theta^a \\ -\theta^a & 0 \end{pmatrix},$$

(5.32)

where $\theta^a = \eta_{ab}\theta^b$, and we find explicitly

$$\Gamma^a_b = \tilde{\omega}^a_b + \frac{t^{a}Dt_b - t_bDt^a}{1 + t^4},$$

(5.33)

$$\theta^a = t^4\tilde{\theta}^a + Dt^a - \frac{t^a(dt^4 - \tilde{\theta}^b_t^b)}{1 + t^4}.$$  

(5.34)

Here we denoted $Dt^a = dt^a + \tilde{\omega}^a_t^b t^b$.

Under the action of the de Sitter group $S_{10}$, which is now realized nonlinearly, the gauge fields are transformed as

$$\Gamma'^{a}_{b} = L'^{c}_{a} \Gamma^{d}_{c} L'^{-1}_{d} b + L'^{c}_{a} d L'^{-1}_{b}, \quad \theta'^{a} = L'^{a}_{b} \theta^{b},$$

where $S_L' = \begin{pmatrix} L'^{a}_{b} & 0 \\ 0 & 1 \end{pmatrix}$, and the nonlinear function $L' = L'(t^A, S)$ is determined from $S\xi = S S_T = S'TS'L$.

Thus, we can finally identify the 1-form $\Gamma^a_b$ with the local Lorentz connection, and the 1-form $\theta^a$ with the canonical form. As a result, the Riemann-Cartan geometry naturally arises on $M_4$. Note that, in general, the field $\theta^a$ depends on the additional Goldstone variables, corresponding to the embedding of the de Sitter group $S_{10}$ into the general linear group $GL(5, \mathbb{R})$, similarly to the $\mathcal{P}_{10}$-theory. However, we will not distinguish them explicitly, since we are more interested in the specific features of the $S_{10}$-theory, which are important for the construction of the consistent dynamical scheme.
Conclusions: the general scheme of the gauge gravity theory

Let us summarize the essential points of the gauge approach to the gravity theory, distinguishing it from the theory of internal symmetries.

1. The spacetime \( M_4 \) is not considered as a fixed background. The determination of its geometrical structure is a main issue of the gauge-theoretic scheme.

2. The gauge field (potential) in the gravity theory is the connection in the principal bundle \( P(M, G) \).

3. The principal bundle of the gauge gravity theory is not abstract, but is realized as a certain structure on \( M_4 \). For the \( \mathcal{P}_{10} \)-theory, this is the bundle of affine frames \( A(M) \).

4. The specific realization of bundles in the gauge gravity theory implies their soldering to the base. There exists a cross-section of the \((G/H)\)-bundle and the soldering 1-form is defined by the translational part of the connection in \( P(M, G) \).

5. The soldering establishes a homomorphism of the \( P(M, G) \) bundle structure onto the tangent bundle \( TM \), thereby introducing the geometrical structure on the spacetime \( M_4 \), with the metric and connection defined in terms of the gauge gravity field.

6. The explicit construction of the geometrical spacetime structure is performed in terms of the connection in \( P(M, G) \) in the nonlinear realization of the gauge group. The Goldstone fields, arising in the reduction \( G \to H \) (the latter being closely related to the soldering), also contribute to the metric and the linear connection in \( TM \).

It is worthwhile to note that the group of general coordinate transformations is not related directly to the construction of the gauge gravity theory. However, the gauge transformations of special form may induce the infinitesimal transformations of coordinates on \( M_4 \), if we introduce the latter with the help of the reduction map \( \varphi \) defined in (5.9). Indeed, using the cross-section \( \sigma : M \to A(M) \), we can define the local coordinates \( x^\mu \) as \( x = \varphi \circ \sigma \). Then under the action of translations \( T = \begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix} \), the cross-section \( \sigma \to R_T \sigma \), and we obtain \( x^\mu \to x^\mu + a^\mu(x) \) as the local general coordinate transformations.

Within the framework described above, the Riemann structure of GR cannot be obtained without additional restrictions of the natural geometrical structure arising on the spacetime. Such restrictions (for example, a priori putting the torsion equal zero) contradict the consistent gauge-theoretic scheme, and they do not follow from any (even non-gauge-theoretic) fundamental physical principles. In this relation, it is instructive to recall Einstein’s words, that “...the question whether this continuum has a Euclidean, Riemannian, or any other structure is a question of physics proper which must be answered by experience, and not a question of a convention to be chosen on grounds of mere expediency.” [39].
6  

Gauge gravity models with dynamical $(\Gamma-S)$-interaction

6.1. Principles for construction of dynamics

In the previous Chapter, we saw that the gauge approach based on the Poincaré group $\mathcal{P}_{10} = T_4 \supset L_6$ in general yields a non-Einstein theory of gravity. One can choose the tetrad field $h_\mu^a$ and the local Lorentz connection $\Gamma_{b\mu}^a$ field on $M_4$ as the main dynamical variables of the $\mathcal{P}_{10}$-theory. In accordance with the general gauge scheme, their source is the $\mathcal{P}_{10}$-current consisting of the canonical energy-momentum tensor $t_\mu^a$ and the canonical spin tensor $S^{\mu a}_b$, which correspond to subgroups $T_4$ and $L_6$, respectively. As compared to GR, the new feature of the $\mathcal{P}_{10}$-theory is the existence of an additional $(\Gamma_{b\mu}^a-S^{\mu a})$-interaction, along with Einstein’s $(h_\mu^a-t_\mu^a)$-interaction. Since the Poincaré group is not simple, these interactions are characterized by the two different coupling constants — one with the dimension $[l_0^2] = m^2$ for translations and the dimensionless one $\lambda$ for Lorentz rotations.

The prediction of a new type of the gravitational interaction (spin-connection) is the main physical consequence of the gauge gravity theory for the Poincaré group.

The realization of this interaction in the Einstein-Cartan model was reviewed in Chapters 3 and 4. However, one cannot consider the ECT as a satisfactory gauge gravity theory. The Lagrangian $\mathcal{L}_{ECT} \sim R(\Gamma) \sim (\partial \Gamma + \Gamma \partial) hh$ is non-dynamical both for $\Gamma$ and for $h$, in the sense that the field equations do not contain derivatives of the gauge field strengths, unlike the Maxwell or Yang-Mills theories. Due to the contact nature of the $(\Gamma-S)$-interaction, one can eliminate the torsion from the theory, thereby arriving at the effective non-renormalizable
$S^2$-self-interaction of matter. The degeneracy of ECT is also manifest in the absence of the second coupling constant for the $(\Gamma-S)$-interaction, as result of which the spin effects turn out to be negligibly small. Another sign of degeneracy is the mismatch of the gauge kinematics and the dynamics of the model: we have $R \sim t$, $Q \sim S$ in ECT field equations, linking the $L_6$ field strength (the curvature) to the $T_4$ current (the energy-momentum), and the $T_4$ field strength (the torsion) to the $L_6$ current (the spin).

The consistent nondegenerate dynamical scheme of the gauge gravity theory should be free of such deficiencies. However, its construction is complicated due to the essential non-uniqueness of theory’s action, which is a consequence of the nonlinear realization of the gauge symmetry that allows for non-polynomial (in the nonlinear fields $h^a_{\mu}$ and $g_{\mu\nu}$) interactions. We can try to overcome this difficulty by supplementing the gauge theory with a number of conditions (natural from physical and geometrical viewpoints) that help to select the basic Lagrangian of the theory.

The appropriate general requirements can be formulated as follows:

1. Correspondence principle: in the absence of spin and torsion, the field equations should allow for a limiting transition to GR.
2. Both gravitational interactions, $(h-t)$ and $(\Gamma-S)$, should be dynamical.
3. The dynamics of the theory should be consistent with the gauge kinematics in the sense that the source of the gauge field of a certain spacetime symmetry is the corresponding Noether current.

Let us apply these criteria to analyze the possible choices of the Lagrangians (more exactly, of the action functionals) of the gravitational theory considered earlier in the literature. To keep the correct dimension of the action, $[S] = [h]$, we write down all the dimensional factors explicitly.

a) $S_{GR} = \frac{1}{2\kappa c} \int d^4x \sqrt{g} R$, $S_{ECT} = \frac{1}{2\kappa c} \int d^4x \sqrt{g} R(\Gamma)$.
These models with the Lagrangians linear in the curvature do not satisfy the requirements 2 and 3.

b) $S_{YM} = -\frac{\hbar}{4} \int d^4x \sqrt{g} R_{\mu\nu}(\Gamma) R_{\mu\nu}(\Gamma)$.
The Yang-Mills Lagrangian obviously provides the fulfilment of the requirements 2 and 3. However, this model has two serious problems: the absence of GR limit and the conformal invariance of $S_{YM}$ which imposes a strong constraint on matter (the vanishing of the trace of the energy-momentum tensor). Therefore, this model can be viewed only as a microscopic limit of a fundamental theory.

c) $S_0 = S_{GR} + \beta S_{YM}, \quad \beta = \text{const}$.
This choice does not completely satisfy the requirement 3, since this model arises in the gauge theory of the Lorentz group with the assumption that the metric is a priori defined on the base.

d) $S_1 = S_{ECT} + \beta S_{YM}, \quad \beta = \text{const}$.
Also in this model the requirement 3 is not fulfilled in the sense that the variation with respect to the tetrads yields non-dynamical equations.

e) $S_2 = -\frac{\alpha}{2\kappa c} \int d^4x \sqrt{g} Q^\alpha_{\mu \nu} Q^\mu_{\nu} + \beta S_{YM}, \quad \alpha, \beta = \text{const}$.
This model does not satisfy the principle of correspondence with GR. However, one can come up with a modification
\[
S_3 = \frac{1}{2\kappa c} \int d^4 x \sqrt{g} (\alpha Q^\mu_\mu + 2Q^\mu Q^\mu) + \beta S_{YM},
\]
that contains the Schwarzschild solution as a torsionless limiting case.

f) The model, satisfying all the requirements above, is given by the action
\[
S_P = \int d^4 x \sqrt{g} \left\{ \frac{1}{2\kappa c} (R(\Gamma) - 2\alpha Q^a_{\mu\nu} Q^a_{\mu\nu}) - \frac{1}{4\lambda} R^{ab}_{\mu\nu} (\Gamma) R_{\mu\nu}^{ab} (\Gamma) \right\}. \tag{6.1}
\]
Here \(\alpha\) is a dimensionless coupling constant, and the dimension of \(\lambda = \frac{1}{\hbar}\).

Below we demonstrate, how this dynamical framework naturally arises in \(\mathcal{P}_{10}\)- and in \(S_{10}\)-theory, so that one can avoid a nontrivial analysis of all possible Lagrangians quadratic in the curvature and torsion with arbitrary coefficients.

6.2. Choice of the action in the gauge gravity theory

**Difficulties of the Poincaré gauge theory**

Let \(\Omega\) be the 1-form of the generalized affine connection on \(M_4\) (we assume a cross-section \(\sigma: M \rightarrow O(M)\)), interpreted as the linear \(\mathcal{P}_{10}\)-gauge gravitational field. The field strength is identified with the curvature 2-form
\[
\mathcal{R} = d\Omega + \Omega \wedge \Omega = \begin{pmatrix} \bar{R} & \bar{\Theta} \\ 0 & 0 \end{pmatrix},
\]
where \(\bar{R} = \bar{\omega} + \bar{\omega} \wedge \bar{\omega}, \bar{\Theta} = \bar{\theta} + \bar{\omega} \wedge \bar{\theta}\). However, the latter is not invariant with respect to the decomposition of the curvature into the translational and the Lorentz parts, and to improve this we proceed to the nonlinear realization. The nonlinear gauge field
\[
A = \begin{pmatrix} \Gamma^a_b \\ \theta^a \end{pmatrix}
\]
(6.2)
encompasses the local Lorentz connection \(\Gamma^a_b\) (the \(L_6\) gauge field) and the tetrad field \(h^a\) (the \(T_4\) nonlinear gauge field) is introduced by \(\theta^a = \frac{1}{l_0} h^a\), where the factor \(l_0\) is included to get the correct dimension, \([h^a] = 1\). The curvature 2-form
\[
\mathcal{R} = \begin{pmatrix} R^a_{\mu\nu} & \frac{1}{l_0} \hat{Q}^a \\ 0 & 0 \end{pmatrix}, \quad R^a_{\mu\nu} = d\Gamma^a_{\mu\nu} + \Gamma^a_c \wedge \Gamma^c_{\mu\nu}, \quad \hat{Q}^a = dh^a + \Gamma^a_b \wedge h^b \tag{6.3}
\]
consists of the Lorentz curvature 2-form \(R^a_{\mu\nu}\) and the torsion 2-form \(\hat{Q}^a\), which are the \(L_6\) and the \(T_4\) field strengths of the \(\mathcal{P}_{10}\) gauge field, respectively. Note that \(\hat{Q}^a = \frac{1}{2} \hat{Q}^a_{\mu\nu} dx^\mu \wedge dx^\nu\) with \(\hat{Q}^a_{\mu\nu} = -2Q^a_{\mu\nu}\), cf. (1.3).
One can construct the dynamics of the \( \mathcal{P}_{10} \)-theory by two methods:

1. We can use the irreducible parts \( \Gamma \) and \( h \) of the nonlinear gauge field \( (6.2) \) to derive the covariant objects \( (6.3) \) and then to build the Lagrangian as an arbitrary polynomial of invariants formed from the curvature, torsion and tetrad, \( R, Q, h \). This procedure is essentially non-unique.

2. Using the analogy with the Yang-Mills fields of the internal symmetries, one can try to construct the Lagrangian depending on \( A \) (and \( R \)) as a whole. In this case, the simplest choice is the Yang-Mills Lagrangian \( \frac{1}{2} \text{Tr}(R \wedge \ast R) \).

The trace operation “\( \text{Tr} \)” refers to the \( \mathcal{P}_{10} \)-“internal” indices, and the Hodge operator (dualization, \( \ast \)) is defined with the help of the metric \( g_{\mu \nu} = h_a^{\mu} h_b^{\nu} \eta_{ab} \), resulting from the nonlinear realization of the \( \mathcal{P}_{10} \) group. Since the Poincaré group is semi-simple, its Cartan-Killing metric (encoded in the trace operation \( \text{Tr} \)) is degenerate, and as a result, the variational derivation of the field equation does not commute with the decomposition of \( A \) and \( R \) into the \( L_6 \)- and \( T_4 \)-irreducible parts. Indeed, if we make such decomposition at first, then

\[
\frac{1}{2} \text{Tr}(R \wedge \ast R) = \frac{1}{2} \text{Tr}\left( \begin{pmatrix} R & \frac{1}{2} \tilde{Q} \\ 0 & 0 \end{pmatrix} \right) \wedge \left( \begin{pmatrix} \ast R & \frac{1}{2} \ast \tilde{Q} \\ 0 & 0 \end{pmatrix} \right) = \frac{1}{2} \text{Tr}(R \wedge \ast R),
\]

hence the torsion completely drops out, and the remaining Lagrangian is described as the case b) in Sec. 6.1.. The resulting field equations are unsatisfactory for the macroscopic gravitation theory [100].

On the other hand, if the action \( \frac{1}{2\lambda} \int \text{Tr}(R \wedge \ast R) + S_{\text{mat}} \) is formally varied with respect to \( A \), we find the field equation

\[
\mathcal{D} \ast \mathcal{R} \equiv d \ast \mathcal{R} + A \wedge \ast \mathcal{R} - \ast \mathcal{R} \wedge A = \lambda \ast J. \tag{6.4}
\]

Here the \( \mathcal{P}_{10} \)-current is obtained as a variational derivative \( \ast J := \delta S_{\text{mat}} / \delta A \) of the matter action with respect to \( A \). The structure of the source reflects its relation to the Poincaré group,

\[
J = \begin{pmatrix} S_{ab} & l_0 t^a \\ 0 & 0 \end{pmatrix}, \tag{6.5}
\]

where the 1-forms \( S_{ab} \) and \( t^a \) are the \( L_6 \)- and \( T_4 \)-irreducible parts of \( J \), respectively. With an account of (6.2) and (6.3), we recast (6.4) into the \( L_6 \)- and \( T_4 \)-components:

\[
\mathcal{D} \ast R_a^b \equiv d \ast R_a^b + \Gamma_c^a \wedge \ast R_c^b - \ast R_c^a \wedge \Gamma_c^b = \lambda \ast S_{ab}, \tag{6.6}
\]

\[
\mathcal{D} \ast \tilde{Q}^a - \ast R_a^b \wedge h^b = \lambda l_0^2 \ast t^a. \tag{6.7}
\]

Let us investigate the vacuum solutions \( (S_{ab} = 0, \ t^a = 0) \). We switch to the coordinate notation in components and use the decomposition (1.12) to write
the field equations (6.6) and (6.7) in terms of the metric and contortion tensors \((g_{\mu\nu}, T^\lambda_{\mu\nu})\) instead of the original variables \((h^a_{\mu}, \Gamma^a_{\mu\nu})\):

\[
R_\alpha^\beta \mu\nu(\Gamma) \cdot_\nu + T^\alpha_{\sigma\nu} R^\sigma_{\beta\mu\nu}(\Gamma) - T^\sigma_{\beta\nu} R^\sigma_{\alpha\mu\nu}(\Gamma) = 0,
\]  
(6.8)

\[
T^{a[\nu\beta]} \cdot_\nu + T^a_{\sigma\nu} T^{a[\nu\beta]} + \frac{1}{2} R^{\alpha\beta}(\Gamma) = 0.
\]  
(6.9)

Recall that the semicolon \(\cdot\) denotes the Riemannian covariant derivatives defined by the Christoffel symbols.

The field equations (6.6) and (6.7) is the system of differential second order equations for the tetrad and the local Lorentz connection. Therefore, both \((h-t)\) and \((\Gamma-S)\)-interactions are dynamical in this model. Decomposing the curvature tensor into the Riemannian (depending only on the metric) part and the post-Riemannian (torsion-dependent) part, we find the system (6.8) and (6.9) of differential equations of the second order for the metric and contortion \((g_{\mu\nu}, T^\lambda_{\mu\nu})\).

When the torsion equal to zero, \(T^\alpha_{\mu\nu} = 0\), the vacuum system (6.8) and (6.9) reduces to the Einstein equation \(R_{\alpha\beta} = 0\). But in general case, the torsion is propagating, and thus the \((\Gamma-S)\)-interaction has a non-zero effective radius, in contrast to ECT.

The physical interpretation of the sources in the right-hand sides of the field equations (6.6) and (6.7) is a difficult issue. This primarily refers to (6.7), since the variational derivation of the equation (6.6), corresponding to \(L_0\)-subgroup, does not depend on the order of separation of translations in \(P_{10}\). Accordingly, the source \(S^a_{\mu\nu}\) can be consistently identified with the canonical spin density tensor of matter [99, 100]. However, \(t^a_{\mu}\) in (6.5) cannot be identified with the energy-momentum tensor, since when deriving the system (6.4), the dependence of the metric \(g_{\mu\nu}\) (in the Hodge operator \(\ast\)) on the translational field \(\theta\) was not taken into account (recall that the formal variation of the gravitational and matter actions was performed only with respect to the gauge field \(A\) as a whole).

Therefore, although the second method of construction of the \(P_{10}\)-theory dynamics allows to highlight some important features typical for the gauge gravity theory (such as the correspondence with GR, the dynamical nature of all interactions, and the partial consistency of the dynamics with the kinematics), it does not lead to the physically satisfactory results. This is related to the fact that the Poincaré group \(P_{10}\) is not non-semisimple.

Thus, we have to turn to the first method to construct the dynamics of the \(P_{10}\)-theory. However, in this case, there are too many arbitrary coupling constants in the theory. For their concretization and for the fixation of the structure of the gravitational Lagrangian, we now analyse the construction of the gauge theory dynamics for de Sitter group \(S_{10}\) – the nearest semisimple extension of \(P_{10}\)
Dynamics of $S_{10}$-theory of gravity

In the discussion of kinematics of the gauge gravity theory in Sec. 5.4., we identified the gauge potential for de Sitter group theory with the nonlinear connection (5.32):

$$A = \begin{pmatrix}
\Gamma^a_b & \frac{1}{l_0^2} h^a_b \\
-\frac{1}{l_0} h^a_b & 0
\end{pmatrix}.$$

The corresponding gauge field strength is identified with the curvature 2-form

$$\mathcal{R} = dA + A \wedge A = \begin{pmatrix}
R^a_b - \frac{1}{l_0^2} \pi^a_b & \frac{1}{l_0} \hat{Q}^a \\
-\frac{1}{l_0} \hat{Q}_b & 0
\end{pmatrix},$$

where $\pi^a_b = h^a \wedge h_b = \frac{1}{2} \pi^a_{b\mu\nu} dx^\mu \wedge dx^\nu$. Obviously, $\pi^a_{b\mu\nu} = h^a_b \pi^a_{b\nu} - h^a_b \pi^a_{b\mu}$.

Since the de Sitter group is semisimple, its Cartan-Killing form is non-degenerate, and one can choose the action of the theory in the standard Yang-Mills form

$$S_{dS} = \frac{1}{4\lambda} \int d^4x \sqrt{g} \text{Tr}(\mathcal{R} \wedge \ast \mathcal{R}).$$

With an account of (6.10), we have

$$S_{dS} = \int d^4x \sqrt{g} \left\{ \frac{1}{2\kappa c} [R(\Gamma) - 2\Lambda - 2Q^a_{\mu\nu}Q_a^{\mu\nu}] - \frac{1}{8\lambda} R_{\mu\nu}^{ab}R_{\mu\nu}^{ab} \right\},$$

where we redefined the coupling constants as $\kappa c = l_0^2 \lambda, \Lambda = \frac{3}{l_0^2}$.

The action (6.11) is invariant with respect to the nonlinear realization of the de Sitter group $S_{10}$. Its structure coincides, up to a $\lambda$-term, with that of $Sp(6.1)$ with $\alpha = 1$. Therefore, the non-degeneracy of the Cartan-Killing form for $S_{10}$ makes it possible to construct the gauge theory dynamics that satisfies the requirements 1-3 formulated in Sec. 6.1.

From the action (6.11), the equations of the gravitational field are derived with the help of variation with respect to the local Lorentz connection $\Gamma^a_{b\mu}$ and the tetrad $h^a_\mu$. In vacuum (in the absence of matter) they read

$$-\frac{1}{2\lambda} \mathcal{D}_a R_{\mu\nu}^{ab}(\Gamma) - \frac{1}{\kappa c} \left( Q^a_{\mu\nu} + h^a_\mu Q_b - h^a_\nu Q_a + 2Q_{[ab]}^{\mu\nu} \right) = 0,$$

$$\frac{1}{\kappa c} \left( 2\mathcal{D}_a Q_a^{\mu\nu} - R^a_{\mu\nu}(\Gamma) + \frac{1}{2} h^a_\mu R(\Gamma) - \Lambda h^a_\mu \right) + \frac{c}{\kappa} \tau_a^{\mu\nu} = 0,$$

where modified covariant derivative $\mathcal{D}_a$ is defined using $\Gamma^a_{b\mu}$ for the Lorentz (tetrad) indices and using $\{^\lambda_{\mu\nu}\}$ for the world (coordinate) indices $(\mu, \nu, \ldots)$.

The tensor $\tau_a^{\mu\nu}$ in (6.13) reads

$$\tau_a^{\mu\nu} = \frac{c}{2\lambda} \left( R_{\alpha\beta}^{b\nu}(\Gamma) R_{\nu\mu}^{\alpha\beta}(\Gamma) - \frac{1}{4} h^a_\mu R_{\alpha\beta}^{b\nu}(\Gamma) R_{\nu\mu}^{\alpha\beta}(\Gamma) \right) + \frac{4}{\kappa} \left( Q_{b\mu\nu} Q^{b\mu\nu} - \frac{1}{4} h^a_\mu Q^a_{b\alpha} Q_b^{b\alpha} \right).$$
This is the energy-momentum tensor of $S_{10}$-gauge field, that takes into account the influence on the spacetime geometry of the gravitational field itself. The field equation (6.13) differs from (6.7) by the term $\tau^a$, making the physical meaning of the nonlinear field equation (6.13) differs from (6.7) by the term the influence on the spacetime geometry of the gravitational field itself. The $\partial^a$ coupling principle (as a result of the gauge approach) by means of a substitution momentum of matter, which coincide with (1.47) and (1.45) when we notice that these are the conservation laws of the total angular momentum and the energy-momentum of matter, that coincide with (1.47) and (1.45) when we notice that $\partial^a \rightarrow \h^a$. This amounts to supplementing the action $S_{dS}$ by the interaction term

$$S_{int} = -\int d^4x \sqrt{g} \mathcal{L}_{int},$$

$$\mathcal{L}_{int} = -\Gamma^a_{\mu\nu} S^b_{\mu} + \frac{1}{c} t^a \h^a,$$

$$J = \left( \begin{array}{cc} S^a_{\mu} & \frac{1}{2} t^a \\ -\frac{1}{2c} t^b & 0 \end{array} \right).$$

Variation of total action $S_{dS} + S_{int}$ with respect to $\Gamma^a_{\mu\nu}$ and $\h^a$ gives

$$\mathcal{D}_\mu R_{\mu\nu}(\Gamma) + \frac{2}{g} (Q^a_{\mu\nu} + \h^a Q_b + \h^a Q_a + 2Q_{[ab]}^a = 2\lambda S^a_{\mu\nu},$$

$$-2\mathcal{D}_\mu Q^a_{\mu\nu} + R_{\mu\nu} = \frac{1}{2} \h^a R(\Gamma) + \Lambda \h^a = \kappa (t^a + \tau^a).$$

This is the final system of the field equations of the $S_{10}$-gauge gravity theory.

In terms of the metric and contortion, the equations (6.16), (6.17) read

$$\frac{\ell^D}{2} \left\{ R_{\alpha\beta\mu\nu}(\Gamma) + T_{\alpha\beta\mu\nu}(\Gamma) + T_{\lambda\beta\mu\nu}(\Gamma) \right\}$$

$$+ 3g^{\mu\nu} T_{\alpha\beta\mu\nu} + \delta^\mu_{\alpha} T_{\beta} = \kappa c S^a_{\alpha\beta},$$

$$R_{\mu\nu} = \frac{2}{g} \epsilon_{\mu\nu} R + \epsilon_{\mu\nu} \Lambda - T_{(\mu\nu)} + g_{\mu\nu} T_{(\lambda\mu\nu)} - \frac{3}{2} T_{(\alpha\beta)} T^{\alpha\beta} + \frac{1}{2} g_{\mu\nu} (T_{\lambda} T^{\lambda} - 3T_{(\lambda\mu\nu)} T^{\lambda\mu\nu})$$

$$- \frac{\ell^D}{2} \left\{ R_{\alpha\beta\lambda}(\Gamma) - \frac{1}{2} g^{\mu\nu} R_{\alpha\beta\mu\nu}(\Gamma) R_{\alpha\beta\rho\sigma}(\Gamma) \right\} = \kappa t_{(\mu\nu)},$$

$$3T_{(\mu\nu\lambda)} - T_{[\mu\nu\lambda]} + T_{\lambda} T_{(\mu\nu\lambda)} + 3T_{(\alpha\mu\nu)} T^{\alpha\mu\nu} - 3T_{(\beta\mu\nu)} T^{\beta\mu\nu} = \kappa t_{[\mu\nu]}.$$

Taking the covariant derivatives of the field equations (6.16) and (6.17), after some algebra we find

$$t_{[\alpha\beta]} = c \mathcal{D}_\mu S^\mu_{\alpha\beta},$$

$$\mathcal{D}_\mu t^\mu_{\alpha} + 2Q^b_{\alpha\mu} t^\mu_{b} + c S^\mu_{\alpha\beta} R^{cd}_{\mu\alpha} = 0.$$
When the matter is spinless and the matter Lagrangian does not depend on the tetrad otherwise than in terms of the metric, (6.22) reduces to the ordinary Einstein’s conservation law of the energy-momentum tensor.

### 6.3. Physical consequences of the gauge gravity theory

The study of the $S_{10}$-theory generalizing the gauge approach based on the Poincaré group, allows one to fix the structure of gravitational field action via (6.10) and (6.11).

Now we investigate the physical content of the gauge gravity theory with the dynamical ($\Gamma$-$S$)-interaction described by the system (6.16) and (6.17). It is worthwhile to note that these field equations represent a particular case of the general gauge gravity model with the action

\[
S_g = \int d^4x \sqrt{g} \mathcal{L}_g, \quad \mathcal{L}_g(h^a, Q^a, R^{ab}) \text{ is an arbitrary scalar function of the tetrad (the metric), and the gravitational field strengths – the curvature and the torsion.}
\]

The general field equations read

\[
\mathcal{D}_\nu H_{ab}^{\mu\nu} - E_{ab}^{\mu} = S_{ab}^{\mu},
\]

\[
\mathcal{D}_\nu H_a^{\mu\nu} - E_a^{\mu} = \frac{1}{c} t_a^{\mu}.
\]

Here the Lorentz and the translational gauge momenta are defined by

\[
H_{a}^{\mu\nu} := 2 \frac{\partial \mathcal{L}_g}{\partial R^a_{\mu\nu}}, \quad H_a^{\mu\nu} := \frac{\partial \mathcal{L}_g}{\partial Q^a_{\mu\nu}},
\]

and the gravitational energy-momentum and spin are introduced as

\[
E_a^{\mu} := h_a^{\mu} \mathcal{L}_g - 2Q^b_{a\mu}H_b^{\mu\nu} - R^c_{bav}H_c^{b \mu\nu}, \quad E_{ab}^{\mu} := H_{[ab]}^{\mu}.
\]

**Correspondence with GR**

Let us investigate the issue of the correspondence of the vacuum field equations (6.18)-(6.20) and the Einstein field equations under the assumption of vanishing torsion. For the sake of generality, we do not confine ourselves to the de Sitter gravity with the action (6.17), but study the most general Lagrangian without the parity-violating terms that leads to the second-order equations for the gauge gravitational fields ($h_{\mu}^a$, $\Gamma_{a\mu}^\nu$):

\[
\mathcal{L}_g = \frac{1}{2\kappa} \left\{ a_0 \mathcal{R}(\Gamma) - 2\Lambda + a_1 Q_{a\mu\nu}Q^{a\mu\nu} + a_2 Q_{a\mu\nu}Q^{\alpha\mu\nu} + a_3 Q^a \right\} + b_1 R_{a\beta\mu\nu}(\Gamma)R^{a\beta\mu\nu}(\Gamma) + b_2 R_{a\beta\mu\nu}(\Gamma)R^{\mu\nu\alpha\beta}(\Gamma) + b_3 R_{a\beta\mu\nu}(\Gamma)R^{a\beta\mu\nu}(\Gamma) + b_4 R_{a\beta}(\Gamma)R^{a\beta}(\Gamma) + b_5 R_{a\beta}(\Gamma)R^{\beta\alpha}(\Gamma) + b_6 R^2(\Gamma).
\]

We have the 4 dimensionless coupling constants $a_0, a_1, a_2, a_3$, and the 6 coupling constants $b_1, \ldots, b_6$ with the dimension $[h]$, in addition to the cosmological term $\Lambda$. When $a_0 = 0$, we have the purely quadratic model.
Not all terms in the Lagrangian (6.23) are independent, since the expression
\[
\sqrt{g} \left\{ R_{\alpha\beta\mu\nu}(\Gamma) R^{\mu\nu\alpha\beta}(\Gamma) - 4 R_{\alpha\beta}(\Gamma) R^{\beta\alpha}(\Gamma) + R^2(\Gamma) \right\}
\]
is the complete divergence representing the Gauss-Bonnet identity. The integral of this scalar density describes the topological invariant. Therefore, any of the constants \( b_2, b_3, b_6 \) may be eliminated. However, here this possibility is not used.

Let us analyse the field equations for the Lagrangian (6.23) for the case of the vanishing torsion, \( Q^\lambda_{\mu\nu} = 0 \). Then we find \( H_\alpha^{\mu\nu} = 0 \), hence \( E_{\alpha\mu\nu} = 0 \), and in vacuum \((S^\mu_{\alpha\beta} = 0 \text{ and } t^\mu_{\alpha} = 0)\) the field equations reduce to
\[
D_\nu H_{\alpha\mu\nu} = - h_\alpha^a h_b^\beta \left( \varphi_1 C_{\alpha\beta\mu\nu} + \varphi_2 \delta^\beta_a [ R_{\alpha\beta} ] \right) = 0,
\]
\[
- E^{\mu}_{\alpha} = \frac{1}{\kappa c} \left( a_0 R^\mu_{\alpha} - \frac{a_0}{2} R h^\mu_{\alpha} + \Lambda h^\mu_{\alpha} \right) + \varphi_1 C_{\alpha\nu}^{\mu\beta} R^\nu_{\beta} + \varphi_2 R R^\mu_{\alpha} = 0.
\]
Here \( C_{\alpha\beta}^{\mu\nu} \) is the Weyl tensor introduced in the irreducible decomposition of the curvature tensor (1.14), and \( \rightarrow R_{\mu\nu} = R_{\mu\nu} - \frac{1}{4} g_{\mu\nu} R \) is the traceless part of the Ricci tensor. It is worthwhile to mention that the Weyl tensor is double-self-dual, as a result, it satisfies the quadratic identity
\[
C_{\alpha\sigma}^{\mu\nu} C_{\beta\sigma}^{\mu\nu} = \frac{1}{4} \delta^a_\alpha \delta^\beta_\sigma C_{\rho\sigma}^{\mu\nu} C_{\rho\sigma}^{\mu\nu},
\]
which was taken into account in the derivation of the field equations.

The form of the dynamical equations essentially depends on the parameters \( \varphi_1 \) and \( \varphi_2 \) that are constructed from the coupling constants:
\[
\varphi_1 = 2(4\lambda_1 + \lambda_2), \quad \varphi_2 = \frac{4}{3}(\lambda_1 + \lambda_2 + 3\lambda_3),
\]
\[
\lambda_1 = b_1 + b_2 + \frac{1}{2} b_3, \quad \lambda_2 = b_4 + b_5, \quad \lambda_3 = b_6.
\]
The last line determines the structure of the effective Lagrangian obtained from (6.23) for the vanishing torsion, \( Q^\lambda_{\mu\nu} = 0 \):
\[
L_\gamma = \frac{1}{2\kappa c} (a_0 R - 2\Lambda) + \lambda_1 R_{\alpha\beta\mu\nu} R^{\alpha\beta\mu\nu} + \lambda_2 R_{\alpha\beta} R^{\alpha\beta} + \lambda_3 R^2.
\]
When \( \varphi_1 = 0 \) and \( \varphi_2 = 0 \), the field equations reduce to Einstein’s equation with the cosmological term. This happens when \( \lambda_1 = \lambda_3 \) and \( \lambda_2 = -4\lambda_3 \) which corresponds to the Gauss-Bonnet topological action.

The resulting system for all values of \( \varphi_1 \) and \( \varphi_2 \) yields \( a_0 R = 4\Lambda \) (to check this, contract the second equation). Consequently, the torsionless \((Q^\lambda_{\mu\nu} = 0)\) vacuum field equations for the whole class of models (6.23) read
\[
a_0 R = 4\Lambda, \quad \varphi_1 C_{\alpha\beta\mu\nu} R^{\alpha\beta} + \left( \frac{a_0}{\kappa c} + \varphi_2 R \right) R_{\mu\nu} = 0, \quad \varphi_1 \rightarrow R_{\mu[\alpha;\beta]} = 0.
\]
The last equation follows from the Bianchi identity \((1.17)\).

It is obvious that the vacuum Einstein spaces (with \(\lambda\)-term)
\[
a_0 R = 4 \Lambda, \quad R_{\mu \nu} = \frac{1}{4} g_{\mu \nu} R,
\]
are solutions of the system \((6.24a)-(6.24c)\). The question is: are there other solutions or \((6.25)\) represents the unique solution? This is a non-trivial problem \([137, 187]\). Let us analyse the case of a nonvanishing \(a_0 \neq 0\) first; we then can put \(a_0 = 1\) without the loss of generality. There are several situations depending on the values of \(\varphi_1\) and \(\varphi_2\).

1. When \(\varphi_1 = 0\), the system \((6.24a)-(6.24c)\) reduces to
\[
R = 4 \Lambda, \quad \left( \frac{1}{\kappa c} + 4 \varphi_2 \Lambda \right) R_{\mu \nu} = 0.
\]
(6.26)

Then we have one of the two possibilities.

1.1. If \(\frac{1}{\kappa c} + 4 \varphi_2 \Lambda \neq 0\), the system \((6.26)\) coincides with Einstein’s field equations \((6.25)\).

1.2. In the special case \(\frac{1}{\kappa c} + 4 \varphi_2 \Lambda = 0\), equations \((6.24b), (6.24c)\) are satisfied identically, and the solutions are arbitrary spaces subject to the only constraint \(R = 4 \Lambda\).

2. If \(\varphi_1 \neq 0\), we introduce
\[
\xi := \frac{\frac{1}{\kappa c} + 4 \varphi_2 \Lambda}{\varphi_1},
\]
and the system \((6.24a)-(6.24c)\) is then recast into
\[
R = 4 \Lambda, \quad C_{\alpha \beta \mu \nu} R^{\alpha \beta} = -\xi R_{\mu \nu},
\]
(6.27a)
\[
R_{\mu [\alpha ; \beta]} = 0.
\]
(6.27c)

One can prove \([193]\) that the only solutions of the system \((6.27a)-(6.27c)\) are Einstein spaces \((6.25)\), unless \(\xi = 0\), or \(\xi = \frac{2 \Lambda}{3}\), or \(\xi = -\frac{4 \Lambda}{3}\). In these exceptional cases, the solutions of the system \((6.27a)-(6.27c)\) are known with \(R_{\mu \nu} \neq 0\) which are not Einstein spaces \([189]\).

For completeness, let us mention that similar conclusions can be derived for the purely quadratic model with \(a_0 = 0\) in which case \(\xi = \varphi_2 R / \varphi_1\).

The \(S_{10}\)-theory with the action \((6.11)\) belongs to the class of quadratic models \((6.23)\). As one can immediately check, \(\xi = \frac{2 \Lambda}{3}\) in this de Sitter model, and hence it does not belong to the exceptional cases above. Accordingly, we conclude that the only torsionless solutions of the de Sitter theory \((6.11)\) are Einstein spaces \((6.25)\), thus demonstrating the correspondence of the \(S_{10}\)-theory with GR.

Let us estimate the coupling constants \(\lambda\) and \(l_0\) for the \((\Gamma-S)\)-interaction. Using the torsionless solutions \((6.25)\) (Einstein spaces), one can try to identify \(\Lambda\) directly with the cosmological term of GR. According to the modern observations, \(\Lambda \approx 10^{-52}\) m\(^{-2}\), consequently we find \(l_0 \approx 10^{26}\) m, which is of order of
a typical size of the de Sitter world. As a result, assuming that \( \kappa \) is the usual Einstein’s gravitational constant, we obtain for the coupling constant of the (\( \Gamma - S \))-interaction an estimate \( 1/\lambda \approx 10^{120} \hbar \). Thus, we can conclude that the spin of matter under ordinary physical conditions practically does not affect the gravitational field, and with the high accuracy we can treat (6.16) as the vacuum field equation with the zero right-hand side. However, in this case, the last term in (6.11) quadratic in the curvature would dominate at macroscopic distances, which obviously contradicts the idea to identify \( \Lambda \) with the cosmological term. Consequently, such identification and the resulting estimate of the coupling constants are not consistent with observations.

One should expect the coupling constant \( \lambda \) to be much larger (with the upper limit for \( \lambda \) to be fixed by experiment), so that the curvature quadratic term dominates only at small distances, whereas at the macroscopic scales it is essentially smaller than the term \( (R + Q^2) \). Then the resulting huge value of \( \Lambda \) does not allow to interpret it as a cosmological term. The correspondence with Einstein’s equations was derived here in the absence of matter, which is a special physical situation. When \( \pi_{\mu \nu} \neq 0 \), the condition of the vanishing torsion turns out to be too strong even in the absence of spin \( S_{\mu \nu} = 0 \), and it does not lead to reasonable results, in general [100, 129, 140, 141].

**Instanton solutions**

The Yang-Mills structure of the action (6.10) of the de Sitter theory underlies the existence of the self-dual or the instanton solutions. To analyse this important feature, we turn to the Euclidean formulation of the theory. Technically this is done via the replacements \( SO(1, 4) \rightarrow SO(5) \), \( SO(1, 3) \rightarrow SO(4) \), and \( \eta_{ab} \rightarrow \delta_{ab} \); one achieves this by the Wick rotation of the time coordinate into the purely imaginary domain \( t \rightarrow it \). The self-duality conditions for the de Sitter curvature (6.10) read

\[
R = *R,
\]

and this is written explicitly as the system

\[
\left( R_{ab}^b - \frac{1}{l_0^2} \pi_{a}^b \right) = * \left( R_{ab}^a - \frac{1}{l_0^2} \pi_{a}^b \right), \tag{6.29}
\]

\[
\hat{Q}^a = * \hat{Q}^a. \tag{6.30}
\]

On the self-dual configurations (6.29), (6.30), the vacuum field equations (6.12) and (6.13) are satisfied due to generalized Ricci (1.15) and Bianchi (1.17) identities which can be conveniently recast into

\[
2\mathcal{D}_{[\lambda} Q_{a] \mu \nu} = - R_{a[b} h_{c]}^b, \quad \mathcal{D}_{[\lambda} R_{a b \mu \nu]} = 0.
\]

The solutions of (6.29), (6.30) are called gravitational instantons. The Euclidean action

\[
S = \frac{1}{8\lambda} \int \left\{ \left( R_{ab}^b - \frac{1}{l_0^2} \pi_{a}^b \right) \wedge * \left( R_{ab}^a - \frac{1}{l_0^2} \pi_{a}^b \right) + \frac{2}{l_0^2} \hat{Q}^a \wedge * \hat{Q}_a \right\}, \tag{6.31}
\]
reaches a local extremum (minimum) on the instanton configurations, when it becomes proportional to the topological invariant

\[-\int \text{Tr}(\mathcal{R} \wedge \mathcal{R}) = \int \left\{ R^{ab} \wedge R_{ab} + \frac{2}{l_0^2} \left( \hat{Q}^a \wedge \hat{Q}_a - R^{ab} \wedge \pi_{ab} \right) \right\}.\]

The first term (with an appropriate numeric coefficient) is equal to the Pontryagin index, whereas the second term describes the Nieh-Yan topological invariant.

In the theory under consideration, the Euclidean action is well defined (unlike the Euclidean action in GR and ECT), leading to the formally convergent path integral in the quantum theory.

The absolute minimum of (6.31) is realized on the Euclidean de Sitter space

\[\mathcal{R} = 0 \quad \left\{ Q^a_{\mu \nu} = 0, \quad R^{ab}_{\mu \nu} = \frac{2}{l_0^2} h^{[a}_{\mu} h_{b \nu]} \right\}, \quad (6.32)\]

on which the action vanishes, \( S = 0. \)

It is worthwhile to note that using the ansatz \( R^a_{\beta \mu \nu} = \frac{2}{l_0^2} \delta^a_{[\mu} g_{\nu] \beta} \) when solving the complete system of the vacuum field equations (6.18)-(6.20) automatically yields the zero torsion. This result agrees with the original assumptions: in the \( S_{10} \)-theory, the simplest solution is the de Sitter space of the constant curvature, but not the Minkowski flat space.

One can straightforwardly describe the torsionless instantons. When \( Q = 0, \) the contraction of the equations (6.29) gives (due to the Ricci identity) the Einstein equations with the \( \lambda \)-term. Thus, in the theory under consideration, the Einstein spaces constitute the subset of the instanton solutions, and it is quite natural to interpret the solutions of Einstein’s equations as gravitational instantons in this approach.

Let us consider in detail the important special case: the spherical \( SO(4) \)-symmetric instanton solutions of (6.29) and (6.30). The general spherically symmetric (in the four-dimensional \( SO(4) \) sense) ansatz for the local Lorentz connection and the tetrad reads as follows:

\[
\Gamma^a_{b \mu} = A(x^a \delta_{b \mu} - x_{b} \delta^a_{\mu}) + B \varepsilon^a_{b \mu \nu} x^\nu, \quad (6.33)
\]

\[
h^a_{\mu} = f \delta^a_{\mu} + g x^a x_{\mu}, \quad (6.34)
\]

where \( A = A(\rho), \quad B = B(\rho), \quad f = f(\rho), \quad g = g(\rho), \) are the scalar functions of \( \rho = x^a x_a = \delta_{ab} x^a x^b. \)

Substituting (6.33), (6.34) to (6.29), we obtain the system

\[
2f A' f + 2Af = \left( 2B - 2AB \rho + \frac{f^2}{l_0^2} \right) (f + \rho g) = 0, \quad (6.35)
\]

\[
2(A \pm B)' f - 2(A \pm B) g + (A \pm B)^2 (f + \rho g) = 0. \quad (6.36)
\]

Here the dash denotes derivatives with respect to \( \rho, \) and the upper signs correspond to the self-duality \( (R - l_0^{-2} \pi) = * (R - l_0^{-2} \pi), \) whereas the lower signs
refer to the anti-self-duality \((R - l_0^{-2}) = - *(R - l_0^{-2})\). In a similar way, for the torsion we derive from (6.30) one more equation

\[2f' - g + (A \pm 2B)(f + \rho g) = 0. \quad (6.37)\]

The spherical \(SO(4)\)-symmetric ansatz (6.33) and (6.34) excludes the traceless part of the torsion. However, the trace and pseudotrace are non-trivial:

\[Q_\mu = - \frac{3}{2f} x_\mu \{2f' - g + A(f + \rho g)\}, \quad \bar{Q}_\mu = - \frac{3}{2f} x_\mu B(f + \rho g),\]

and the self-duality equation (6.37) establishes their relation (it is obvious that for such configuration the trace \(Q^\mu\) is proportional to the pseudotrace \(\bar{Q}^\mu\)). This implies the first result: any self-dual \(SO(4)\)-symmetric solution (6.33), (6.34) describes the Riemannian spacetime (there is no torsion), if either \(B = 0\) or \(f + \rho g = 0\).

All Riemannian \(SO(4)\)-instantons may be described explicitly.

When \(f + \rho g = 0\), we obtain a degenerate geometry. Indeed, then \(g = - f/\rho\) and thus the tetrad \(h^a_\mu = f(\delta^a_\mu - x^a x_\mu/\rho)\) is a projector with the vanishing determinant. The same applies to the corresponding Euclidean spacetime metric \(g_{\mu\nu} = f^2(\delta_{\mu\nu} - x_\mu x_\nu/\rho)\) which represents an unphysical geometry.

When \(B = 0\), there exists an infinite family of the \(SO(4)\)-symmetric solutions of the system (6.35)-(6.37):

\[(A\rho - 1)^2 = 1 \mp \frac{f^2 \rho}{l_0^2}, \quad g(1 - A\rho) = 2f' + Af.\]

We can express \(A\) and \(g\) in terms of an arbitrary function \(f\), provided \(A\rho \neq 1\). For the self-dual case, we have an exceptional solution \(A = 1/\rho, f = l_0/\sqrt{\rho}\) with an arbitrary \(g\). This is a singular solution, but the infinite family encompasses the regular configurations, as well. The famous de Sitter instanton is recovered from the family above by fixing \(g = 0\):

\[A = \frac{2}{\rho \pm \rho_0}, \quad f = \frac{2l_0 \sqrt{\rho_0}}{\rho \pm \rho_0}, \quad g = 0, \quad B = 0.\]

As usual, the upper (lower) sign refers to the self-duality (anti-self-duality). For the positive real integration constant \(\rho_0\), the self-dual solution is everywhere regular, but the anti-self-dual solution diverges at the radius \(\rho = \rho_0\).

The search for the non-Riemannian \(Q^a_{\mu\nu} \neq 0\) \(SO(4)\)-instantons represents a nontrivial task. There is no their complete description yet, and specific solutions are difficult to interpret. A preliminary analysis can be found in [194].

Possible existence of the instanton configurations is consistent with the idea of Hanson and Regge [145] who noticed the analogy of the gravity with \(Q \neq 0\) and the superconductivity theory. The vanishing torsion is then dynamically realized as a phase of the Meissner type effect, and the local regions with nontrivial torsion \(Q \neq 0\) are the analogs of Abrikosov's vortices.
Torsion dynamics and universal spin-spin interaction

Let us consider the physically most important particular type of matter – the fermion fields (describing quarks, leptons, hadrons) as a source of the gauge gravitational field. Then the tensor of spin is totally antisymmetric and has the following form:

\[ S_{\lambda\mu\nu} = \varepsilon_{\lambda\mu\nu\kappa} \tilde{S}^\kappa, \quad \tilde{S}_\mu = \frac{i\hbar}{4} \nabla_{\mu} \gamma_5 \Psi. \]

Substitution of the ansatz \( R^a_{b c} = \frac{1}{l^2} \pi^a_{b c} \) into the field equations of the de Sitter model (6.18)-(6.20) yields the Riemannian de Sitter space in the absence of matter, but for the nontrivial fermion sources this ansatz gives rise to the algebraic relation between the torsion and the spin \( \tilde{T}_\mu = \kappa c \tilde{S}_\mu \) via (6.18), with the “translational” equation (6.19) reducing to the Einstein equation with the \( \lambda \)-term and the symmetrical part of the canonical energy-momentum tensor on the right-hand side.

Let us qualitatively investigate the dynamics of the torsion field in the de Sitter gauge gravity theory. For this purpose, we consider a simplified model: we assume the metric to be flat \( g_{\mu\nu} = \eta_{\mu\nu}, \ h^a_{\mu} = \delta^a_{\mu} \), and the torsion is represented only by the pseudotrace vector \( \tilde{T}_\mu \), which is natural if the sources are fermions.

Then the action describing the torsion dynamics takes the following form:

\[ S_T = \int d^4 x \sqrt{g} \left\{ \frac{1}{\lambda} \left[ \partial_\mu \tilde{\nabla} \partial^\mu \tilde{T}^\nu + \frac{1}{2} (\partial_\mu \tilde{T}^\nu)^2 + \mu^2 \tilde{T}_\mu \tilde{T}^\mu - \frac{3}{2} (\tilde{T}_\mu \tilde{T}^\mu)^2 \right] + \tilde{S}_\mu \tilde{T}^\mu \right\}, \]

where \( \mu^2 = 9\lambda/\kappa c = 9/l_0^2 \).

We can use the method of Gupta [195] to calculate the static non-relativistic interaction potential between the two fermions with spins \( \sigma_1 \) and \( \sigma_2 \) exchanging by the \( \tilde{T} \)-quanta. In the linear approximation we obtain

\[ V_T (r) = -\frac{\lambda c \hbar^2}{8\pi} \left\{ (\sigma_1 \cdot \sigma_2) \frac{e^{-\mu r}}{r} - \frac{1}{\mu^2} (\sigma_1 \cdot \nabla)(\sigma_2 \cdot \nabla) \left( \frac{e^{-\mu r}}{r} - \frac{e^{-\sqrt{2} \mu r}}{r} \right) \right\}. \]

In the limit of the zero transferred momentum in \( \tilde{T} \)-propagators (\( \mu \neq 0 \)), one finds the contact potential like in the ECT. Therefore, the presence of the Yang-Mills (quadratic in the curvature) term in (6.11) modifies ECT in the same way as the theory of intermediate bosons modifies the Fermi theory of weak interactions.

Qualitatively the same result for the interaction potential arises also for the most general quadratic gauge gravity model (6.23), where one needs to replace the parameters \( \lambda, \mu^2 \) and 2/3 with the appropriate algebraic expressions constructed from the coupling constants \( a_1, \ldots, a_3 \) and \( b_1, \ldots, b_6 \), see the details in [196, 197, 198].

The spin-spin potential is also found [146] in GR for the fermions interacting via the graviton exchange, where \( V_G \sim \kappa c^2 \hbar^2 (\sigma_1 \cdot \sigma_2)/r^3 \) (analog of the Breit potential). In contrast to \( V_G \), the dynamical gauge (\( T \)-S)-interaction predicts the \( 1/r \) spin-spin interaction (similar to Coulomb’s and Newton’s potentials).
6.4. Model description of microscopic gravitational interactions

We now turn to the study of the microscopic limit of the gauge gravity theory. Addressing this problem one can relax the requirements formulated in Sec. 6.1. for the construction of the dynamics, in particular it makes no sense to talk about the correspondence with GR at small distances. Moreover, in accordance with the gauge principles in the framework of the geometrodynamical approach (following Einstein’s idea that the spacetime geometry is not fixed ad hoc, but is determined by the interaction and the motion of matter), one can assume that at small scales we may need to consider new microscopic physical properties of matter, which did not manifest themselves at the macroscopic level. Understanding the geometrization program in a general sense, the new physics should lead to the new geometry. Accordingly, the minimal model based on the Poincaré group (or its extension to the de Sitter group) may turn out to be too narrow in the area of the high energies (very small distances).

Most probably, the $\mathcal{P}_{10}$-theory provides a good description for the gravitational interactions of leptons, since their spacetime symmetries are exhausted by the Poincaré group. However, there are additional symmetries in the physics of strongly interacting particles [147] besides the $\mathcal{P}_{10}$. In particular, the systematics of hadrons within the Regge approach [148] suggests the description of all particles lying on a Regge trajectory as excitations of one physical object. Such states are classified by the unitary representations of $SL(3,R)$ group, or in relativistic theory by $SL(4,R)$. In addition, in the high energy strong interactions one observes another (asymptotic) regularity – the Bjorken scaling [149] as a manifestation of the dilational symmetry [140]. Together, the dilations and the $SL(4,R)$ constitute the general linear group and, adding the translations, we end up with the general affine group $GA(4,R) = GL(4,R) \supset T_4$.

Therefore, one can suggest that the gravitational interactions of the hadrons should be described not by the $\mathcal{P}_{10}$-theory, but rather by a more general theory based on $GA(4,R)$. The corresponding gauge theory is called a metric-affine gravity; for its comprehensive overview see [199].

In Secs. 5.3. and 5.4., we constructed the kinematics of the gauge theory for the general affine group. The gauge potential is identified with the generalized affine connection in the principal bundle of affine frames $A(M)$, which introduces the most general geometrical structure on the spacetime manifold $M_4$: the independent linear connection $\Gamma^\alpha{}_{\beta\mu}$ and the metric $g_{\mu\nu}$ (the tetrad fields $h^a_{\mu}$). The sources of such a microscopic gravitational field are the corresponding Noether $GA(4,R)$-currents of matter: the energy-momentum and the “hypermomentum” [147, 199]

\[
J^\mu_{ab} = \frac{\partial L_m}{\partial \partial_\mu \Phi} I_{ab} \Phi. \tag{6.39}
\]
Here $\mathcal{L}_m$ is the Lagrangian of the matter fields $\Phi$, and $I_{ab}$ are the generators of $GL(4,R)$. The irreducible parts of the hypermomentum (6.39) are the spin $S_{\mu ab} = J^\mu_{[ab]}$, the dilation current $J^\mu = J^\mu_{ab}\eta^{ab}$, and the “proper hypermomentum” $\mathcal{J}^\mu_{ab} = J^\mu_{(ab)} - \frac{1}{4} J^\mu \eta_{ab}$, corresponding to shear deformations of the frame.

### Microscopic gravitation as a Yang-Mills field

Now we consider the microscopic gravitation model $(\Gamma^a_{\mu b}, h^a_\mu)$ assuming that the linear connection $\Gamma^a_{\mu b}$ is the most general one. The dynamics of the theory at small distances can be formally obtained by discarding the “translational” low-energy terms in the basic Lagrangian (6.11) or taking a formal limit $l_0 \to \infty$. Thus, we will describe the model microscopic theory of the gravitational interactions by the Lagrangian of the Yang-Mills type for the $GL(4,R)$-gauge field

$$L_{YM} = \frac{1}{8\lambda} R^a_{b\mu\nu}(\Gamma) R^{b\mu\nu}_a(\Gamma). \quad (6.40)$$

The invariance group of the theory (6.40) is very wide. It includes the conformal Weyl group $g_{\mu\nu} \to g'_{\mu\nu} = e^{2\sigma(x)} g_{\mu\nu}$ (see Sec. 4.2); the local linear $GL(4,R)$-group $h^a_\mu \to h'_a\mu = L_{ab} h^b_\mu$, $\Gamma^a_{\mu b} \to \Gamma^a_{\mu b}' = L^a_{c\mu} \Gamma^c_{ab} L^{-1}_{c\mu}$, with $L \in GL(4,R)$ and the group of general coordinate transformations: for example, $h^a_\mu \to h'^a_\mu = \frac{\partial x^a}{\partial x'^\mu} h^a_\mu$.

The field equations are derived by the variation of (6.40) with respect to the independent $\Gamma^a_{\mu b}$ and $g_{\mu\nu}$. In vacuum, they read

$$\nabla^\mu R^a_{b\mu\nu}(\Gamma) = \frac{1}{\sqrt{g}} \partial^\mu \left\{ \sqrt{g} R^a_{b\alpha\beta}(\Gamma) g^{\alpha \mu} g^{\beta \nu} \right\}$$

$$+ g^{\alpha \mu} g^{\beta \nu} \left\{ \Gamma^a_{c\mu} R^c_{b\alpha\beta}(\Gamma) - \Gamma^c_{b\mu} R^c_{a\alpha\beta}(\Gamma) \right\} = 0, \quad (6.41)$$

$$\tau_{\mu\nu} = - \frac{c}{2\lambda} \left\{ R^a_{b\mu\lambda}(\Gamma) R^{b\nu\lambda}_a(\Gamma) - \frac{1}{4} g_{\mu\nu} R^a_{b\alpha\beta}(\Gamma) R^b_{a\alpha\beta}(\Gamma) \right\} = 0. \quad (6.42)$$

Compare the last equation with (6.14).

Technically, the equation (6.42) means that the energy-momentum tensor of the Yang-Mills field vanishes, which therefore is often called a ghost field. It is convenient to introduce the conformally invariant tensor density $\pi^{\mu\nu\alpha\beta} = \sqrt{g} (g^{\mu\alpha} g^{\nu\beta} - g^{\mu\beta} g^{\nu\alpha})$. The metric enters the theory only in terms of $\pi^{\mu\nu\alpha\beta}$, in particular, the action reads $S_{YM} = \frac{1}{16\lambda} \int d^4x \pi^{\mu\nu\alpha\beta} R^a_{b\mu\nu}(\Gamma) R^b_{a\alpha\beta}(\Gamma)$ and the field equation (6.41) can be rewritten as

$$\partial^\mu \left\{ R^a_{b\alpha\beta}(\Gamma) \pi^{\mu\nu\alpha\beta} \right\} + \left\{ \Gamma^a_{c\mu} R^c_{b\alpha\beta}(\Gamma) - \Gamma^c_{b\mu} R^c_{a\alpha\beta}(\Gamma) \right\} \pi^{\mu\nu\alpha\beta} = 0. \quad (6.43)$$

The importance of the Yang-Mills fields motivates the interest in the exact solutions of the classical gauge field equations. It is known [150, 151] that any spherically symmetric solution of the Yang-Mills equations for the internal groups $[SU(2)$ etc.] in the Minkowski space has the “Coulomb” type behaviour.
Let us study the spherically symmetric solutions of (6.43) in the Riemannian space-time with a given background metric $g_{\mu\nu}(x^\lambda)$. In the local spherical coordinates $(x^0 = t, x^1 = r, x^2 = \theta, x^3 = \varphi)$ for an arbitrary spherically symmetric metric $ds^2 = -\xi(r,t)dt^2 + \eta(r,t)dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\varphi^2$, the most important component of the density $\pi^{\mu\nu\alpha\beta}$ reads as follows:

$$\pi^{0101}(t, r, \theta, \varphi) = f(r, t)\chi(\theta, \varphi), \quad (6.44)$$

where $\xi, \eta, f, \chi$ are the functions of the specified arguments. Under the condition of the spherical symmetry, we assume that $\Gamma_{\alpha\beta\gamma}^0(r, t)$ depends on the time and the radial coordinates, whereas $\Gamma_{\alpha\beta\gamma}^1 = \Gamma_{\alpha\beta\gamma}^2 = \Gamma_{\alpha\beta\gamma}^3 = 0$. Then one can check that the solution of (6.43) reads

$$\Gamma_{\alpha\beta\gamma}^0(r, t) = A_{\alpha\beta} \int_{r_0}^r \frac{dr'}{f(r', t)}, \quad R_{\alpha\beta\gamma}^0 = \frac{\partial \Gamma_{\alpha\beta\gamma}^0}{\partial r}, \quad (6.45)$$

where the constant matrix $A \in GL(4, R)$. In [152], it is proved that this is a general (up to a gauge transformation) spherically symmetric solution of the equations (6.43).

In addition, when the matrix $A$ is such that $A_{\alpha\beta}A_{\beta\alpha} = 0$, eq. (6.45) describes the exact solutions of the complete system of the microscopic gravitation (6.41) and (6.42), since then $\tau_{\mu\nu} = 0$ identically. Such matrices do exist in $GL(4, R)$.

To illustrate (6.45), we consider several special background metrics.

1. Let $f(r, t) = -r^2$. This is true for the Schwarzschild, the Reissner-Nordström, the de Sitter, the Kottler metrics (the spherical vacuum solution of Einstein’s equations with the cosmological term), the conformally flat metrics (in particular, for the Friedman metric) and for some others. In this case, $R_{\alpha\beta\gamma}^0 = -A_{\alpha\beta}/r^2$, and $\Gamma_{\alpha\beta\gamma}^0 = A_{\alpha\beta}(1/r - /r_0)$. The solution thus has the “Coulomb” form, generalizing the result of [150] to the aforementioned Riemannian spaces.

2. Let us consider the cosmological metrics

$$ds^2 = -dt^2 + a^2(t) \left\{ \frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\varphi^2) \right\}, \quad k = \text{const}.$$ 

Then $f(r, t) = -a(t)r^2\sqrt{1-kr^2}$ and

$$\Gamma_{\alpha\beta\gamma}^0(r, t) = \frac{A_{\alpha\beta}}{a(t)} \left( \frac{\sqrt{1-kr^2}}{r} - \text{const} \right), \quad R_{\alpha\beta\gamma}^0 = -\frac{A_{\alpha\beta}}{a(t)r^2\sqrt{1-kr^2}}.$$

For particular cases, we have:

a) $k = 1/r_0^2$ is a closed universe, with $a(t) = 1$ this is the static Einstein’s world. The gravitational field strength $R_{\alpha\beta\gamma}^0(r, t)$ has singularities at $r = 0$ and at $r = r_0$, at the boundary or the world. The region $r > r_0$ is unphysical. Thus, in the Einstein world, the spherically symmetric solutions of the Yang-Mills equations have a non-Coulomb short-range behaviour.
b) $k = -1/r_0^2$ is an open universe, $a(t) = 1$. The solution is asymptotically Coulomb one.

c) $k = 0$, $a(t) = e^{Ht}$, $(H = \text{const})$ is the inflationary cosmological de Sitter metric in the Lemaitre-Robertson form [120]. In this case

$$R_{b0}(r,t) = -\frac{A_{b}^{a}}{r^2}e^{-Ht}, \quad \Gamma_{b0}(r,t) = A_{b}^{a}(1/r - 1/r_0)e^{-Ht}.$$  

A local observer interprets this solution as a Coulomb type configuration with a decreasing gauge charge (compare this with Dirac’s hypothesis [68] about the change of constants in time).

Introducing the interaction of the microscopic gravity with matter, the field equations are modified by the relevant sources: the hypermomentum (6.39) is added in the right-hand side of (6.41) and the energy-momentum tensor $t_{\mu\nu}$ in (6.42). The conformal symmetry of the model (6.40) give rise to the restriction $t_{\mu\mu} = 0$ that allows only the massless matter fields as the sources of the $(\Gamma, g)$-gravity. Examples of such fields are the massless spinor fields and the Yang-Mills fields in the theories of “grand unification” (before a spontaneous symmetry breaking), which underlines the essentially microscopic nature of model under consideration.

Below, we investigate the question how this microscopic picture is related to the macroscopic theory of gravity.

*Spontaneous breaking of conformal symmetry and macroscopic limit*

The structure of the microscopic gravity theory does not formally allow for the interaction with the macroscopic, massive matter. Therefore, in order to preserve the conformal invariance of a massive field action, we introduce an auxiliary scalar field $\varphi(x^{\lambda})$ which transforms under the action of the Weyl group as

$$\varphi \rightarrow \varphi' = e^{-\sigma(x)} \varphi(x).$$

If we now define the masses of a vector $A_{\mu}$ and a spinor $\Psi$ fields as $m = \varphi \bar{m}$, with $\bar{m} = \text{const}$, one can add the corresponding mass term to the complete matter Lagrangian.

It is important to determine the Lagrangian of the dynamical field $\varphi$ to set the transition to the macroscopic theory. There are two possibilities, which are based on the different interpretation of the Weyl symmetry that we analyzed in Sec. 4.2.. Let us consider both cases.

a) *Canonical Weyl approach and correspondence with GR*

The usual interpretation of Weyl’s transformations $g_{\mu\nu} \rightarrow e^{2\sigma}g_{\mu\nu}$ is based on the expansion of lengths, and in particular, on the scaling of the interval $ds \rightarrow ds' = e^{\sigma}ds$. If we understand the conformal symmetry of the microscopic gravity theory in this sense, the most general Weyl-invariant action for the
scalar field $\varphi$ has the form:

$$S = \int d^4x \sqrt{g} \left\{ \frac{1}{2} \partial_{\mu} \varphi \partial^{\mu} \varphi + \frac{R}{12} \varphi^2 - \Lambda_0 \varphi^4 \right\}, \quad \Lambda_0 = \text{const.}$$

We assume that the scalar field has the canonical dimension of an inverse length, whereas $\Lambda_0$ is dimensionless. Fixing the Weyl freedom by the gauge condition $\varphi = \varphi_0 = \text{const}$ [73, 149] (which corresponds to spontaneous breaking of conformal symmetry) one finds the complete action of the theory

$$S = \int d^4x \sqrt{g} \left\{ \mathcal{L}_M(\Gamma) + \frac{1}{c} \mathcal{L}_m(\Psi, A) + \frac{1}{2\kappa c}(R - 2\Lambda) - \frac{1}{2}\varphi_0^2 m_\lambda A_\mu A^\mu + \varphi_0^2 m_\Psi \overline{\Psi} \Psi \right\}. \quad (6.46)$$

Here $\mathcal{L}_m(\Psi, A)$ is the free Lagrangian of the massless vector $A_\mu$ and spinor $\Psi$ fields, and $\kappa c = 6/(\hbar \varphi_0^2)$, $\Lambda = 6\Lambda_0 \varphi_0^2$.

On the macroscopic scales $\gg 1/\varphi_0$, we have $\frac{1}{2\kappa} R \gg \mathcal{L}_M(\Gamma)$, i.e., the Hilbert-Einstein term is dominating. Consequently, the macroscopic limit in this case is described by GR with the cosmological term.

This procedure, that introduces the masses and the Lagrangian $R$, is the classical analog of the quantum mechanism of the dynamical breaking of the conformal symmetry, and the field $\varphi$ plays the role of the Goldstone boson [153].

Assuming that the coupling self-interaction constant of the scalar field is small ($\Lambda_0 \to 0$), one can ignore the $\Lambda$-term in (6.46), and hence the gravitational Lagrangian has the form $\mathcal{L}_0 = \frac{1}{2\kappa} R + \frac{1}{\kappa} R^a_{\mu\nu}(\Gamma) R^b_{\mu\nu}(\Gamma)$, described as the case (c) in Sec. 6.1.. The corresponding field equations read:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \kappa (t_{\mu\nu} + \tau_{\mu\nu}), \quad (6.47)$$

$$\nabla_{\nu} R^a_{b\mu\nu}(\Gamma) = 2\lambda J^a_{b\mu}, \quad (6.48)$$

The exact spherically symmetric solution of the vacuum equations (6.47), (6.48) (for the similar solutions for an arbitrary gauge group see in [99, 100, 152]) is straightforwardly derived:

$$ds^2 = e^\nu dt^2 + e^{-\nu} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2, \quad e^\nu = 1 - \frac{2Gm}{c^2 r} + \frac{\kappa c q^2}{4\Lambda r^2},$$

for the metric, and for the gauge field $\Gamma^a_{b\mu}$:

$$R^a_{b01} = \frac{u^a_b}{r^2}, \quad R^a_{b23} = v^a_b \sin \theta,$$

where $q^2 = -u^a_b v^b_a - v^a_b u^b_a$ and the dimensionless constant matrices $u^a_b$, $v^a_b \in \text{GL}(4, \mathbb{R})$. This solution is of the Reissner-Nordström type. The essential difference is that the value of the “effective charge” $q^2$ may be greater, equal or
DYNAMICS OF GAUGE GRAVITY THEORIES

smaller than zero, in view of the non-compactness of the group $GL(4, R)$. Formally, $q^2$ is composed of the three irreducible parts, since $u$ and $v$ are connected with the spin, the dilation current and the proper hypermomentum of the central body: $u_{ab} = \frac{1}{4} \eta_{ab} u + u_{[ab]} + \pi_a$, $\pi^a = 0$. The case $q^2 = 0$ belongs to the class of the abovementioned ghost solutions of the Yang-Mills equations.

The non-compactness of the symmetry group may have some specific consequences in the gauge gravity theory. For example, the indefiniteness of the sign of $q^2$ may prevent the formation of a horizon, as well as singularities may be avoided in cosmology. This is related to the fact that the sign of the energy $\tau_{00}$ of the field $\Gamma$ is indefinite, which leads to the violation of the energy conditions of the Penrose-Hawking theorems. Indeed, denoting $E^{ab}_{i} = R^{ab}_{ij}$ and $B^{ab}_{ij} = R^{ab}_{ij}$ (Latin indices from the middle of the alphabet $i, j, k, \cdots = 1, 2, 3$ label 3-space components), we can decompose the curvature into the four pieces

$$
\tilde{E}_i^k = E_{ik}, \quad \tilde{B}^{kj} = B^{kj}, \quad \tilde{E}_i^k = B_{ik}^0, \quad \tilde{B}^{kl}_{ij} = B^{kl}_{ij}.
$$

Then using the self-evident symbolic notation we find

$$
\tau_{00} = \frac{c}{2\lambda} \left( \tilde{E}^2 + \tilde{B}^2 - \tilde{E}^2 - \tilde{E}^2 \right),
$$

that clearly demonstrates that the energy density is not positive definite.

This can be observed as an “antigravity” effect, or a gravitational repulsion, that would be manifest as an effective reduction of mass. One can qualitatively illustrate this as follows. Suppose the matter source in (6.47), (6.48) is a massive cylinder made of a ferromagnetic material that can be magnetized along its axis $z$. The dilation current and the proper hypermomentum are both equal zero and the model thus reduces to an effective $T^4 \times L^6$-theory. Under the assumption of an approximately flat metric $g_{\mu \nu} \approx \eta_{\mu \nu}$, the polarized state of a homogeneous sample can be modeled by the spinor field $\Psi = \sqrt{n} e^{imt} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, so that the spin pseudovector is $\hat{S}_i^\mu = \{0, 0, 0, -\frac{\hbar}{2} n\}$, where $n$ is the particle density of the polarized spins. In the linear approximation, neglecting the boundary effects, we then find the internal solution of the equation (6.48) for the components of the curvature: $R^0_{ab} = \frac{\lambda}{2} \lambda x^i S^0_{ab}$, $R_{ab}^{ik} = \lambda \left( x^i S^k_{ab} - x^k S^a_{ab} \right)$. This yields $\tau_{00} = -\lambda c^2 n^2 (5\pi^2 + 9\pi^2) / 36$, and we obtain the correction to the energy caused by the $\Gamma$-field by evaluating the integral over the cylinder source:

$$
\Delta E = \int d^3 x \tau_{00} = -\frac{\lambda c n^2}{36} V \left( \frac{5 \rho_o^2}{2} + \frac{14 \ell_o^3}{3} \right) < 0,
$$

where sample’s geometry is specified by the radius $\rho_o$, the length $\ell_o$, and the volume $V = \pi \rho_o^2 \ell_o$. To make a numeric estimate, let us take $\rho_o \sim 10^{-2}$ m, $\ell_o \sim 10^{-1}$ m, and $n = 10^{29}$ m$^{-3}$. Then we conclude, that upon an instantaneous magnetization of the sample, its mass should effectively reduce by $\Delta M = \Delta E / c^2 \approx -\lambda c \times 10^{-8}$ kg, or the relative reduction should be
\( \Delta M/M \approx -\lambda h \times 10^{-7} \). This is a tiny effect for the value \( \lambda h = 10^{-120} \) of the coupling constant estimate obtained in Sec. 6.3. The current qualitative analysis shows that the gauge gravity coupling constant can be significantly larger, and even \( \lambda h \sim 1 \) is a viable estimate. The actual value of \( \lambda \) should be fixed by future experiments.

b) Internal conformal symmetry and model \( GA(4, R) \)-theory

In another approach, which was developed earlier in Sec. 4.2., the Weyl transformations are interpreted as the tetrad scaling \( h^\mu_a \rightarrow c^\sigma h^\mu_a \). In this case, the most general conformally invariant action for the auxiliary scalar field \( \varphi \) has the form

\[
\tilde{S}_\varphi = \int d^4x \sqrt{\mathcal{g}} \left[ R(\Gamma) \varphi^2 - \Lambda_0 \varphi^2 \right] \tag{6.49}
\]

The peculiar feature of this \((g_{\mu\nu}, \Gamma^{a\mu})\)-theory is the presence of an additional Abelian long-distance field in it – the Weyl vector \( K^\mu \). Using the irreducible decomposition of \( GL(4, R) \)-connection \( \Gamma^{a\mu} = \Gamma^{a\mu}_b dx^b \) 1-form

\[
\Gamma^{ab} = \tilde{\Gamma}^{ab} - \frac{1}{2} K^{ab}
\]

into the local Lorentz connection \( \tilde{\Gamma}^{ab} = \Gamma^{(ab)} \) and the nonmetricity\(^1\) 1-form \( K^{ab} := D\eta_{ab} = -2\tilde{\Gamma}_{(ab)}^\mu \), we find the relevant decomposition of \( GL(4, R) \)-curvature \( R^{a\mu}_b = \frac{1}{2} R^{a\mu\nu}_b (\Gamma) dx^\mu \wedge dx^\nu \) 2-form

\[
R^{a\mu}_b = \tilde{R}^{a\mu}_b + \frac{1}{4} K^{a\mu}_c \wedge K^{c\mu}_b - \frac{1}{2} \partial \tilde{K}^{a\mu}_b + \frac{1}{4} \delta^a_b \Omega \tag{6.50}
\]

into the local Lorentz curvature 2-form \( \tilde{R}^{a\mu}_b = d\tilde{\Gamma}^{a\mu}_b + \tilde{\Gamma}^{c\mu}_a \wedge \tilde{\Gamma}^{c\mu}_b \) and the homothetic curvature\(^2\) 2-form \( \Omega = -\frac{1}{2} dK = -\frac{1}{2} \partial \eta_{\mu\nu} \) is easy to show that the vector field \( K^\mu \) decouples from the other fields and interacts with matter and with the metric similarly to the electromagnetic field. Furthermore, the homothetic curvature plays the role of the “Maxwell tensor”. Indeed, substituting (6.50) into (6.49) we recast the Yang-Mills part of the Lagrangian

---

\(^1\) Note that \( K^{ab} = K^{\mu\nu}_a dx^\mu \) where \( K^{\mu\nu}_a = h^\alpha_a \phi^\beta \nabla_\mu g_{\alpha\beta} \), cf. the definition of the tensor of nonmetricity (1.7). Furthermore, \( 2\Gamma^{(ab)} = D\eta^{ab} = -\eta^{ac} \eta^{bd} D\eta_{cd} \).

\(^2\) \( \Omega = \frac{1}{4} \Omega_{\mu\nu} dx^\mu \wedge dx^\nu \), see (1.2) and (1.10) for the definition of the homothetic curvature tensor.
into
\[ R^a_b \wedge \ast R^b_a = \frac{1}{4} \Omega \wedge \ast \Omega + \frac{1}{4} \hat{D} \mathcal{K}^a_b \wedge \ast \hat{D} \mathcal{K}^b_a \]
\[ + \left( R^a_b + \frac{1}{4} \mathcal{K}^a_c \wedge \mathcal{K}^c_b \right) \wedge \ast \left( R^b_a + \frac{1}{4} \mathcal{K}^b_d \wedge \mathcal{K}^d_a \right), \]

while the curvature scalar \( R(\Gamma) \), depends only on the first term from the right-hand side of (6.50), not containing \( K^\mu \). The matter source of the quasi-Maxwell Weyl field is identified with the dilation current \( J^\mu \) which is the trace part of the hypermomentum (6.39). The latter is conserved due to the invariance of the \( \mathcal{L}_1 \) Lagrangian (6.49) with respect to Einstein’s \( \lambda \)-transformations [70].

The remaining irreducible non-Riemannian parts of \( GL(4, R) \)-connection \( \Gamma \) have the short-range behaviour due to the effective mass arising for them from the Hilbert-Einstein term in (6.49). One can show that in the weak field approximation for the Lagrangian (6.49), the translational field \( h^\mu_a \) has the standard Einstein dynamics, thus providing the correspondence with GR.
Quantization of gravity

7.1. Methods of quantum gravity

Construction of the quantum theory of a physical system starts with the conversion of its degrees of freedom into operators on which the canonical commutation relations are then imposed. Obviously this should be preceded by the construction of the canonical formalism which was analyzed at the classical level in the previous chapters. The results of this analysis, caused by the special features of the canonical formalism in the curved spacetime and the degeneracy of the gravitational action, reveal a number of new features of the quantum theory.

A basic feature of gravity theory is that it is degenerate in virtue of its gauge invariance under local diffeomorphisms, so that we need a method of quantization of degenerate systems. As was shown in Chapter 2, the degenerate theory described by the canonical action (2.27) with the first-class constraints (2.28) after choosing the canonical gauges (2.32) reduces to the effective non-degenerate system with the action (2.40) in terms of independent physical degrees of freedom \( \Phi^* = (q^*, p^*) \) satisfying (2.39). The quantization of this theory means that \( q^* \) and \( p^* \) become operators in the Hilbert space of states and are subject to canonical commutation relations

\[
\left[ \hat{\Phi}^*, \hat{\Phi}^* \right] = i\hbar \{ \Phi^*, \Phi^* \}^*,
\]

(7.1)

where \( \{ \cdot, \cdot \}^* \) are the Poisson brackets calculated for the set of the physical phase variables \( \Phi^* \). In what follows we consider the Planck constant \( \hbar \) to be one in universal units, but later will reinstate it explicitly when considering the semiclassical expansion in powers of \( \hbar \).
The variables of the original total phase space $\Phi = (q, p)$ also become operators as the operator functions of $\hat{\Phi}^*$ (2.38). However, their commutation relations differ essentially from the “naive” canonical commutators defined relative to the Poisson brackets $\{\ ,\ \}$ in $(q, p)$-space

$$\left[\hat{\phi}, \hat{\phi}'\right] = i \{ \hat{\Phi}(\hat{\phi}^*), \hat{\Phi}'(\hat{\phi}^*) \}^* \neq i \{ \hat{\phi}, \hat{\phi}' \}$$

and essentially depend on a choice of a gauge in the complete set of constraints $\psi^a = (\chi^a, H_\mu)$.

Thus, within the framework of a certain set of gauges, the quantum system is described by the operators $\Phi^*$, and its state is determined in the Schrödinger picture by the vector of the space of states $|\Psi_{\text{phys}}(t)\rangle$, evolving according to the Schrödinger equation with the physical Hamiltonian

$$i \frac{\partial |\Psi_{\text{phys}}(t)\rangle}{\partial t} = \hat{H}_{\text{phys}}(\hat{\Phi}^*) |\Psi_{\text{phys}}(t)\rangle.$$  

(7.3)

When applied to the gravity, this approach is a direct realization of the quantization program in the ADM procedure. Solution of the constraints in any given gauge defines the chronologically ordered family of space-like hypersurfaces and the corresponding family of quantum states representing the quantum evolution.

However, this approach has serious disadvantages. Solution of the constraints is in general a complicated problem, the resulting physical Hamiltonian turns out to be non-local, and the whole method is not manifestly covariant. Moreover, a priori there is no guarantee that such quantization performed with different gauge conditions will give unitarily equivalent results. The first step to the understanding how this equivalence can be attained was done in [202] within the path integral formalism.

The time evolution operator for the Schrödinger equation (7.3) has a form of the path integral over the physical phase space [154, 155]:

$$U_{\text{phys}}(g_1, t_1 | g_0, t_0) = \int D[g^A, p_A] \exp \int_{t_0}^{t_1} dt \left( p_A \dot{g}^A - H_{\text{phys}}(g^A, p_A) \right),$$

(7.4)

$$D[g^A, p_A] = \prod_{t, A} dg^A(t) \prod_{t' B} dp_B(t'),$$

(7.5)

where the integration is done with respect to the Liouville measure\(^1\) and with the boundary conditions for the phase space coordinate $g^A(t)$ at the initial $t_0$ and final $t_1$ moments of time

$$g^A(t_0) = g_0^A, \quad g^A(t_1) = g_1^A.$$  

(7.6)

---

\(^1\)Here and in what follows we will disregard trivial purely numerical factors like $\prod_i (1/2\pi)$ in the path integral measure, in the Fourier representation of relevant functional delta functions, etc.
One should use the relation between the phase space integration measure on the physical subspace of $g^A$ and $p_A$ and the delta-function type measure on the original phase space of $q^i$ and $p_i$,

$$\prod_A dq^A dp_A = \prod_i dq^i dp_i \delta(\chi(q,p)) \delta(H(q,p)) J(q,p). \tag{7.7}$$

Here the delta functions of gauge conditions and constraints have a support on the subspace where they are vanishing,

$$\delta(\chi(q,p)) = \prod_\mu \delta(\chi_\mu(q,p)), \tag{7.8}$$
$$\delta(H(q,p)) = \prod_\mu \delta(H_\mu(q,p)) = \int \left( \prod_\mu dN_\mu \right) e^{iN_\mu H_\mu(q,p)}, \tag{7.9}$$

and the factor $J(q,p)$ is the determinant of the functional matrix of the canonical Faddeev-Popov operator $J_\mu^\nu(q,p)$ given by the Poisson brackets of the set of gauge conditions and constraints (cf. equation (2.33) or equation (3.7) for the case of a special choice of canonical coordinates $\Phi = (X^\mu, \Pi_\mu; g^A, p_A)$)

$$J(q,p) = \det J_\mu^\nu(q,p), \quad J_\mu^\nu(q,p) = \{ \chi^\mu, H_\nu \}. \tag{7.10}$$

Therefore, making according to (7.7) the change of integration variables in (7.3) from $(g^A, p_A)$ to $(q,p)$ and using the relation (2.39) we get

$$U_{\text{phys}} \sim \int D[q,p] DN \delta(\chi(q,p)) |J[q,p]| \exp i \int_{t_0}^{t_1} dt \left( p_i \dot{q}^i - H_0 - N_\mu H_\mu(q,p) \right), \tag{7.11}$$

where the linear combination of first class constraints in the exponentiated canonical action was achieved due to the integral representation of their delta function (7.9). Here the functional delta function $\delta[\chi(q,p)]$ and the functional determinant $J[q,p]$ obviously represent products over time moments of their local analogues,

$$\delta[\chi(q,p)] = \prod_t \delta(\chi(q(t),p(t))); \quad J[q,p] = \prod_t J(q(t),p(t)), \tag{7.12}$$

while the path integral measure $D[q,p]$ is of course similar to that of the physical phase space (7.5),

$$D[q,p] = \prod_{t,i} dq^i(t) \prod_{t',k} dp_k(t'), \quad DN = \prod_{t',\mu} dN_\mu(t'). \tag{7.13}$$

The canonical path integral representation of the unitary evolution operator (7.11) allows one to perform the gauge dependence analysis, because the variation of the gauge in this expression can be achieved by a canonical transformation of integration variables [202], and thus its gauge independence can
formally be proven. This gives a first hint on the unitary equivalence of quantization schemes in different sets of gauge conditions. Moreover, the contribution of the functional determinant can be represented as a functional integral over the anticommuting (Grassman) variables – the Faddeev-Popov ghost fields $C^\nu(t)$ and $\overline{C}_\mu(t)$ [203],

$$J[q,p] = \int D\overline{C} DC \exp \left\{ i \int dt C^\nu(t) J^\nu(q(t),p(t)) \right\}, \quad (7.14)$$

so that the total action becomes local in spacetime, and the path integral becomes manageable by a standard technique of renormalization of ultraviolet divergences. Integration over the canonical momenta converts it to the Lagrangian form which within the generalized set of special (relativistic) gauges becomes manifestly covariant and admits covariant renormalization.

All these advantages of the path integral representation remain, however, incomplete because the change of integration variables $(g^A, p_A) \rightarrow (q^i, p_i)$ is not consistent at the end points of the time interval $[t_0, t_1]$. This happens in virtue of the boundary conditions (7.6) and unequal footing of the coordinates and momenta at these points – while the coordinates are kept fixed, the momenta are being integrated over. For these reasons strict equality in (7.11) was replaced by the similarity sign. In scattering theory applications this difficulty is easily circumvented due to the simplicity of the linear regime for asymptotic states at $t_0 \rightarrow -\infty$ and $t_1 \rightarrow +\infty$ (S-matrix theory), but for finite times and nontrivial quantum states this issue represents a real problem. And this problem becomes critically important in the cosmological context because the initial state of the very early quantum Universe cannot be formulated in terms of asymptotic states of scattering theory.

This problem, in particular, leads to the question of how the boundary conditions on integration histories in the right hand side of (7.11) should be associated with the arguments of the kernel $U_{\text{phys}}$ on the left and side, how the composition law for the path integrals over $(q,p)$ in the time intervals $[t_0, t_1]$ and $[t_1, t_2]$ should look like, or what will be the analogue of the unitary propagation law for quantum states in the physical (ADM) sector:

$$\Psi_{\text{phys}}(g_1, t_1) = \int dg_0 U_{\text{phys}}(g_1, t_1 | g_0, t_0) \Psi_{\text{phys}}(g_0, t_0), \quad (7.15)$$

not to mention how the physical wave functions $\Psi_{\text{phys}}(g, t)$ should be associated with the quantum states $\Psi(q, t)$ in the coordinate representation of the operators $(\hat{q}, \hat{p})$. The answers to this set of questions will be partially given in the rest of this chapter, but before that we will briefly consider the ADM quantization method in terms of the physical sector of phase space variables $\Phi^\nu$. 
7.2. Arnowitt-Deser-Misner quantization method

In the ADM formalism (see Chapter 3), the physical degrees of freedom of arbitrary gravitating system are the pairs of the phase variables $\Phi^A = (g^A, p_A)$ (3.4), which in accordance with the previous section become operators satisfying the canonical commutation relations (7.1). The state of the system in the set of the gauges (3.11) is the quantum state on one of the space-like hypersurfaces defined in the four-dimensional manifold with the help of the same gauge conditions. This state is described by the vector $|\Psi_{\text{phys}}(t)\rangle$ which satisfies the Schrödinger equation

$$i \frac{\partial |\Psi_{\text{phys}}(t)\rangle}{\partial t} = H_{\text{phys}}(\hat{g}^A, \hat{p}_A) |\Psi_{\text{phys}}(t)\rangle,$$

(7.16)

where the Hamiltonian is given by the formula (3.14).

With the solution of this equation, one can find the quantum averages of all phase variables $\Phi = (\hat{g}_{ab}, \hat{p}^a; \hat{\phi}, \hat{p})$ of the system, as well as of the lapse and shift functions. Indeed, according to (3.5), (3.12) and (3.15), we have

$$\langle \hat{\Phi}(t) \rangle = \langle \Psi_{\text{phys}}(t) | \Phi[f, -P_\mu[f; \hat{g}^A, \hat{p}_A]; \hat{g}^A, \hat{p}_A] | \Psi_{\text{phys}}(t) \rangle,$$

$$\langle \hat{N}_\mu(x) \rangle(t) = \langle \Psi_{\text{phys}}(t) | \int d^3 y J^{-1(\mu)}(x, y) | \Psi_{\text{phys}}(t) \rangle f^\nu(t, y),$$

(7.17)

where in the last equation the operator nature of $J^{-1(\mu)}(x, y)$ implies that it is taken as a function of operators of physical variables and $X^\mu = f^\mu(t, x)$ and $\Pi_\mu = -P_\mu[f; \hat{g}^A, \hat{p}_A]$. The proper time of the observer, whose world line is described by the equation $x = \text{const}$, becomes the physical observable. Its average is given by the expression

$$\langle \tau \rangle = \int dt \langle N \sqrt{1 - N_a N^a/N^2} \rangle.$$

(7.18)

Similarly, the spatial intervals become the quantum observables. In accordance with the operator nature of the lapse and shift functions in (7.17), the position of a space-like hypersurface itself is an observable. This fact reflects special features of the quantization of the gravitational field, whose universal interaction determines the properties of space and time, and thereby affects the observer who cannot be shielded from this influence that has a quantum nature.

Example of ADM quantization: Friedmann universe with a scalar field

The dynamics of the closed homogeneous universe filled with the scalar matter will be briefly considered here as a demonstration of the ADM quantization formalism. The subject of quantization will not only be the matter field but
also the gravitational configuration of the Friedman model [164] described by
the metric (3.39).

The physical setup we will consider has basically a toy-model nature designed
more for the demonstration of the limitations of the ADM approach, rather
than for revealing some interesting new quantum physics undoubtedly inherent
to the early quantum cosmology. Mainly we will consider the situation when
the dynamics of the model is only slightly different from the evolution of the
classical Universe. The qualitative behaviour of closed Friedmann models in the
Einstein theory is the picture of the world that expands up to some maximal
cosmological radius from the initial singularity, and then contracts back to the
singular state. The expansion process is characterized by the suppression of
the originally present inhomogeneous gravitational modes. Conversely, when
contraction to the singularity, the Universe becomes essentially inhomogeneous
and anisotropic. If we restrict ourselves to the dynamical stage ranging from the
maximal radius to the domains which are sufficiently far from the singularity,
one can take the homogeneous and isotropic model as an approximation of the
closed Universe.

We can assume that the homogeneous gravitational field at the maximal ex-
ansion is generated by the large number of randomly moving particles that fill
uniformly the entire volume of the Universe being in a thermodynamic equi-
librium. To describe particles, it is necessary to introduce the inhomogeneous
matter field, whose separate oscillatory modes correspond to the particles of
different energies. Thus, the consistency condition of the model reduces not to
the requirement of the homogeneity of the matter field, as it is done in the
majority of the toy models [178] which claim to give a self-consistent dynamics
of the homogeneous matter and gravitational fields, but rather to the require-
ment of constancy in the three-dimensional space of such characteristics as the
energy density, the number of particles, etc.

The action of the massless scalar field \( \phi \) reads

\[
S_m = -\frac{1}{2} \int \! d^4 x \, g^{1/2} \, g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi. \tag{7.19}
\]

In view of the form of the metric (3.39), we derive the Lagrange function in terms
of the dimensionless scalar field \( \phi = l_0 \phi \) and the dimensionless cosmological
radius \( a \) (3.41)

\[
L_m = \frac{\varepsilon_0}{2} \int \! d^3 x \sqrt{\gamma} \left\{ \frac{\chi^2}{N} l_0^2 a^3 + a \phi \Delta \phi \right\}, \tag{7.20}
\]

where the covariant Laplacian \( \Delta \) is defined by the constant three-dimensional
metric \( \gamma_{ik} \). Decomposing the field \( \phi(x, t) \) with respect to the orthonormal system
of eigenfunctions \( Z_n(x) \) similarly to (4.56), \( \phi(x, t) = \sum_n \varphi_n(t) Z_n(x) \), we
obtain the Lagrange function (7.20) in terms of discrete set of degrees of freedom
of the scalar field

\[
L_m = \frac{\varepsilon_0}{2} \sum_n \left\{ \frac{\varphi_n'}{N} l_0^2 a^3 - N a \omega_n^2 \varphi_n' \right\}. \tag{7.21}
\]
7.2. Arnowitt-Deser-Misner quantization method

The superhamiltonian which corresponds (7.21), reads

\[ \int d^3x \mathcal{H}_\phi = \frac{\varepsilon_0}{2a} \sum_n \left\{ \frac{P_n^2}{a^2} + \omega_n^2 a^2 \phi_n' \right\}, \tag{7.22} \]

where \( P_n \) are momenta conjugated to \( \phi_n' \).

We take the Hamiltonian constraint (3.42), or (2.125) (rather, its integral analogue) in the form

\[ H_\perp = \frac{\varepsilon_0}{2} \left\{ -\frac{1}{2} (p^2 + a^2) + \frac{1}{2} \sum_n \left[ \frac{P_n^2}{a^2} + \omega_n^2 a^2 \phi_n'^2 \right] \right\} \tag{7.23} \]

and perform the canonical transformation of the general ADM procedure (3.4) in two stages. At first, we make the transformation

\[
\begin{align*}
(p, a; P_n, \phi_n') &\rightarrow (p', a'; p_n, \phi_n), \\
p &= p' + \frac{V}{a'}, & a = a', & V \equiv \sum_n p_n \phi_n, \\
P_n &= a' p_n, & \phi_n' = \phi_n / a',
\end{align*}
\tag{7.24}
\]

and then similarly to (3.48)-(3.49), we substitute \((a', p')\) by the new phase variables \(T\) and \(\Pi\):

\[ a' = \sqrt{-2\Pi} \cos T, \quad p' = \sqrt{-2\Pi} \sin T. \tag{7.25} \]

As a result, the constraint (7.23) takes the following form:

\[ H_\perp = \frac{\varepsilon_0}{\sqrt{-2\Pi} \cos T} \left\{ \Pi + H + \frac{V^2}{4\Pi \cos^2 T} \right\}, \tag{7.26} \]

where

\[ H = \frac{1}{2} \sum_n \left( p_n^2 + \omega_n^2 \phi_n^2 - 2 \phi_n p_n \tan T \right). \tag{7.27} \]

In accordance with the general ADM procedure, let us impose the gauge condition

\[ T = t \tag{7.28} \]

and solve the constraint equation (7.26) \( H_\perp = 0 \) with respect to \( \Pi \) approximately in the region of \( H \gg 1 \). Then with an accuracy \( O(H^{-1}) \), we obtain the action

\[ S = \int dt \left\{ \sum_n p_n \dot{\phi}_n - H(t) + \frac{V^2}{4H(t) \cos^2 t} + \ldots \right\} \tag{7.29} \]

in terms of the physical degrees of freedom \((\varphi_n, p_n)\).

From now on, we will consider such a range of the parameter \( t \) and such states for which \( H(t) \cos^2 t \gg 1 \), therefore, we can neglect the last term in (7.29).
Consequently, the physical Hamiltonian of the problem may be approximated by
\[
H(t) = \frac{1}{2} \sum_n \left( p_n^2 + \omega_n^2 \varphi_n^2 - 2 p_n \varphi_n \tan t \right).
\]
(7.30)

According to (3.54), (7.26), and (7.25) with an accuracy of \( O(H^{-1}) \), the expressions for the lapse function and scale factor read
\[
N = t_0 \left( \sqrt{2H(t)} + O(H^{-1}) \right) \cos t, \quad a = \left( \sqrt{2H(t)} + O(H^{-1}) \right) \cos t,
\]
(7.31)

since from the solution of the constraint equation we have \(-\Pi = H(t) + O(H^{-1})\).

We will quantize the system described by the Hamiltonian (7.30) in the Heisenberg representation, and as a first step we write the Heisenberg equation for \( \varphi_n \) that follows from (7.30):
\[
\ddot{\varphi}_n + \Omega_n^2 \varphi_n = 0,
\]
(7.33)

where \( \Omega_n^2 = \omega_n^2 + 1 \). As in Sec. 4.4., in order to introduce the particle interpretation of the system, we decompose the Heisenberg operator \( \varphi_n(t) \) with respect to the two complex-conjugated solutions of this equation \( u_n \) and \( u_n^* \):
\[
\varphi_n(t) = a_n u_n(t) + a_n^* u_n^*(t).
\]
(7.34)

The basis functions \( u_n \) and \( u_n^* \) are chosen so that they satisfy the normalization conditions
\[
u_n \dot{u}_n^* - u_n u_n^* = -i,
\]
therefore, for the creation and annihilation operators \( a_n^* \) and \( a_n \) the commutator reads \([a_n, a_m^*] = \delta_{nm}\).

Further analysis of the particle interpretation of the system can be done with the help of the standard procedure of Hamiltonian’s diagonalization [61] described in Sec. 4.4.. Diagonalization of the Hamiltonian at the moment of the maximal expansion of the Universe \( t = 0 \) takes place in terms of basis functions of the equation (7.33)
\[
u_n(t) = \frac{1}{\sqrt{2\Omega_n}} \left\{ \sqrt{\frac{\Omega_n}{\omega_n}} \cos \Omega_n t - i \sqrt{\frac{\omega_n}{\Omega_n}} \sin \Omega_n t \right\},
\]
(7.35)

which are nontrivial because the frequency \( \omega_n \) in the Hamiltonian and the frequency in the Heisenberg equation of motion \( \Omega_n \) do not coincide (for \( \Omega_n = \omega_n \)).
7.2. Arnowitt-Deser-Misner quantization method

these basis functions would reduce to the standard positive frequency ones \( \exp(-i\omega_n t)/\sqrt{2\omega_n} \). For \( n = 0 \) with \( \omega_0 = 0 \) the basis function is not well defined, but this mode does not have an oscillator nature and should be quantized in the coordinate (or momentum) representation which we will not consider in much detail\(^2\). The modes with \( n > 0 \) have a particle interpretation in terms of the creation and annihilation operators associated with the modes which would diagonalize the Hamiltonian at later moments of time and which do not coincide with (7.35).

In our toy model setup the initial state of the system is prescribed at the moment \( t = 0 \) of the maximal expansion of the Universe when the non-stationarity of the metric and of spacetime curvature is minimal. The proper cosmological time that the Universe spends in the region of the maximal expansion \( \tau = \int dtN = \int dtR(t) \) is large, therefore, despite the fact that we neglected the interaction of particles from the very beginning, this interaction is sufficient for the collision relaxation to bring the system of particles in the Universe into the state of the thermal equilibrium. Therefore, it is reasonable to define the initial state as the equilibrium or thermal density matrix \( \hat{\rho} \) at some temperature.

The average value of any physical observable \( \hat{O} \) may be defined by formula

\[
\langle \hat{O}(t) \rangle = \text{tp} (\hat{\rho} \hat{O}(t)) = \langle \overline{\hat{O}(t)} \rangle,
\]

where \( \langle \hat{O}(t) \rangle \) should be understood as the quantum averaging of the Heisenberg operator \( \hat{O}(t) \) over a pure multiparticle state from the space of occupation numbers, and the bar means the statistical averaging with respect to \( \hat{\rho} \). The details of such calculations for a thermal density matrix can be found in [71, 168].

Here we want to stress that in the approximation of large \( H(t) \) the calculation of the quantum average of the scale factor gives

\[
\langle \hat{a}(t) \rangle = \langle \sqrt{2H(t) + O(H^{-1}(t))} \rangle \cos t = \sqrt{2E(t)} \cos t \{1 + O(E^{-1}(t))\},
\]

where \( E(t) \) is the evolving in time quantum average of the Hamiltonian

\[
E(t) = \langle \overline{H(t)} \rangle.
\]

This quantum average is ultraviolet divergent, and because of the essential lack of covariance and nonlinearity of the physical observable \( \hat{a} \) it is hopeless to have a renormalization of the arising divergences by local covariant counterterms in the original action of the theory. Therefore, the obtained result makes sense only in the situation when the contribution of large occupation numbers in \( E(t) \) is a priori dominating the vacuum polarization contribution subject to UV renormalization. This setup is not applicable in the physically interesting

\(^2\)One can check that the dynamics of the quasi-Gaussian wave packet in the coordinate representation of \( \phi_0 \) with the growth of time will be qualitatively described by a quantum spreading of this packet.
case of the early quantum cosmology when the vacuum polarization effects and nontrivial scaling behavior play the dominant role. ADM quantization in view of its nonlocality, nonlinearity and lack of manifest covariance (which can hardly be restored even at the implicit level) is incapable of solving this important class of quantum problems. For this reason, we go over to another quantization approach to gravity theory.

7.3. Dirac-Wheeler-DeWitt quantum geometrodynamics

The Einstein-Hamilton-Jacobi theory provides an intermediate link between the classical theory of gravity and Dirac-Wheeler-DeWitt quantization, historically called a quantum geometrodynamics, and we begin the review of this approach with its analysis. For simplicity of presentation, we restrict ourselves with the case of the pure gravitational field and consider only the closed cosmology systems with $H_0 = 0$.

**Einstein-Hamilton-Jacobi theory as a semiclassical approximation to quantum geometrodynamics**

The Hamilton-Jacobi equations (2.44), (2.45) for spatially closed cosmological systems with $q^i = g_{ab}(\mathbf{x})$ and $\partial / \partial q^i = \delta / \delta g_{ab}(\mathbf{x})$, in accordance with (2.94) and (2.95), take the following form:

$$
\begin{align*}
G_{abcd} \frac{\delta S}{\delta g_{ab}} \frac{\delta S}{\delta g_{cd}} - g^{1/2} 3 \frac{\delta R}{\delta g_{ab}} &= 0, \\
2g_{ac} \left[ \frac{\delta S}{\delta g_{cb}} \right]_b &= 0, \\
\frac{\partial S}{\partial t} &= 0.
\end{align*}
$$

(7.39) \hspace{1cm} (7.40) \hspace{1cm} (7.41)

The last equation means that the Hamilton-Jacobi function $S$ of the gravitational field does not depend on time $t$. This is a manifestation of the “frozen” formalism for closed cosmological systems, as described in Chapter 3. Time independence of $S$ means that formally the time evolution can be incorporated in the system only with the aid of the gauge conditions (2.48) that by construction explicitly depend on time.

Thus, $S$ is a functional of the three-dimensional metric and of the $6 - 4 = 2$ functions $\alpha^4(\mathbf{x})$ which represent the independent integration constants in the complete integral of the system of equations (7.39), (7.40).

The equations (7.40) have a simple geometrical meaning since they reflect the invariance of the functional $S[g_{ab}(\mathbf{x}), \alpha^4(\mathbf{x})]$ with respect to the coordinate
7.3. Dirac-Wheeler-DeWitt quantum geometrodynamics

transformations of the metric

\[ \Delta^f g_{ab}(x) = 2 f_{(a|b)}(x), \]  

\[ \Delta^f S = \int d^3x \frac{\delta S}{\delta g_{ab}(x)} \Delta^f g_{ab}(x) = -2 \int d^3x \left[ \frac{\delta S}{\delta g_{ab}(x)} \right]_{|b} f_{a}(x) \equiv 0. \]  

Therefore, \( S \) is actually the functional not of the three-dimensional metric, but of three-dimensional geometry that encompasses the equivalence class of the three-dimensional metrics connected by the coordinate transformations.

The equation (7.39) is called the Einstein-Hamilton-Jacobi equation (EHJ) and it represents the basic equation of the classical geometrodynamics. Actually, this equation contains the entire dynamics of the classical theory of gravity. In accordance with the results of Sec. 2.1., the evolution of a gravitating system is restored as a solution of systems (2.47), (2.48), where the gauges are chosen to be explicitly time-dependent.

The Einstein-Hamilton-Jacobi theory is a semiclassical approximation to the quantum theory. However, the gravitation is a degenerate system with constraints and this approximation has the special gauge-theoretic features which are follows.

In the semiclassical approximation, the amplitude of the state of the quantum system is described in the coordinate representation by the superposition of the functions of the form \( \exp\{iS(q, \alpha)\} \) with different values of parameters \( \alpha \), and with the corresponding amplitudes \( P(\alpha) \). This superposition forms the wave-package

\[ \Psi = \int d\alpha P(\alpha) e^{iS(q,\alpha)}. \]  

(7.44)

The center of gravity of such package or the point in the coordinate space describing the classical system position is determined as a stationary point of the integral (7.44) with respect to the variables \( \alpha \),

\[ \frac{\partial S}{\partial \alpha^A} = \beta_A, \]  

(7.45)

where the parameters \( \beta_A = -\frac{1}{i} \partial \ln P(\alpha)/\partial \alpha^A \) depend on the form of the amplitude \( A(\alpha) \). In non-degenerate systems, the equations (7.45) contain all the necessary information to find the classical evolution. In systems with constraints, the equations (7.45) or (2.47) are not sufficient, since they define the surface in the space of the phase coordinates that consists of a set of the classical trajectories describing the same physical state and connected by the transformations of the “group” of invariance of the action (2.30). To select a single physical representative, it is necessary to impose the gauge (2.32), i.e., to introduce another surface in the space of the variables \( q \), described by the equation (2.48), that have a single intersection point with the initial one.
The space of three-dimensional geometries, where the action $S$ is determined as a solution of EHJ equation (7.40), is usually called a superspace\(^3\). However, since we can operate in a constructive way only in terms of the three-dimensional metrics, we use the term “superspace” to call the manifold of the three-dimensional metrics $g_{a b}(\mathbf{x})$, where the action $S$ satisfying equations (7.40)-(7.42) is determined.

Now let us proceed to the construction of the exact quantum theory, whose quasi-classical approximation would give the EHJ theory.

**Dirac quantization method and the Wheeler-DeWitt equation**

An alternative to the ADM quantization approach is the Dirac quantization method for theories subject to constraints \([40]\). The operators of the original phase space $\Phi = (\hat{q}, \hat{p})$, in contrast to the previous ADM method, are treated as independent and satisfying the usual canonical commutation relations

$$[\hat{\Phi}, \hat{\Phi}'] = i\{\Phi, \Phi'\}. \tag{7.46}$$

These operators are defined on a maximally wide space of vectors. However, the vectors of the physically admissible states may not be arbitrary, but only those that satisfy the quantum constraints as the equations

$$\hat{H}_\mu(\hat{q}, \hat{p}) |\Psi\rangle = 0 \tag{7.47}$$

and the Schrödinger equation

$$\frac{\partial |\Psi\rangle}{\partial t} = \hat{H}_0(\hat{q}, \hat{p}) |\Psi\rangle. \tag{7.48}$$

The latter reduces to the independence $|\Psi\rangle$ on $t$ for the special case of a closed world with $\hat{H}_0 \equiv 0$.

The consistency of the system of equations (7.47), (7.48) is guaranteed by the involution relations (2.29), provided that one choose such order of operators in the constraints $\hat{H}_\mu$ and in the Hamiltonian $\hat{H}_0$, that the Poisson brackets in (2.29) are mapped into the quantum commutators constructed on the basis of (7.46):

$$[\hat{H}_\mu, \hat{H}_\nu] = \hat{U}^\alpha_{\mu \nu} \hat{H}_\alpha, \quad [\hat{H}_0, \hat{H}_\mu] = \hat{V}^\nu_{\mu} \hat{H}_\nu, \tag{7.49}$$

with the operators of structure functions $\hat{U}^\alpha_{\mu \nu}$ and $\hat{V}^\nu_{\mu}$ standing to the left of the operators of constraints.

In the coordinate representation of the commutation relations (7.46), the momenta $\hat{p} = \partial/\partial \hat{q}$, the coordinates $\hat{q}$ are $c$-numbers, and $|\Psi\rangle = \Psi(q) = \Psi[g_{a b}(\mathbf{x})]$, so that the equation (7.47) takes the form

$$H_\mu(\hat{q}, \hat{\partial}/\partial \hat{q}) \Psi(q) = 0, \tag{7.50}$$

\(^3\)It should not be confused with the superspace in the modern supersymmetric theories. These are different notions.
with some particular choice of the operator realization $H_\mu(q, \partial/i\partial q)$ of $\hat{H}_\mu$, which includes a particular operator ordering of $q$ and $\hat{p}$ and possible quantum correction terms $O(\hbar)$.\footnote{The inclusion of these terms will require to reinstate explicitly the Planck constant $\hbar$ which will be done later.}

Thus, in quantum geometrodynamics with $\partial/i\partial q \equiv \delta/i\delta g_{ab}(x)$ at least naively we have

$$\left\{-G_{abcd} \frac{\delta}{\delta g_{ab}} \frac{\delta}{\delta g_{cd}} - g^{1/2} 3R\right\} \Psi[g_{ab}] = 0,$$

(7.51)

$$-2g_{ac} \left[ \frac{\delta \Psi[g_{ab}]}{\delta g_{cb}} \right]_{b} = 0.$$

(7.52)

The first of these equations is a widely acclaimed Wheeler-DeWitt equation \cite{53} which, of course, represents an infinite set of equations – one per each spatial point. Its operator realization is thus far purely symbolic and will be discussed later. On the contrary, the operator ordering in (7.52) has a sufficiently strong geometrical ground because the equation (7.52), similarly to (7.43), incorporates the local coordinate invariance of $\Psi[g_{ab}(x)]$, which rather firmly fixes the operator realization of quantum momentum constraints.

The Wheeler-DeWitt equation is a basic equation of quantum geometrodynamics. Obviously, the quasi-classical approximation leads to its solution of the form

$$\Psi = e^{iS},$$

where $S$ satisfies the equations (7.40) and (7.42).

Thus, comparing this approach with the ADM quantization in reduced phase space, we find the two different situations. In quantum geometrodynamics – the Dirac quantization scheme – the operators $(\hat{q}, \hat{p})$ are independent but the physical states are subject to the constraints (7.47). In the ADM (or reduced phase space) quantization the operators $(q, p)$ identically satisfy the constraints but the states are arbitrary. And no unitary equivalence exists between these two approaches, because the unitary equivalent representations have the same commutation relations, whereas the commutation relations (7.46) are fundamentally different from (7.2).

Another difficulty is that the Dirac quantization scheme is intrinsically incomplete, because a priori it lacks the algorithm for the physical inner product of quantum states that would give quantum expectation values and probabilities. The usual $L^2$ inner product

$$\langle \Psi' | \Psi \rangle = \int dq \Psi'^\ast(q) \Psi(q)$$

(7.53)

does not make sense, because the physical states $\Psi(q)$, which are annihilated by constraints, have a distributional nature $\sim \delta(\hat{H}_\mu)$ and make this integral
divergent, $[\delta(\hat{H})]^2 \sim \infty$. This inner product should have a nontrivial integration measure restricting the range of integration over $q$ to a certain subspace corresponding to a given time slice of the whole spacetime (as long as physical time is hidden among the variables $q$). The attempt to interpret as the physical inner product the flux of the conserved current of the Wheeler-DeWitt equation (7.51) [53, 165, 182] – the analogue of the conserved current of the second-order Klein-Gordon equation,

$$j_{ab}(x) = G_{abcd} \left\{ \psi^* \frac{\delta \psi}{i \delta g_{cd}(x)} - \frac{\delta \psi^*}{i \delta g_{cd}(x)} \psi \right\},$$

(7.54)

requires justification from first principles of quantization. Moreover, it demands specification of the surface in the space of 3-metrics $q = g_{ab}(x)$ through which the flux of this current is running, this specification establishing the above mentioned association with the time slice of spacetime.

Consistency between the ADM quantization and quantum geometrodynamics as the Dirac quantization of gravity theory was attained in the series of works [204, 205, 206, 207, 208, 209, 210] where the questions of the above type were clarified and checked at least in the first nontrivial order of semiclassical expansion – one-loop approximation. It turned out that direct unitary equivalence between these quantization methods is impossible and can only be achieved by embedding the Dirac scheme into a wider quantization framework. This framework is the Batalin-Fradkin-Vilkovisky (BFV) operator quantization of constrained systems [213, 46, 214, 215, 216, 217, 218, 219, 220] stemming from the BRST method\(^5\) [211, 212].

7.4. Dirac quantization from Batalin-Fradkin-Vilkovisky formalism

Here we give a brief overview of the BRST/BFV formalism in gauge systems with first class constraints [211, 212, 213, 46, 214, 215, 217, 218, 219, 220]. Within the technique of the relativistic phase space of gauge and ghost fields [213, 46] we build the operator of the unitary evolution in their representation space and formulate the BRST invariant physical states. A special Batalin-Marnelius (BM) procedure of gauging out the BRST symmetry in the subspace of these states [221, 222, 223] leads to a well-defined physical inner product in the space of BRST singlets [204, 206, 208]. Then, the Dirac quantization scheme turns out to be a special truncation of the BFV formalism [206, 208, 210] – a particular realization of this BM gauge fixing procedure – which allows one to construct a path integral representation for the solution of the Wheeler-DeWitt equations. In the subsequent section these results are explicitly checked in the

\(^5\)BRST refers to Becchi, Rouet, Stora, and Tyutin [211, 212].
first one-loop order of semiclassical expansion which, in particular, allows one to formulate a unitary map from the Wheeler-DeWitt (Dirac) wave functions \( \Psi(q) \) to the physical states \( \Psi_{\text{phys}}(g) \) of the reduced phase space quantization.

**BFV formalism**

We begin by reminding the reader that the ADM canonical formalism incorporates splitting the full configuration space of \( g_{\mu\nu} \) into phase space coordinates \( q^i = g_{ab}(x) \), and non-dynamical Lagrange multipliers \( N^\mu = (N^a(x), N^n(x)) \) in the canonical action (2.27) or (2.124). The momenta \( p_i \) are conjugated to \( q^i \), whereas the variables \( N^\mu \) do not have conjugated momenta. In what follows we will use condensed canonical notations when the field labels carry together with discrete indices also spatial coordinates, and contraction of indices \( i \) or \( \mu \) implies spatial (but not time) integration. In open models with asymptotically flat (or other boundaries like horizons, etc.) \( H_0(q,p) \) represents the relevant surface integral specified by boundary conditions which are the part of physical setting for the gravitational system. Below we consider the case of spatially closed cosmology with \( H_0(q,p) = 0 \).

The diffeomorphism invariance of the theory has a manifestation in Poisson brackets algebra of constraints (2.29) with the structure functions \( U^\lambda_{\mu\nu} = U^\lambda_{\mu\nu}(q) \) which depend on phase-space coordinates \( q \). At the quantum level the classical constraints take the form of equations on physical quantum states (7.47). In the functional coordinate representation of quantum gravity they form the system of the Wheeler-DeWitt equations on the wave function \( \langle q | \Psi \rangle = \Psi(q) \). Their consistency requires the commutator algebra (7.49) to hold with the operator structure functions \( \hat{H}_\lambda \) standing to the left of \( \hat{H}_\lambda \).

This BFV quantization generalizes this Dirac quantization scheme by extending the representation space of the original phase space variables \( q^i, p_i \) to that of the relativistic phase space variables. The latter include together with \( q^i, p_i \) the canonically conjugated pairs of Lagrange multipliers and their momenta \( p_\mu \) and canonical pairs of Grassman (fermionic) ghosts \( C^\mu, \overline{P^\mu} \) and anti-ghosts \( \overline{C^\mu}, P^\mu \),

\[
Q^I, P_I = \hat{q}^i, p_i; \quad N^\mu, p_\mu; \quad C^\mu, \overline{P^\mu}; \quad \overline{C^\mu}, P^\mu,
\]

\[
[Q^I, P_J] = i \delta^I_J. \tag{7.56}
\]

Here \([A, B] \) denotes a supercommutator taking into account the Grassman parity \( n \) of \( A \) and \( B \), \([A, B] \) \( = AB - (-1)^{n(A) n(B)} BA \). In gravity theory with bosonic matter fields \( n(q) = n(N) = 0 \) and \( n(C) = n(\overline{C}) = n(P) = n(\overline{P}) = 1 \). The canonical commutation relations (7.56) represent the quantum version of classical Poisson superbrackets commutators \( \{Q^I, P_J \} = i \delta^I_J \) (we use units with \( \hbar = 1 \)). Ghost variables have Hermiticity properties compatible with these commutation relations

\[
C^\mu \dagger = C^\mu, \quad P_\mu \dagger = -P_\mu, \quad \overline{C^\mu} \dagger = -\overline{C^\mu}, \quad \overline{P^\mu} \dagger = \overline{P^\mu}. \tag{7.57}
\]
In what follows we will regularly omit the hat notation for operators (7.55) and only use it in case when they have to be distinguished from their \( c \)-number eigenvalues. The hat notations will as a rule be used for composite operators. State vectors in the representation space of all relativistic variables will be denoted by double ket notations \(|| \Psi \rangle \rangle\) (contrary to vectors \(|\Psi \rangle\) in the representation space of the original operators \((\hat{q}, \hat{p}_i)\)). The coordinate representation will be introduced as follows:

\[
|| Q \rangle \rangle \equiv || q, N, C, \bar{C} \rangle \rangle, \hat{Q}^I \langle\langle Q \rangle \rangle = Q^I \langle\langle Q \rangle \rangle, \Psi(Q) = \langle\langle Q |\langle\langle \Psi \rangle \rangle, (7.58)
\]

\[
\langle\langle \Psi_1 |\langle\langle \Psi_2 \rangle \rangle = \int dQ \Psi_1^\ast(Q) \Psi_2(Q), (7.59)
\]

where the BRST inner product is defined in the sense of Berezin integration over \( Q \). To finish description of notations for BFV formalism we mention that we will also need the momentum representation in the sector of the Lagrangian multipliers with the interchanged roles of \( N^\mu \) and \( p_\mu \). It will be denoted by tilde, and the corresponding set of variables will look like

\[
\tilde{Q}^I, \tilde{P}^I = q^i, p^i, -N^\mu; C^\mu, \bar{P}^\mu; \bar{C}^\mu, P^\mu; \bar{C}^\mu, \bar{P}^\mu, \tilde{\Psi}(\tilde{Q}) = \langle\langle \tilde{Q} \rangle \rangle. (7.60)
\]

The basic object of the BRST/BFV technique is the nilpotent fermionic BRST operator \( \hat{\Omega} \) acting in the space of \(|\langle\langle \Psi \rangle \rangle\) and satisfying the master equation

\[
[\hat{\Omega}, \hat{\Omega}] \equiv \hat{\Omega}^2 = 0. (7.61)
\]

This equation allows one to look for the solution as an expansion in powers of the ghosts \( C^\mu \) and their momenta \( P_\mu \) starting with the combination \( \hat{\Omega} = p_\alpha \bar{P}^\alpha + C^\mu \hat{H}_\mu + O(PC^2) \). The coefficients of this expansion \( \hat{H}_\mu, \hat{U}_{\lambda}^\lambda, \hat{U}_{\mu\lambda}^\lambda, ... \) are the structure functions of the gauge algebra of constraints beginning with (7.49) – higher order structure functions follow from applying the Jacobi identity to multiple commutators of (7.49) with \( \hat{H}_\sigma \) [214, 215, 219]. In non-supersymmetric gravity theory, which we consider here, this sequence terminates at \( \hat{U}_{\mu\nu\alpha}^\lambda = 0 \), and \( \hat{\Omega} \) takes the form

\[
\hat{\Omega} = p_\alpha \bar{P}^\alpha + C^\mu \hat{H}_\mu + \frac{1}{2} C^\nu C^\rho \hat{U}_{\mu\nu}^\lambda P_\lambda, \quad \hat{\Omega}^\dagger = \hat{\Omega}. (7.62)
\]

It is Hermitian in the BRST inner product (7.59) in accordance with the Hermiticity properties of ghost variables (7.57), provided the quantum Dirac constraints have the anti-Hermitian part\(^6\)

\[
\hat{H}_\mu - \hat{H}_\mu^\dagger = i \hat{U}_{\mu\lambda}^\lambda. (7.63)
\]

\(^6\)In higher rank gauge theories with nonvanishing higher order structure functions this Hermiticity properties are modified by their higher order contributions.
In models with \( H_0(q,p) \neq 0 \) this BRST operator determines also its BRST extension \( \hat{\mathcal{H}} = \hat{H}_0 + O(CP) \) by the equation \([\hat{\Omega}, \hat{H}_0] = 0\) and the so-called unitarizing Hamiltonian
\[
\hat{\mathcal{H}}_\Phi = \hat{H}_0 + \frac{1}{i} [\hat{\Phi}, \hat{\Omega}].
\] (7.64)

It explicitly depends on the gauge fermion \( \Phi \), \( n(\hat{\Phi}) = 1 \), which provides gauge fixing in the BRST/BFV formalism. In parametrization invariant closed cosmology \( \hat{H}_0 = 0 \), and the unitarizing Hamiltonian reduces to the commutator of the BRST operator and gauge fermion.

The unitary evolution operator \( \hat{U}_\Phi(t, t_-) \) acting in the space of \( ||\Psi|| \) is a solution of the following Cauchy problem:
\[
i\hbar \frac{\partial}{\partial t} \hat{U}_\Phi(t, t_-) = \hat{\mathcal{H}}_\Phi \hat{U}_\Phi(t, t_-), \quad \hat{U}_\Phi(t_-, t_-) = I.
\] (7.65)

It is obvious that from \([\hat{\Omega}, [\hat{\Phi}, \hat{\Omega}]] \equiv 0\) and \([\hat{\Omega}, \hat{\mathcal{H}}_\Phi] = 0\) the BRST operator is a constant of motion in this evolution,
\[
[\hat{\Omega}, \hat{U}_\Phi(t, t_-)] = 0,
\] (7.66)

so that it plays the role of the conserved BRST charge and serves as a generator of BRST transformations in the relativistic phase space.

In the coordinate representation the kernel of the unitary evolution has a representation of the canonical path integral
\[
\hat{U}_\Phi(t_+, Q_+ | t_-, Q_-) = \langle \langle Q_+ || \hat{U}_\Phi(t_+, t_-) || Q_- \rangle \rangle
\]
\[
= \int_{Q(t_\pm) = Q_\pm} \mathcal{D}[Q, P] \exp \left\{ i \int_{t_-}^{t_+} dt \left( P_1 \dot{Q}_1 - \mathcal{H}_\Phi(Q, P) \right) \right\},
\] (7.67)

where \( \mathcal{H}_\Phi(Q, P) \) is the \(QP\)-symbol of the unitarizing Hamiltonian given by the Poisson superbracket of \( c \)-number symbols \( \Phi \) and \( \Omega \) of the operator gauge fermion and BRST charge \( \mathcal{H}_\Phi(Q, P) = \{ \Phi, \Omega \} \) (remember that \( H_0 = 0 \)). Also, \( \mathcal{D}[Q, P] \) is a Liouville integration measure in the full boson-fermion phase space of \( c \)-number histories:
\[
\mathcal{D}[Q, P] = \prod_t dQ(t) \prod_{t^*} dP(t^*).
\] (7.68)

The difference between the set of points \( t = (t_N, ... t_1) \) and \( t^* = (t_{N+1}^*, ... t_1^*) \), \( N \to \infty \), over which the product of integration measure factors is taken, reflects the typical slicing of the path integral into a sequence of multiple integrals in the decomposition of the full time segment \([t_+, t_-]\) into infinitesimal pieces. This decomposition, \( t_+ > t_N > t_{N-1} > ... > t_1 > t_- \), \( t_+ > t_{N+1}^* > t_N > t_N^* > ... > t_1 > t_1^* > t_- \), implies that the points \( t_i^* \), at which the integrated
momenta are taken, are associated with “interiors” of the segments \([t_{i+1}, t_i]\)
whose boundaries carry the integrated coordinates – so that the number of
momentum integrations is by one larger than those of coordinate ones.

In the momentum representation for Lagrangian multipliers \(\tilde{Q}^I = q^i, \pi^\mu, C^\mu, \bar{C}_\mu\), cf. (7.60), the unitary evolution kernel has a similar path integral rep-
resentation:

\[
\tilde{U}_\Phi(t_+, \tilde{Q}_+ | t_-, \tilde{Q}_-) \equiv \langle \langle \tilde{Q}_+ \| \tilde{U}_\Phi(t_+, t_-) \| \tilde{Q}_- \rangle \rangle = \int_{\tilde{Q}(t_\pm) = \tilde{Q}_\pm} \tilde{D}[Q, P] \exp \left\{ i \int_{t_-}^{t_+} dt \left( \tilde{P}_I \dot{\tilde{Q}}^I - \mathcal{H}_\Phi(Q, P) \right) \right\}, \tag{7.69}
\]

(7.70)

and is of course related by the Fourier transform to the kernel (7.67)

\[
\tilde{U}_\Phi(t_+, \tilde{Q}_+ | t_-, \tilde{Q}_-) = \int dN_+ dN_- e^{-ip_+ N_+} U_\Phi(t_+, Q_+, t_-, Q_-) e^{ip_- N_-} \tag{7.71}
\]

in full accordance with the fact that two symplectic forms in the integrands of
path integrals on the left and right hand sides here are related by

\[
\int_{t_-}^{t_+} dt \tilde{P} \tilde{Q} = \int_{t_-}^{t_+} dt P \dot{Q} - p_+ N_+ + p_- N_-, p_\pm \equiv \pi(t_\pm), N_\pm \equiv N(t_\pm). \tag{7.72}
\]

The principal theorem of the BFV quantization is that the matrix elements of
the unitary evolution operator \(\hat{U}_\Phi(t, t_-)\) between the BRST-invariant physical
states annihilated by \(\hat{\Omega}\) are independent of the choice of the gauge fermion [219]:

\[
\hat{\Omega} \| \Psi_{1,2}) = 0 \implies \delta_{\Phi} \langle \langle \Psi_1 \| \hat{U}_\Phi(t_+, t_-) \| \Psi_2 \rangle \rangle = 0. \tag{7.73}
\]

The logic of the above BRST/BFV construction is based on the observation
that relativistic gauge conditions, involving time derivatives of Lagrange multi-
pliers, make the latter propagating and having nonvanishing canonical momenta
\(p_\mu\) which are absent in the original action. To compensate the contribution of
these artificially introduced degrees of freedom and the degrees of freedom which
have to be excluded by first class constraints one introduces dynamical ghosts
and antighosts of the statistics opposite to those of \(\hat{H}_\mu\). Due to statistics they
effectively subtract in quantum loops the contribution of these gauge degrees
of freedom. However, a similar subtraction should be done in external lines of
Feynman diagrams, which means that not all quantum states in BRST space
are physical. Physical states form a subspace belonging to the kernel of the
BRST operator. In this subspace due to the above theorem the transition am-
plitudes and quantum averages are independent of the choice of gauge fixing
procedure – the corner stone of quantizing the gauge invariant systems.
Batalin-Marnelius gauge fixing and the physical inner product

Physical states should be BRST invariant \( \hat{\Omega} \langle \Psi \rangle = 0 \). This equation does not uniquely select its solution because in view of the nilpotent nature of \( \hat{\Omega} \) the BRST transformed state \( \langle \Psi \rangle' = \langle \Psi \rangle + \hat{\Omega} \langle \Phi \rangle \),

\[
\langle \Psi \rangle' = \langle \Psi \rangle + \hat{\Omega} \langle \Phi \rangle,
\]

(7.74)

with an arbitrary \( \langle \Phi \rangle \) also satisfies the BRST equation. This invariance results in the problem of constructing (or regulating) the physical inner product.

Problem is that the original inner product (7.59) for physical states represents the \( 0 \times \infty \)-indeterminacy. The essence of this indeterminacy can be qualitatively explained by the fact that squaring of the physical state \( \Psi(Q) \sim \delta(\hat{\Omega}) \) in (7.59) gives a divergent factor whereas the integration over Grassman variables multiplies it by zero. This inner product can be regulated by transforming the BRST-invariant state \( \langle \Phi \rangle \) to a special gauge as it was suggested by Batalin and Marnelius in [221, 222] (see also [223]):

\[
\langle \Psi \rangle \rightarrow \langle \Psi_{BM} \rangle : \hat{P}_{\mu} \langle \Psi_{BM} \rangle = 0, \hat{N}^{\mu} \langle \Psi_{BM} \rangle = 0,
\]

(7.75)

\[
\hat{P}_{\mu} \langle \Psi_{BM} \rangle = 0
\]

(7.76)

(The last condition is in fact a corollary of the second one, \( [\hat{\Omega}, N_{\mu}] \langle \Psi_{BM} \rangle = 0 \)).

The consistency of this gauge with the BRST-invariance of \( \langle \Psi_{BM} \rangle \) implies that it also satisfies the quantum Dirac constraints

\[
0 = \frac{1}{i} [ \hat{\Omega}, \hat{P}_{\mu} \langle \Psi_{BM} \rangle ] = (\hat{H}_{\mu} + C^{\nu} \hat{U}_{\nu}^{\lambda} \hat{P}_{\lambda} \langle \Psi_{BM} \rangle = \hat{H}_{\mu} \langle \Psi_{BM} \rangle,
\]

(7.77)

and its wave function is independent of ghost variables,

\[
\frac{\partial}{\partial C^{\nu}} \Psi_{BM}(Q) = 0, \quad \frac{\partial}{\partial C^{\nu}} \Psi_{BM}(Q) = 0.
\]

(7.78)

Together with (7.75) this means that \( \Psi_{BM}(Q) \) has the form

\[
\langle Q \langle \Psi_{BM} \rangle \rangle = \langle q \mid \Psi \rangle \delta(N) = \Psi(q) \delta(N),
\]

(7.79)

where the “matter” part \( \Psi(q) \) satisfies quantum Dirac constraints in the coordinate representation,

\[
\hat{H}_{\mu} \Psi(q) = 0.
\]

(7.80)

According to [221, 222] the physical inner product of wave functions \( \langle \Psi_{BM} \rangle \) can be regularized by a special operator-valued measure which is explicitly built with the aid of the gauge fermion \( \Phi \)

\[
\langle \langle \Psi' \mid \Psi \rangle \rangle_{phys} = \langle \langle \Psi'_{BM} \mid e^{[\Phi, \hat{\Omega}]} \langle \Psi_{BM} \rangle \rangle.
\]

(7.81)
The choice of this fermion is immaterial because
\[ \delta \phi \langle (\Psi' \| \Psi) \rangle_{\text{phys}} = \int_0^1 ds \langle (\hat{\Psi}_{BM}'' \| e^{(1-s)[\hat{\phi}, \hat{\Omega}]} [\delta \hat{\phi}, \hat{\Omega}] e^{s[\hat{\phi}, \hat{\Omega}]} \| \Psi_{BM}) \rangle = 0, \]  
(7.82)
since \([\hat{\Omega}, [\hat{\phi}, \hat{\Omega}]] = 0\) and \([\hat{\Omega} \| \Psi_{BM}) = 0\). However, with a special choice of this fermion the physical inner product (7.81) resolves the \(0 \times \infty\) uncertainty and becomes well defined [221].

This is easy to show if this fermion is constructed with the aid of gauge conditions functions \(\hat{\chi} = \chi^\mu(\hat{q}, \hat{p})\) which commute with themselves, \([\hat{\chi}^\mu, \hat{\chi}^\nu] = 0\), and with the structure functions operators, \([\hat{\chi}^\mu, \hat{U}_\alpha^\lambda = 0\) (in gravity theory this is coordinate gauge conditions \(\chi^\mu(q)\) commuting with \(U^\lambda_{\alpha\beta}(q)\)). For the fermion \(\hat{\Phi}_{BM} = \hat{C}_\mu \hat{\chi}^\mu\) we have
\[ [\hat{\Phi}_{BM}, \hat{\Omega}] = i\hbar \hat{\chi}^\mu + i\hat{C}_\mu \hat{J}^\nu C^\nu, \quad \hat{J}^\nu = \frac{1}{i} [\hat{\chi}^\mu, \hat{H}_\nu]. \]  
(7.83)

Then the physical inner product for Batalin-Marnelius wave functions (7.79) in the coordinate representation with \(p_\mu = \partial / i \partial N^\mu\) takes the form
\[ \langle (\Psi' \| \Psi) \rangle_{\text{phys}} = \int dq \ dN \ dC \ d\bar{C} \ d\bar{\Psi}^\nu(q) \delta(N) e^{-i\bar{C}_\mu \hat{J}^\nu C^\nu + \bar{\chi}^\mu \bar{\nabla}^\nu} \delta(N) \Psi(q) = \langle \Psi' \| \int d\pi dC d\bar{C} e^{-i\bar{C}_\mu \hat{J}^\nu C^\nu + i\bar{\chi}^\mu} \mid \Psi \rangle. \]  
(7.84)

This leads to the physical inner product as a special operator valued measure \(\hat{M}\) acting in the space of Dirac wave functions \(|\Psi'\rangle\) and \(|\Psi\rangle\) endowed with the auxiliary \(L^2\) inner product (7.53),
\[ \langle \Psi' \| \Psi \rangle_{\text{phys}} = \langle (\Psi' \| \Psi) \rangle_{\text{phys}} = \langle \Psi' \| \hat{M} \| \Psi \rangle, \]  
(7.85)
\[ \hat{M} = \int d\pi dC d\bar{C} e^{-i\bar{C}_\mu \hat{J}^\nu C^\nu + i\bar{\chi}^\mu} = \delta(\bar{\chi}) \det(\hat{J}_\mu^\nu (1 + O(\{\hat{\chi}, \hat{J}\}))). \]  
(7.86)

This measure is known in quadratures as an explicit integral over ghost fields and Lagrangian multipliers momenta. In the leading semiclassical order in \([\hat{\chi}, \hat{J}] = O(h)\), this integral allows one to disentangle the delta function of gauge conditions, \(\delta(\hat{\chi}) = \prod_{\mu} \delta(\hat{\chi}^\mu)\), which is well defined in view of their commutativity.\(^7\)

The expression (7.86) is the analogue of the time-local measure in the canonical Faddeev-Popov path integral [202] with the operator (7.83) semiclassically represented by the Poisson bracket \(\hat{J}^\nu_\mu = \{\chi^\mu, H_\nu\}\). As we will show below by direct calculations in the one-loop (subleading in \(h\)) order this inner product [205, 206, 208, 209]

\[ \langle \Psi' \| \Psi \rangle_{\text{phys}} = \int dq \Psi'^\nu(q) \delta(\hat{\chi}(q)) \det(\hat{J}_\mu^\nu) \Psi(q) + O(h) \]  
(7.87)

\(^7\)This, however, does not save us from extra corrections, because \(J^\nu_\mu(q, \hat{p})\) depends on the momentum \(\hat{p}_i\) and is a differential operator acting in the space of \(q\).
for semiclassical wave functions $\Psi(q)$ and $\Psi'(q)$ is independent of the choice of $\chi^\mu(q)$ and is consistent with the Hermiticity property\(^8\) (7.63) [209].

**Path integral representation in Dirac quantization scheme**

Transition to the BM gauge can in fact be obtained by a simple procedure of truncation of BRST invariant wave functions to the sector of “matter” variables, that was suggested in [206, 208]. Introduce a wave function $\Psi(q)$ in the representation space of $\hat{q}^i, \hat{p}_i$ which can be obtained from the solution $|\Psi\rangle$ of the BRST equation $\hat{\Omega}(\Psi) = 0$ by

$$\Psi(q) = \int dN \Psi(q, N, C, \bar{C}) \bigg|_{C=0}. \quad (7.88)$$

As we show below, from the BRST equation it follows that this function satisfies quantum Dirac constraints and is independent of the antighost variable $\bar{C}$ (that is why the argument $\bar{C}$ of the right-hand side is omitted on the left-hand side of this definition):

$$\hat{H}_\mu \Psi(q) = 0, \quad \frac{\partial}{\partial C_\mu} \Psi(q) = 0. \quad (7.89)$$

Both properties follow from the BRST equation which in the coordinate representation reads as

$$\hat{\Omega}(\Psi) = \left( \frac{\partial}{\partial N^\mu} \frac{\partial}{\partial C_\mu} + C^\mu \hat{H}_\mu + \frac{i}{2} C^\nu C^\mu \hat{U}_{\mu\nu}^{\lambda} \frac{\partial}{\partial C^\lambda} \right) \Psi(Q) = 0. \quad (7.90)$$

Integrating this equation over the Lagrange multipliers $N$ in infinite limits and assuming that $\Psi(Q)$ falls off sufficiently rapidly at $N \to \pm \infty$, one finds that the first term of (7.90) vanishes. Then one can differentiate the result with respect to the ghost field $C_\mu$ and subsequently put $C = 0$. This proves the first of relations (7.89). The second relation follows from multiplying Eq. (7.90) by $N^\mu$, putting $C = 0$ and integrating over $N$ by parts in the remaining first term of this equation

$$0 = \int dN \ N^\mu \hat{\Omega}(\Psi) \bigg|_{C=0} = \int dN \ N^\mu \frac{\partial}{\partial N^\mu} \frac{\partial}{\partial C_\nu} \Psi(Q) \bigg|_{C=0}. \quad (7.91)$$

This truncation of the BRST quantization scheme to the Dirac quantization suggested in [206, 208] serves in fact as a realization of the Batalin-Marnelius gauge fixing (7.75) of the BRST symmetry (7.74). To put the generic BRST state into the Batalin-Marnelius gauge it is enough to take the bosonic “body”

---

\(^8\)Even though this Hermiticity property holds with respect to another – auxiliary – inner product (7.53) different from (7.87).
of its wave function, integrate it over the Lagrange multipliers argument $N$ and multiply by $\delta(N)$:

$$\langle\langle \Psi \rangle\rangle \rightarrow \langle\langle \Psi_{BM} \rangle\rangle : \quad \Psi_{BM}(Q) = \delta(N) \int dN' \Psi(q, N', C, \bar{C}) \bigg|_{C=0}. \quad (7.92)$$

Truncation similar to (7.88) for the kernel of the unitary evolution (7.67) reads

$$U(q_+, q_-) = \int dN_+ dN_- U_\phi(t_+, Q_+ | t_-, Q_-) \bigg|_{C_\pm=0}. \quad (7.93)$$

This object can be represented as a matrix element of $\hat{U}_\phi(t_+, t_-)$ between the following two states $\langle\langle \Psi_\pm \rangle\rangle$ which are both zero vectors of the Lagrangian multiplier momentum and trivially satisfy the BRST equation:

$$\delta_\Phi U(q_+, q_-) \equiv \langle\langle \Psi_+ || \hat{U}_\phi(t_+, t_-) || \Psi_- \rangle\rangle = \langle\langle \Psi_+ || \frac{1}{i} [\hat{\phi}, \hat{\Omega}] \hat{U}_\phi(t_+, t_-) || \Psi_- \rangle\rangle = 0. \quad (7.94)$$

which guarantees the uniqueness of its definition. The second important property is that the kernel (7.93) is independent of $t_\pm$ in parametrization invariant theory with $H_0 = 0$ because

$$i \frac{\partial}{\partial t_+} U(q_+, q_-) = \langle\langle \Psi_+ || \frac{1}{i} [\hat{\phi}, \hat{\Omega}] \hat{U}_\phi(t_+, t_-) || \Psi_- \rangle\rangle = 0 \quad (7.98)$$

in view of the Schrödinger equation (7.65) for $\hat{U}_\phi(t_+, t_-)$. This allows one to omit $\Phi$ and $t_\pm$ labels in the left-hand side of the definition (7.93).

Finally, applying the same derivations as in (7.90)-(7.91) to the main BRST equation (7.66) for $\hat{U}_\phi(t_+, t_-)$, one proves that this kernel is independent of antighosts $\bar{C}_\mu^\pm$ and satisfies quantum Dirac constraints with respect to both arguments

$$\hat{H}_\mu U(q, q') = 0, \quad U(q, q') \overset{\text{triv}}{\rightarrow} H_\mu^\prime = 0. \quad (7.99)$$

Integration over $N_\pm$ in (7.93) implies that this kernel can be interpreted as the unitary evolution kernel in the momentum representation of Lagrange multipliers (7.71) at zero values of $p_\pm$,

$$U(q, q') = \int dN_+ dN_- e^{-ip_+ N_+} U_\phi(t_+, Q_+ | t_-, Q_-) e^{ip_- N_-} \bigg|_{p_\pm=C_\pm=0}$$

$$= \hat{U}_\phi(t_+, Q_+ | t_-, Q_-) \bigg|_{p_\pm=C_\pm=0}. \quad (7.100)$$
Therefore, it has the path integral representation (7.69) with the symbol of the unitarizing Hamiltonian – the Poisson superbracket of the gauge fermion Φ and BRST charge Ω,

\[
U(q_+, q_-) = \int \tilde{D}[Q, P] \exp \left[ i \int_{t_-}^{t_+} dt \left( \tilde{P}_I \dot{Q}^I - \{\Phi, \Omega\} \right) \right] \bigg|_{\tilde{Q}(t_{\pm}) = \tilde{Q}_{\pm}} \quad (7.101)
\]

There exists a special choice of the gauge fermion, \( \Phi = P^\mu N^\mu + \bar{C}^\mu \chi^\mu(q) \), which generates in the Lagrangian formalism (after integrating out the phase space momenta) calculationally most useful relativistic gauge conditions of the form \( \dot{N}^\mu - \dot{\chi}^\mu(q) = 0 \) [213, 46, 217, 218]. With this fermion (note that it differs from the BM gauge fermion \( \Phi = \bar{C}^\mu \chi^\mu(q) \)) the path integral becomes

\[
U(q_+, q_-) = \int \tilde{D}[Q, P] \exp \left[ i \int_{t_-}^{t_+} dt \left( \tilde{P}_I \dot{Q}^I - N^\mu H_\mu - p_\mu \chi^\mu - \bar{C}^\mu J_\nu^\mu C_{\nu} - P_\alpha (\tilde{P}^\alpha + U_{\mu\nu} N^\mu C^\nu) \right) \right] \bigg|_{\tilde{Q}(t_{\pm}) = \tilde{Q}_{\pm}, p_{\pm} = C_{\pm} = 0} \quad (7.102)
\]

A standard procedure of transition to the unitary gauge \( \chi^\mu(q) = 0 \) consists in rescaling the gauge function \( \chi^\mu \) by a small numerical parameter \( \varepsilon \), \( \chi^\mu \rightarrow \chi^\mu / \varepsilon \) and making the change of integration variables \( p_\mu \) and \( \bar{C}^\mu \) (with a unit Jacobian)

\[
p_\mu \rightarrow \varepsilon p_\mu, \quad \bar{C}^\mu \rightarrow \varepsilon \bar{C}^\mu.
\quad (7.103)
\]

Note that boundary conditions at \( t = t_{\pm} \) admit this change of variables, and all this does not affect the answer for \( U(q, q') \) in view of its gauge independence. Then in the limit \( \varepsilon \rightarrow 0 \) the kinetic terms \( -N^\mu \dot{p}_\mu \) and \( \tilde{P}^\alpha \tilde{C}_\alpha \) disappear from the integrand of (7.102), Gaussian integration over ghost momenta does not contribute any field-dependent measure, and the projector to Dirac states takes the form of the usual canonical Faddeev-Popov path integral

\[
U(q_+, q_-) = \int D[q, p] \, DN \, D\pi \, DC \, D\bar{C} \times \\
\times \exp \left\{ i \int_{t_-}^{t_+} dt \left( p_i \dot{q}^i - N^\mu H_\mu - p_\mu \chi^\mu - \bar{C}^\mu J_\nu^\mu C_{\nu} \right) \right\} \bigg|_{q(t_{\pm}) = q_{\pm}, p_{\pm} = C_{\pm} = 0}
\]

\[
= \int_{q(t_{\pm}) = q_{\pm}} D[q, p] \, DN \left( \prod_{t \neq t_{\pm}} \delta(\chi) \right) \det J_\mu^\nu e^{i \int_{t_-}^{t_+} dt \left( p_i \dot{q}^i - N^\mu H_\mu \right)} \quad (7.104)
\]

in which, however, the gauge fixing factor \( \delta(\chi) \) \( \det J_\mu^\nu \) is absent at the both end points \( t_{\pm} \). This completely specifies the expression in the right-hand side (7.11) formulated above and proves all its properties – projection on the subspace of Dirac constraints (7.99), gauge independence (7.97) and independence of the choice of \( t_{\pm} \) (7.98). Its relation to the physical sector evolution operator standing in the left-hand side of (7.11) is discussed below.
7.5. Semiclassical approximation

The formalism of the above type should stand verification by a viable calculational scheme. Such a scheme applicable to theories of a general type is the semiclassical loop expansion in powers of $\hbar$. So here we reinstate the $\hbar$ parameter (thus far chosen to be 1 in universal units) and develop this expansion in the first (one-loop) order linear in the Planck constant.

**Operator realization of constraints**

The first thing to do is to find the operator realization of constraints $\hat{H}_\mu$ and structure functions $\hat{U}^\lambda_{\mu\nu}$ that would satisfy the quantum algebra (7.49) with the operators of structure functions standing to the left of the constraints (with the Planck constant reinstated the right-hand side of these algebraic relations should be multiplied by $\hbar$). Remarkably, in the approximation linear in $\hbar$ there exists in a closed form the following solution to this problem – the operators are given by the Weyl (symmetric in $\hat{q}$ and $\hat{p}$) ordering of the following expressions [206, 209]:

$$
\hat{H}_\mu = N_W \left\{ H_\mu(\hat{q}, \hat{p}) + \frac{i\hbar}{2} U^\nu_{\mu\nu}(\hat{q}, \hat{p}) + O(\hbar^2) \right\}, \quad (7.105)
$$

$$
\hat{U}^\lambda_{\mu\nu} = N_W \left\{ U^\lambda_{\mu\nu}(\hat{q}, \hat{p}) - \frac{i\hbar}{2} U^{\lambda\sigma}_{\mu\nu\sigma}(\hat{q}, \hat{p}) + O(\hbar^2) \right\}. \quad (7.106)
$$

Here $N_W$ is a symbol of Weyl ordering, and these algorithms are true for a generic theory subject to first class constraints having at the classical level the hierarchy of structure functions of the Poisson bracket algebra beginning with $H_\mu(q, p), U^\lambda_{\mu\nu}(q, p), U^{\lambda\sigma}_{\mu\nu\omega}(q, p), ...$, as mentioned above higher order structure functions following from the Jacobi identities applied to multiple commutators of (7.49) with $\hat{H}_\sigma$ [214, 215, 219]. In Einstein gravity theory, of course, $U^{\lambda\nu}_{\mu\nu} \equiv 0$. Note a nontrivial anti-Hermitian part of the constraint operator generated by the trace of the structure constant, which is compatible with the Hermiticity properties of the BFV operators (7.63) relative to the auxiliary inner product (7.53) in the space of $\Psi(q)$.

A similar algorithm holds for the physical observables – functions $O_I(q, p)$ on the phase space which are classically gauge invariant in a weak sense, that is commuting with classical constraints modulo the constraints themselves, $\{O_I, H_\mu\} = V_I^{\nu\mu} H_\nu$. They read [209]

$$
\hat{O}_I = N_W \left\{ O_I + \frac{i\hbar}{2} V^\lambda_{I\lambda} + O(\hbar^2) \right\}, \quad (7.107)
$$

$$
\hat{V}^\nu_{I\mu} = N_W \left\{ V^\nu_{I\mu} - \frac{i\hbar}{2} V^{\nu\sigma}_{I\mu\sigma} + O(\hbar^2) \right\}. \quad (7.108)
$$
7.5. Semiclassical approximation

This operator realization applies, in particular, to $\hat{H}_0$ and $\hat{V}_\mu^\nu \equiv \hat{V}_0^\mu_\nu$ of Eq.(7.49) in theories with a nonvanishing Hamiltonian $H_0(q, p)$.

**Semiclassical physical states**

Semiclassical wave functions are characterized by the Hamilton-Jacobi function $S(q)$ and pre-exponential factor $P(q)$ expandable in $\hbar$-series beginning with the one-loop order $O(\hbar^0)$:

$$\Psi(q) = P(q) \exp \left[ \frac{i}{\hbar} S(q) \right].$$

(7.109)

The general semiclassical solution of quantum constraints (7.99) with operators (7.105) was found in [205, 206, 208] in the form of the two-point kernel $U(q, q')$ “propagating” the initial data from $q'$ throughout the space of $q$:

$$U(q, q') = P(q, q') \exp \left[ \frac{i}{\hbar} S(q, q') \right].$$

(7.110)

In both expressions (7.109) and (7.110) the phase in the exponential satisfies the Hamilton-Jacobi equation (2.44), while the one-loop pre-exponential factor is subject to continuity type equation originating from the full quantum constraint in the approximation linear in $\hbar$ [205, 206, 209]:

$$\frac{\partial}{\partial q^i} \left( \nabla^i_{\mu} p^2 \right) = U^\lambda_{\mu\lambda} p^2,$$

(7.111)

$$\nabla^i_{\mu} \equiv \frac{\partial H_\mu}{\partial p_i} \Big|_{p = \partial S/\partial q^i}.$$

(7.112)

Note a nontrivial right-hand side in the “continuity” equation, generated by the anti-Hermitian part of $\hat{H}_\mu$.

For a two-point kernel the Hamilton-Jacobi function coincides with the principal Hamilton function $S(q, q')$ (action on the extremal joining points $q$ and $q'$) and the solution of the continuity equation can be found as a generalization of the Pauli-Van Vleck-Morette ansatz for the one-loop pre-exponential factor [224, 208] of the Schrödinger propagator. This generalization is nothing but a Faddeev-Popov gauge-fixing [225] procedure for a matrix of mixed second-order derivatives of the principal Hamilton function:

$$S_{ik'} = \frac{\partial^2 S(q, q')}{\partial q^i \partial q^{k'}}$$

(7.113)

which is degenerate in virtue of the Hamilton-Jacobi equations (2.44) giving rise to the left zero-value eigenvectors (7.112) and analogous right zero-vectors.

---

9If the observables satisfy the Poisson algebra $\{\mathcal{O}_I, \mathcal{O}_J\} = U^I_{IJ} \mathcal{O}_L + U^I_{IJ} H_\lambda, U^I_{IJ} = \text{const}$, the above expressions are also modified by the trace of structure constants $U^I_{IJ} [209]$. 
\[ \nabla_\mu S_{ik'} = 0, \quad S_{ik'} \nabla_\nu = 0, \quad \nabla_\nu S_{ik'} = 0, \quad \nabla_\nu \equiv \frac{\partial H_\nu (q', p')}{\partial p'_k} \bigg|_{p' = -\partial S/\partial q'} . \] (7.114)

The pre-exponential factor reads
\[ P = \left[ \frac{\det F_{ik'}}{J(q) J(q') \det c_{\mu\nu}} \right]^{1/2}, \] (7.115)
where \( F_{ik'} \) is a nondegenerate matrix of the initial action Hessian (7.113) supplied with a gauge-breaking term
\[ F_{ik'} = S_{ik'} + \chi^\mu_i c_{\mu\nu} \chi^\nu_{ik'}, \] (7.116)
and \( J(q) \) and \( J(q') \) are the Faddeev-Popov “ghost” determinants [225] compensating for the inclusion of this term. They are constructed with the aid of two sets of arbitrary covectors \( (\chi^\mu_i, \chi^\nu_{ik'}) \) ("gauge" conditions) satisfying the only requirement of the nondegeneracy of their ghost operators [205, 208, 206]:
\[ J^\mu_\nu (q) = \chi^\mu_i \nabla_\nu, \quad J(q) \equiv \det J^\mu_\nu (q) \neq 0, \]
\[ J^\mu_\nu (q') = \chi^\mu_i' \nabla_\nu', \quad J(q') \equiv \det J^\mu_\nu (q') \neq 0. \] (7.117)

The invertible gauge-fixing matrix \( c_{\mu\nu} \) and its determinant are the last ingredients of the generalized Pauli-Van Vleck-Morette ansatz (7.115).

**Reduction to the physical sector**

The interpretation of the semiclassical state (7.110), (7.115) is rather transparent in the physical sector of the theory [205, 206, 208]. The sector explicitly arises after the reduction to physical variables by disentangling them from the original phase space of \( (q^i, p_i) \) in a unitary gauge. Such a reduction in [205, 206, 208] was given for a special type of gauge conditions imposed only on phase space coordinates \( q^i, \chi^\mu (q) = 0 \). These coordinate gauge conditions determine the embedding of the \( (n - m) \)-dimensional space \( \Sigma \) of physical coordinates directly into the space of original coordinates \( q^i \) \(- \) superspace. This fact strongly simplifies the reduction of the semiclassical kernel (7.110), (7.115) to the physical sector, because this reduction in the main boils down to the embedding of the arguments of \( U(q, q') \) into the physical subspace \( \Sigma \). The geometry of this embedding, considered in much detail in [208], can be better described in special coordinates on superspace \( \bar{q}^i = (\xi^A, \theta^\mu) \), in which \( g^A, \ A = 1, ... n - m, \) serve as intrinsic coordinates on \( \Sigma \) (physical configuration coordinates), and \( \theta^\mu \) is determined by gauge conditions:
\[ q^i \rightarrow \bar{q}^i = (g^A, \theta^\mu), \quad q^i = e^i (g^A, \theta^\mu), \quad \theta^\mu = \chi^\mu(q). \] (7.118)
7.5. Semiclassical approximation

The equation of the surface $\Sigma$ in the new coordinates is $\theta^\mu = 0$, so that its embedding equations coincide with the above reparametrization equations at $\theta^\mu = 0$, $e^i(g) = e^i(g, 0)$

$$\Sigma : q^i = e^i(g), \quad \chi^\mu(e^i(g)) \equiv 0. \quad (7.119)$$

The relation between the integration measures on superspace $dq = d^nq$ and on $\Sigma$, $dg = d^{n-m}g$, involves the Jacobian of this reparametrization, built of the basis of vectors tangential and normal to $\Sigma$:

$$dg = dq \delta(\chi) \mathcal{M}, \quad \delta(\chi) = \prod_\mu \delta(\chi^\mu(q)), \quad \mathcal{M} = (\det[e^i_\mu, e^i_\nu])^{-1}, \quad (7.120)$$

involves the Jacobian of this reparametrization, built of the basis of vectors tangential and normal to $\Sigma$:

$$e^i_A = \frac{\partial e^i}{\partial g^A}, \quad e^i_\mu = \frac{\partial e^i}{\partial \theta^\mu}. \quad (7.121)$$

Note that $m$ covectors normal to the surface can be chosen as gradients of gauge conditions

$$\chi^\mu_i = \frac{\partial \chi^\mu}{\partial q^i}, \quad \chi^\mu_i e^i_\nu = \delta^\mu_\nu, \quad (7.122)$$

that can be identified with auxiliary covectors participating in the algorithm for the pre-exponential factor (7.115). With this identification the Faddeev-Popov operator $J^\mu_i(q, \partial S/\partial q)$ coincides with the operator $J^\mu_i(q)$ of this algorithm (which explains the use of the same notation).

On the same footing with $(e^i_A, e^i_\mu)$ as a full local basis one can also choose the set $(e^i_A, \nabla^i_\mu)$ with vectors $\nabla^i_\mu$ transversal to $\Sigma$ given by eq (7.112). The normal vectors of the first basis when expanded in the new basis, $e^i_\mu = J^{-1}_\mu^\nu \nabla^i_\nu + \Omega^A_\mu e^i_A$, have one expansion coefficient always determined by the inverse of the Faddeev-Popov matrix $J^{-1}_\mu^\nu$ and, thus, independent of the particular parametrization of $\Sigma$ by internal coordinates. The second coefficient is less universal and depends on a particular choice of this parametrization. Missing information about $\Omega^A_\mu$ does not prevent, however, from finding the relation between the determinants of matrices of the old and new bases:

$$\det[e^i_A, \nabla^i_\mu] = \frac{J}{\mathcal{M}}. \quad (7.123)$$

The reduction to physical sector in coordinate gauges follows after identifying $g^A$ with the physical coordinates. The corresponding conjugated momenta $p_A$ can be found from the transformation of the symplectic form restricted to the physical subspace (7.119)

$$\int dt p_i q^i = \int dt \left( p_i e^i_A \dot{g}^A + p_i \frac{\partial e^i(g, t)}{\partial t} \right), \quad (7.124)$$

$$p_A = p_i e^i_A, \quad (7.125)$$
as projections of the original momentum to the tangential components of the basis (7.121). The normal projections of $p_i$ should be found from the constraints $H_\mu(q, p) = 0$, the local uniqueness of their solution being granted by the nondegeneracy of the Faddeev-Popov determinant. Together with (7.119) this solution yields all the original phase space variables $(q^i, p_i)$ as known functions of the physical degrees of freedom $(g^A, p_A)$. The original action (2.124) reduced to physical sector (that is to the subspace of constraints and gauge conditions) acquires the usual canonical form with the physical Hamiltonian contributed by the second term of (7.124) and $H_0(q, p)$ when the latter is nonvanishing.

Note that generically, especially for systems with $H_0(q, p) = 0$, the canonical gauge conditions should explicitly depend on time, $\chi^\mu(q) = \chi^\mu(q, t)$, in order to generate the dynamical evolution in reduced phase space theory [208, 206]. Therefore, the reduced symplectic form generates a nontrivial contribution to the physical Hamiltonian, proportional to the time derivative of the embedding functions (7.119) explicitly depending on $t$, $q^i = e^i(g^A, t)$. The total physical Hamiltonian then takes the form

$$H_{\text{phys}}(g, \pi) = \left[ H_0(q, p) - p_i \frac{\partial e^i(g, t)}{\partial t} \right]_{q=q(g, \pi), p=p(g, \pi)},$$

where we retain for the sake of generality also the nonzero Hamiltonian $H_0(q, p)$.

Canonical quantization of such a classical system runs as usual along the lines of a particular representation and operator realization in the Hilbert space of the theory. In the one-loop (linear in $\hbar$) approximation with the Weyl ordering of the above Hamiltonian this quantization is basically exhausted by the unitary evolution kernel $U(t, g|t', g')$ of the Schrödinger equation. In the coordinate representation it is given by the well-known Pauli-Van Vleck-Morette ansatz [224]

$$U_{\text{phys}}(t, g|t', g') \equiv \left[ \det \frac{i}{2\pi\hbar} \frac{\partial^2 S(t, g|t', g')}{\partial g^A \partial g^{B'}} \right]^{1/2} \frac{1}{\pi} e^{i S(t, g|t', g')},$$

where the principal Hamilton function of the physical variables $S(t, g|t', g')$ is a classical action evaluated at the classical extremal passing the points $g'$ and $g$ respectively at initial $t'$ and final $t$ moments of time. The pre-exponential factor here is built of Van Vleck determinant and guarantees in the approximation linear in $\hbar$ the unitarity of the Schrödinger evolution of the physical states $\Psi(t, g)$

$$\Psi_{\text{phys}}(t, g) = \int dg' U_{\text{phys}}(t, g|t', g') \Psi_{\text{phys}}(t', g')$$

in the Hilbert space with a simple $L^2$ inner product (denoted by round brackets)

$$\langle \Psi_{\text{phys}}' | \Psi_{\text{phys}} \rangle \equiv \int dg \Psi_{\text{phys}}'^* (g) \Psi_{\text{phys}}(g).$$
The unitary map between the reduced phase space quantization of the above type and the Dirac quantization of Sects. 7.2.–7.4. consists in a special relation between the two-point kernel (7.110) with prefactor (7.115) and the Schrödinger evolution operator (7.127) [205, 206, 208]. This relation is based on the equality of the principal Hamilton functions in the original constrained theory and the reduced one and its corollary – the relation between their Van Vleck matrices

\[
S(t, g|t', g') = S(q, q') \bigg|_{q=e(g,t), q'=e(g',t')}, \tag{7.130}
\]

\[
S_{ik'}e^i_A e^{k'}_B = \frac{\partial^2 S(t, g|t', g')}{\partial g^A \partial g^{B'}}. \tag{7.131}
\]

Decomposing the gauge-fixed matrix (7.116) in the basis of vectors \((e^i_A, \nabla^i_\mu)\) and using (7.123) one then easily finds the needed relation [205, 206, 208]

\[
U_{\text{phys}}(t, g|t', g') = \left( \frac{J}{\mathcal{M}} \right)^{1/2} U(q, q') \left( \frac{J'}{\mathcal{M'}} \right)^{1/2} \bigg|_{q=e(g,t), q'=e(g',t')} + O(\hbar). \tag{7.132}
\]

This is the semiclassical (one-loop) implementation of the relation (7.11) formulated above.

This relation implies that the kernel \(U(q, q')\) similarly to the Schrödinger propagator \(U_{\text{phys}}(t, g|t', g')\) can be regarded as a propagator of the Dirac wave function \(\Psi(q)\) in superspace. Indeed, introducing the following map between \(\Psi(q)\) and \(\Psi_{\text{phys}}(t, g)\)

\[
\Psi_{\text{phys}}(g, t) = \left( \frac{J}{\mathcal{M}} \right)^{1/2} \Psi(q) \bigg|_{q=e(g,t)} + O(\hbar) \tag{7.133}
\]

and taking into account the relation (7.120) between the integration measures on superspace and the physical space \(\Sigma\), one finds that the propagation law (7.128) in \(g\)-space can be regarded as a projection onto \(\Sigma(t)\) of the following propagation of the Dirac wave function \(\Psi(q)\) from the initial Cauchy surface \(\Sigma(t')\) in \(q\)-space to the entire superspace:

\[
\Psi(q) = \int dq' U(q, q') \delta(\chi(q', t')) J \left( q', -\frac{\partial S(q, q')}{\partial q'} \right) \Psi(q') + O(\hbar). \tag{7.134}
\]

Here the actual integration runs over the initial physical space \(\Sigma(t')\). However, the integration measure involves not just local quantities at this surface, but also the normal derivatives of the kernel (or the wave function \(\Psi(q')\) itself) arising in the one-loop approximation as a Hamilton-Jacobi argument \(p = -\partial S(q, q')/\partial q'\) of \(J(q, p)\).
Semiclassical physical inner product: Hermiticity and gauge independence

The unitary map from the Dirac quantization to the reduced phase space quantization (7.133) suggests that the simple inner product (7.129) of the former quantization scheme induces a correct physical product in the latter one

$$\langle \Psi' | \Psi \rangle_{\text{phys}} = (\Psi'_{\text{phys}} | \Psi_{\text{phys}})$$

(7.135)
on account of the relation (7.133) and the change of integration variables (7.122). The result for semiclassical states of the form (7.109) looks exactly like (7.81) [204, 205, 206], where the delta-function of gauge conditions determines the \((n-m)\)-dimensional physical subspace \(\Sigma\) embedded in superspace and \(J(q, \partial S/\partial q)\) is a corresponding Faddeev-Popov determinant with the Hamilton-Jacobi value of the momentum\(^{10}\).

Thus, in the one-loop approximation we have a complete agreement between the reduced phase space (ADM) quantization and the BFV induced Dirac quantization method of Sect.7.4. The same can be done for the semiclassical solution of quantum Dirac constraints. While here the two-point prefactor (7.115) was obtained by directly solving the quantum Dirac constraint, the same can be done by calculating in the one-loop approximation the path integral (7.102) in the relativistic gauge of the form \(\hat{N}^\mu - \chi^\mu(q, p) = 0\). This was explicitly done in [226, 227] with the same result (7.115).

It remains to discuss the Hermiticity and gauge dependence properties of relevant observable matrix elements relative to this physical inner product. When it is rewritten in terms of the auxiliary inner product with the delta-function type operator measure

$$\langle \Psi' | \Psi \rangle_{\text{phys}} = \langle \Psi' | \hat{J} \delta(\hat{\chi}) | \Psi \rangle + O(\hbar),$$

(7.136)
the Hermiticity of the observable (7.107) with a real \(O_I(q, p)\) immediately follows from the relation [209] \(\langle \Psi' | [\hat{O}_I, \hat{J} \delta(\hat{\chi})] - i\hbar V^I_\lambda | \Psi \rangle = O(\hbar^2)\). Thus, \(O_I\) is Hermitian with respect to \(\langle \ldots | \ldots \rangle_{\text{phys}},\) even though it has an anti-Hermitian part, \(\hat{O}_I = \hat{O}_I - i\hbar V^I_\lambda + O(\hbar^2)\) with respect to \(\langle \ldots | \ldots \rangle_{\text{phys}}\).

The gauge independence of the physical inner product is based, as shown in [206, 207], on the fact that it can be rewritten as the integral over \((n-m)\)-dimensional surface \(\Sigma\) of a certain \((n-m)\)-form which is closed in virtue of the
7.5. Semiclassical approximation

Dirac constraints on physical states

\[ \langle \Psi' | \Psi \rangle_{\text{phys}} = \int_{\Sigma} \omega^{(n-m)}, \quad d\omega^{(n-m)} = 0, \quad (7.137) \]

\[ \omega^{(n-m)} = \frac{dq^i \wedge \ldots \wedge dq^{i_{n-m}}}{(n-m)!} \epsilon_{i_1 \ldots i_n} \Psi^* \nabla_i \Psi^{i_{n-m+1}} \ldots \nabla_i \Psi. \quad (7.138) \]

It follows then from the Stokes theorem that this integral is independent of the choice of \( \Sigma \) in the class of surfaces the cobordism between which is enforced by some regular \((n-m+1)\)-dimensional subspace of the \( q \)-space.

The closure of \( \omega^{(n-m)} \) is a corollary of the approximate continuity equation for \( \Psi^* \Psi, \quad \partial (\nabla_i \Psi^* \Psi) / \partial q^i = U^\lambda_\mu \Psi^* \Psi + O(\hbar) \), where the correction \( O(\hbar) \) is due to the difference between the momentum arguments related to different Hamilton-Jacobi functions \( p' = \partial S'/\partial q \) and \( p = \partial S/\partial q \), which goes beyond the one-loop approximation (see the footnote 9).

A similar gauge independence property holds for the matrix elements of physical observables, which only differ from (7.137) by the presence of \( O_I(q, \partial S/\partial q) \) in the integrand, and follows from gauge invariance of \( \hat{O}_I \). Thus, as it should have been expected from the theory of gauge fields and BFV formalism the gauge independence of the physical matrix elements or expectation values of observables follows from the gauge invariance of the latter. This is true not only at the formal path-integral quantization level, but also in the operator Dirac quantization scheme.

The exterior form representation (7.137)-(7.138) of the physical inner product gives a correct understanding of the Wheeler-DeWitt conserved current (7.54) – this is a flux of the \((n-m)\)-form through the \((n-m)\)-dimensional surface \( \Sigma \) in \( q \)-space, defined by (7.119). In fact, this surface is a coordinate physical subspace embedded into the coordinate configuration space of \( q \). The codimension of this surface is the range of the index \( \mu \) enumerating the constraints – \( m \), therefore the flux conservation through such surface requires not one, but rather \( m \) continuity type equations. The analogy with the Klein-Gordon type current becomes stronger if we notice that each \( \nabla_i \) factor in the definition (7.138) of \( \omega^{(n-m)} \) can be represented as a result of action of the Wronskian type operator

\[ \frac{1}{2} \left( \left. \frac{\partial H_\mu}{\partial p_i} \right|_{p = \hbar^2 / i \partial q} - \left. \frac{\partial H_\mu}{\partial p_i} \right|_{p = \hbar \delta / i \partial q} \right) + O(\hbar) \quad (7.139) \]

on the wave functions \( \Psi \) and \( \Psi'^* \), with \( \hbar \delta / i \partial q \) acting to the right on the factor \( \Psi \) and \( \hbar \delta / i \partial q \) – to the left on the factor \( \Psi'^* \) [207]. For the Hamiltonian constraint \( H_\mu \) with \( \mu = (\perp, x) \) this is a first order differential Wronskian operator,

\[ \hbar \frac{\delta}{\delta g_{cd}(x)} G^{abcd}(x) - \frac{\delta}{\delta g_{cd}(x)} G^{abcd}(x). \quad (7.140) \]

The product of \( m = 4 \times \infty^3 \) such operator factors in (7.138) is what distinguishes Einstein field theory from the relativistic quantum mechanics of a single Klein-
Quantization of Gravity

Gordon equation corresponding to \( m = 1 \). Unfortunately, this analogy has not yet been extended beyond linear order in \( \hbar \) to get exact conservation law for a flux of many Klein-Gordon type equations subject to a commutator involution of the Dirac constraints type.

In a spatially closed gravitational system with \( H_0 = 0 \) a time evolution arises due to the choice of explicitly time dependent gauge conditions, which according to (7.119) implies the motion of this physical subspace in superspace of 3-metrics \( q \). The one-parameter family of such surfaces \( \Sigma(t) \) induces the nonvanishing values of lapse and shift functions \( N^\mu = -J_{\nu}^{-1}\mu\partial\chi^\nu/\partial t \) (cf. Eq.(3.15)), which according to (2.63) determine the “velocity” of the motion of the 3-dimensional spatial section in spacetime. This is the way how a “frozen” formalism becomes evolutionary due to the imposed time variation of gauge conditions. Such time evolution enters the construction of quantum averages via explicit time dependence of the operator measure in the physical inner product, \( \hat{M}_t = \delta(\chi(q,t))\hat{J}_t + O(\hbar), \hat{J}_t \equiv \det[\chi^\mu(q,t),H^\nu(q,\hat{p})]/i\hbar \). As a result, quantum averages of generic operators become \( t \)-dependent even though the Wheeler-DeWitt wave function \( \Psi(q) \) is constant in time. The nontrivial conservation in time of such averages for matrix elements between different Dirac states is the manifestation of unitarity that was checked above both at the reduced phase space quantization level and in the Dirac-Wheeler-DeWitt formalism.

7.6. Problems and prospects of quantum gravity and cosmology

Thus, the reduced phase space (ADM) quantization and quantum geometrodynamics – Dirac-Wheeler-DeWitt quantization, the latter generated by the truncation of the operator BFV formalism, form equivalent theories. The picture of this equivalence still suffers from a number of inconsistencies and unsolved issues. To begin with, unitary equivalence of these two schemes (7.133) is likely to be restricted to the one-loop approximation for semiclassical states. Apparently the most consistent and universal technique is the operator BFV/BRST quantization. Due to embedding of the theory into extended relativistic phase space of gauge and ghost fields it has the most efficient tools and a lot of gauge fixing flexibility for the analysis of gauge invariance properties, gauge dependence issues, etc. In this regard, the efficiency of the Dirac quantum geometrodynamical approach is much lower, and for the ADM quantization it is nearly absent at all in view of the necessity to solve explicitly complicated nonlinear constraint equations and to deal with spacetime nonlocality.

The problem of the choice of a good gauge-fixing procedure is another fundamental issue in quantum geometrodynamics. Note that, similarly to the Wronskian inner product of the Klein-Gordon equation, the flux of the current form (7.137)-(7.138) is not positive definite. At the classical level this is equivalent to the statement that the determinant factor \( J(q,p) = \det J^\mu_\nu(q,p) \) generally is
7.6. Problems and prospects of quantum gravity and cosmology

not positive definite either. But the degeneration of the Faddeev-Popov functional matrix $J_{\mu}(q, p)$ means the breakdown of the gauge fixing procedure (the analogue of the Gribov copies problem in Yang-Mills gauge theories [228, 229]). The canonical gauges $\chi_{\mu}(q, p)$ that would avoid this problem are, to the best of our knowledge, not known (the discussion of Gribov copies in cosmological context can be found in [230]). This would motivate in analogy with the second quantization of Klein-Gordon equation and secondary quantized string field theory the so called third quantization of gravity targeting once very popular speculations on the physics of multiple baby universes [231].

Another difficulty in quantum geometrodynamics is lacking manifest space-time covariance which is critical for the studies of ultraviolet divergences. The reader might have noticed that the superspace dimensionality and the number of spacetime diffeomorphisms was silently denoted by $n$ and $m$ which in pure 4-dimensional gravity are divergent quantities $n = 6 \times \infty^{3}$ and $m = 4 \times \infty^{3}$, so that majority of constructions underlying the Wheeler-DeWitt equation are purely formal. Commutators of phase space coordinates and momenta in local quantum constraint are divergent $\sim \delta^{(3)}(0)$, traces of structure functions also represented pure divergences $U_{\mu\nu} \sim \delta^{(3)}(0)$, etc. All these expressions require regularization and renormalization, which is hard to render covariant in manifestly noncovariant formalism.

The path integral method which is a central point of the BFV quantization has an efficient means to overcome this difficulty – integration over phase space momenta converts the path integral to the Lagrangian form [213, 46] which in a special class of background covariant gauge conditions allows one to enjoy all the advantages of the manifestly covariant treatment. A covariant regularization allows one to minimize the violation of classical symmetries of the theory at the quantum level. In this respect, the path integral representation is what one needs – calculating this integral by the technique of covariantly regularized Feynman graphs is a right alternative to the attempts of directly solving the Wheeler-DeWitt equation. This really denies disparaging remarks on this equation, mentioned in Preface, as an absolutely useless tool in quantum gravity and explains its first principle founding nature. Like Schrödinger equation in quantum field theory, it lies at the foundation of the theory, though in concrete applications it gives way to a more efficient but a derivative tool – path integral method.

The realization of this strategy for the two point kernel (7.102) gives it as a projector onto the space of solutions of Wheeler-DeWitt equations, acting in superspace of 3-metrics and matter fields $q^i = (g_{ab}(x), \phi(x))$.

$$U[\tilde{g}_{ab}^+(x), \varphi^+(x) | \tilde{g}_{ab}^-(x), \varphi^-(x)] = \int D[g_{\mu\nu}, \phi] e^{iS[g_{\mu\nu}, \phi]}. \quad (7.141)$$

Here $S[g_{\mu\nu}, \phi]$ is the covariant action of the gravitational and matter fields, the details of Faddeev-Popov’s gauge fixing procedure are hidden in the measure $D[g_{\mu\nu}, \phi]$ and the integration runs over gravitational and matter spacetime histories $q^i(t), N^\nu(t) = (g_{\mu\nu}(x), \phi(x))$, $x = (t, x)$, interpolating between config-
urations at initial and final spatial hypersurfaces (labeled by $t_{\pm}$),

\[
g_{ab}(x, t_{\pm}) = g_{ab}^{\pm}(x), \quad \phi(x, t_{\pm}) = \phi^{\pm}(x).
\] (7.142)

Depending on the choice of these configurations and possible analytic continuation of the time integration contour between $t_{-}$ and $t_{+}$ into the complex plane (Euclidean time or “Euclidean quantum gravity”) this path integral generates various prescriptions for a cosmological quantum state. These prescriptions include, in particular, the Hartle-Hawking no-boundary [232, 233] and the tunneling [234, 235, 236, 237] quantum states of the Universe, which are usually associated with creation of the Universe from “Nothing”. The last two decades, with the advent of precision cosmology, are characterized by a growing interest in cosmological inflation theory that is widely recognized as underlying the observable large scale structure of the Universe. This explains the interest in these two quantum states as they, with this or that level of success, describe the origin of inflationary Universe.

Even more appealing sounds a recently suggested construction of the initial cosmological state as a microcanonical density matrix [238, 239]. This density matrix is proposed as a projector on all possible solutions of the Wheeler-DeWitt equation (7.141). The motivation for such a proposal is the principle of Occam razor – minimum set of assumptions, because this is an ultimate equipartition in the full set of states of the theory — “Sum over Everything” [239]. Creation of the Universe from “Everything” is conceptually more appealing than creation from “Nothing”, because the democracy of the equipartition better fits the principle of Occam’s razor than the selection of a concrete state.

The statistical sum of this microcanonical state is the trace of $\hat{U}$ over the physical configuration space, which has a representation of the covariant path integral over periodic in time histories of gravitational and matter fields [210]

\[
Z = \text{tr}_{\text{phys}} \hat{U} \equiv \int dq \hat{M} U(q, q') \big|_{q' = q} = \int_{\text{periodic}} D[g_{\mu\nu}, \phi] e^{i S[g_{\mu\nu}, \phi]}.
\] (7.143)

Semiclassical calculation of this integral by the saddle point method runs in a manifestly covariant formalism which admits covariant UV renormalization underlying nontrivial quantum scaling behavior which in its turn is generated by a conformal anomaly of matter fields [238, 239].

This leads to numerous new physical effects in the theory of very early quantum Universe, which are impossible within the no-boundary and tunneling prescriptions. These effects culminate in the prediction of a new type of inflationary scenario with a bounded sub-Planckian energy scale [238, 241], which incorporates models of the Starobinsky $R^2$-inflation and Higgs inflation [240]. These models give, as is widely recognized, the best explanation of the observable cosmic microwave background (CMB) data and even establish intriguing relation between the basic parameters of this data and the mass value of the recently discovered Higgs boson [242, 243]. There is a lot more in store for us, which is ultimately based on the Dirac-Wheeler-DeWitt quantum geometrodynamics!
A1. Geometry of manifolds

This appendix contains a brief summary of definitions of the basic concepts of the differential geometry of manifolds used in the main text. The details and proofs of facts collected here can be found in [2, 3, 4, 24].

**Manifold**

Let $M$ be a topological Hausdorff space with a countable base (the Hausdorff property means the separability of points in the space $M$ in the following sense: for any $p, q \in M$, there are open neighbourhoods $U_p \ni p$ and $V_q \ni q$, such that $U_p \cap V_q = \emptyset$). The space $M$ is called an $n$-dimensional smooth manifold, if it is locally homeomorphic to $\mathbb{R}^n$, i.e., 1) for each open set $U_i$ (the collection of which covers the whole space $\bigcup U_i = M$), a homeomorphism $\varphi_i : U_i \rightarrow \mathbb{R}^n$ to a region of $\mathbb{R}^n$ is defined, which means the introduction of the local coordinates $x^1(p), \cdots, x^n(p)$ for the points $p \in U_i$; 2) in the intersections $U_i \cap U_j \neq \emptyset$, the composition of maps $\varphi_j \circ \varphi_i^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is defined which is a smooth function of its arguments. For simplicity, the smoothness is understood everywhere as an infinite differentiability. The pair $(U_i, \varphi_i)$ is called a chart, and the complete set of consistent charts (in the sense of the smoothness of transition functions in the intersections) is called an atlas of the manifold $M$.

The local properties of manifolds are clarified when studying the tangent spaces. One can define the tangent vector $X_p$ at the point $p \in M$ as the linear differential operator acting on the set of smooth functions defined on the
manifold \( M \) in the neighbourhood of the point \( p \), that satisfies the condition \( X(fg) = X(f)g(p) + f(p)X(g) \), \( \forall f, g \). The set \( T_pM \) of all tangent vectors at the point \( p \in M \) constitutes an \( n \)-dimensional linear vector space and is called a tangent space to \( M \) at a point \( p \). The operators \( (\partial_\mu)_p = (\partial/\partial x^\mu)_p \) [where \( x^\mu, \mu = 1, \ldots, n, \) are the coordinates in the local chart] and \( (\partial_\mu)_p f = (\partial_\mu f)_p \) constitute the basis in \( T_pM \), which is called the coordinate basis. Any vector \( X_p \) is represented by the linear combination \( X = X^\mu(\partial/\partial x^\mu) \).

Let \( M_1 \) and \( M_2 \) be the smooth manifolds. The mapping \( f : M_1 \to M_2 \) is called a differentiable mapping, if in any local charts \((U_i, \varphi_i) \) in \( M_1 \) and \((V_j, \psi_j) \) in \( M_2 \) the composition \( \psi_j \circ f \circ \varphi_i^{-1} \) is a differentiable function. The differentiable mapping \( f \) determines the mapping of the respective tangent spaces: a tangent vector \( X_p \) at the point \( p \in M_1 \) is mapped to the vector \( Y_p = f_*(X_p) \) at the point \( q = f(p) \in M_2 \) in accordance with the rule \( f_*(X)(g) = X(g \circ f) \) for any smooth function \( g : M_2 \to R \). The map \( f_* \) is called a differential of \( f \); this map is also often denoted as \( df \).

We say that the smooth vector field \( X \) is defined on \( M \), if each point \( p \in M \) is mapped to the vector \( X_p \in T_pM \) and for any smooth function \( f \) on \( M \), \( X_p(f) = (Xf)(p) \) is a smooth function.

The space of linear forms \( T^*_pM \) dual to \( T_pM \) is called a cotangent space at a point \( p \in M \). In the local coordinates \( \{x^\mu\} \) in the neighbourhood of \( p \), the basis in \( T^*_pM \) can be defined by the set of differentials \( \{dx^\mu, dx^2, \ldots, dx^n\} \), the latter basis is dual to the coordinate frame basis \( \{\partial_\mu\} \).

Using \( T_pM \) and \( T^*_pM \), one can consider arbitrary rank tensors as the elements of the tensor products \( T_pM \otimes \cdots \otimes T_pM \otimes T^*_pM \otimes \cdots \otimes T^*_pM \).

In the local chart \( \{x^\mu\} \), any smooth tensor field \( T \) covariant of the rank \( k \) and contravariant of the rank \( l \) is defined by its components \( T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_l} \) – the smooth functions of coordinates

\[
T = T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_l} \partial_{\mu_1} \otimes \cdots \otimes \partial_{\mu_k} \otimes dx^{\nu_1} \otimes \cdots \otimes dx^{\nu_l}.
\]

When changing the coordinates \( x^\mu \to y^\mu(x) \), the tensor components are transformed as

\[
T^{\mu_1 \cdots \mu_k}_{\nu_1 \cdots \nu_l} = T^{\alpha_1 \cdots \alpha_k}_{\beta_1 \cdots \beta_l} \frac{\partial y^{\mu_1}}{\partial x^{\alpha_1}} \cdots \frac{\partial y^{\mu_k}}{\partial x^{\alpha_k}} \frac{\partial x^{\beta_1}}{\partial y^{\nu_1}} \cdots \frac{\partial x^{\beta_l}}{\partial y^{\nu_l}}
\]

in accordance with the transformations of bases of the tangent and cotangent spaces

\[
\frac{\partial}{\partial y^\mu} = \frac{\partial x^\nu}{\partial y^\mu} \frac{\partial}{\partial x^\nu}, \quad dy^\mu = \frac{\partial y^\mu}{\partial x^\nu} dx^\nu.
\]

**Differential forms**

The elements of the cotangent space \( \omega \in T^*_pM \), defined at each point \( p \in M \), are called a smooth differential 1-form, if in the local coordinates \( \{x^\mu\} \) the decomposition coefficients of the form \( \omega \) with respect to the basis of \( T^*_pM \), \( \omega = \omega_\mu(x) dx^\mu \) are differentiable functions.
The exterior product of spaces $T_p^*M$ is defined as a totally antisymmetrized tensor product
\[ T_p^*M \wedge \cdots \wedge T_p^*M = \text{altern}(T_p^*M \otimes \cdots \otimes T_p^*M). \]

The elements of space $\Lambda^k(T_p^*M)$ – of $k$-th exterior power $T_p^*M$ – are the completely antisymmetric covariant tensors
\[ \omega = \frac{1}{k!} \omega_{\mu_1 \ldots \mu_k}(x) \, dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_k}, \quad \omega_{\mu_1 \ldots \mu_k} = \omega[\mu_1 \ldots \mu_k]. \]

If at each point $p \in M$, the element from $\Lambda^k(T_p^*M)$ is defined such that the coefficients $\omega_{\mu_1 \ldots \mu_k}(x)$ are smooth functions, we say that an exterior differential form of rank $k$ is defined on the manifold $M$. The space of all $k$-forms on $M$ is denoted by $\Lambda^k$ (it is clear that $\Lambda^k = \emptyset$, if $k > n = \dim M$). The direct sum $\Lambda^* = \Lambda^0 \oplus \Lambda^1 \oplus \cdots \Lambda^n$ is a linear space (with an obvious definition of the sum and multiplication) and the introduction of the operation of exterior product in it converts $\Lambda^*$ into an algebra. In the local coordinates $\{x^\mu\}$, the exterior product of a $k$-form $\omega = \frac{1}{k!} \omega_{\mu_1 \ldots \mu_k}(x) \, dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_k}$ and an $l$-form $\varphi = \frac{1}{l!} \varphi_{\mu_1 \ldots \mu_l} \, dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_l}$ is defined as a $(k+l)$-form $\sigma = \omega \wedge \varphi$,
\[ \sigma = \frac{1}{(k+l)!} \sigma_{\mu_1 \ldots \mu_{k+l}}(dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_{k+l}}) \]
where
\[ \sigma_{\mu_1 \ldots \mu_{k+l}} = \frac{(k+l)!}{k!!} \omega_{[\mu_1 \ldots \mu_k} \varphi_{\mu_{k+1} \ldots \mu_{k+l}].} \]

The exterior product is non-commutative:
\[ (k) \omega \wedge (k) \varphi = (-1)^{kl} \left( (k) \varphi \wedge (k) \omega \right). \]

For a differentiable mapping $f : M_1 \to M_2$ along with its differential $f_* : T M_1 \to T M_2$, the dual mapping of forms (which called a pull-back map) is always determined: the form $\varphi$ on $M_2$ is mapped to the form $\omega = f^* \varphi$ on $M_1$, so that $f^* \varphi(X_1, \ldots, X_k) = \varphi(f_* X_1, \ldots, f_* X_k)$, where $X_1, \ldots, X_k \in T_p M_1$.

The operation of exterior differentiation is most important notion of the exterior algebra. The mapping $d : \Lambda^k \to \Lambda^{k+1}$ is called the exterior differential, if it has the following properties: the Leibniz rule $d(\omega^l \wedge \varphi) = d\omega \wedge \varphi + (-1)^l \omega \wedge d\varphi$ with $\omega \in \Lambda^l$: for 0-forms (functions on $M$) $d$ coincides with the differential $df$; and $d$ is nilpotent, $d^2 = 0$. In the local coordinates
\[ d \omega = \frac{1}{k!} \partial_{[\mu_1} \omega_{\mu_2 \ldots \mu_{k+1}]}(x) \, dx^{\mu_1} \wedge dx^{\mu_2} \wedge \cdots \wedge dx^{\mu_{k+1}}. \]
If $M$ is a Riemannian space with the metric $g_{\mu\nu}(x)$ of the signature $p$, then the operation of dualization is defined on the exterior algebra, or the Hodge ("star") operator $\ast : \Lambda^k \rightarrow \Lambda^{n-k}$. In the local coordinates, this mapping has the following form:

$$\ast^k \omega = \frac{1}{(n-k)!k!} \varepsilon_{\mu_1...\mu_{n-k}\nu_1...\nu_k} \omega^{\nu_1...\nu_k} dx^{\mu_1} \wedge ... \wedge dx^{\mu_{n-k}},$$

where $\varepsilon_{\mu_1...\mu_n}$ is totally antisymmetric Levi-Civita tensor density. From the definition, we have a property

$$**^k \omega = (-1)^{k(n-k)+p} \ast^k \omega.$$

The theory of integration of exterior forms on manifolds can be studied in [3, 4]. Here, we will confine ourselves to the local definition only. Every smooth $k$-form $\omega$ on the manifold $M^n$ defines on any $k$-dimensional hypersurface $N_k \subset M^n$ a $k$-form $\gamma^* \omega$, where $\gamma : N_k \rightarrow M^n$ is the relevant embedding map. If the local coordinates $(y^1...y^k)$ are chosen in a region $U \subset N_k$, the integral of the form $\omega$ over $U$ is defined as the number $\int_U (\gamma^* \omega)_{12...k} dy^1...dy^k$. This number is denoted as $\int_U \omega$. The invariant integral over the entire $k$-hypersurface $\int_{N_k} \omega$ is then determined using the natural matching of the integrals over the charts covering $N_k$.

**Bundles**

The notion of a bundle space is very important in physics, because it formalizes the idea of an “internal space” of states in which a physical system can be in each point of the spacetime.

Let $V_m$ be an $m$-dimensional vector space, and the action of some group $G : G \times V^m \rightarrow V^m$ is defined on it. Then, the trivial vector bundle over $M$ is called a direct product $V^m \times M = E$, where the action of the group $G : G \times E \rightarrow E$ is naturally defined: $u = (v, x) \in E$ is mapped into $gu = (gv, x)$. The manifold $M$ is called a base of bundle $E$, $V^m$ is a typical fiber of a bundle, and $G$ is a structural group. The generalization of this structure is a locally trivial bundle $E(M, V, G, \pi)$. By definition, this is an $(m+n)$-dimensional smooth manifold $E$ where the projection $\pi : E \rightarrow M$ is defined, and in addition there is a smooth atlas, $\bigcup_i U_i = M$, such that for any point $x \in M$ there is a neighbourhood $U_i \ni x$, the preimage of which $\pi^{-1}(U_i)$ is diffeomorphic to $V \times U_i$ (the local triviality), and the corresponding diffeomorphisms $\varphi_i : \pi^{-1}(U_i) \rightarrow V \times U_i$ are identical, i.e., for $(v, x) \in V \times U_i$ one has $\pi \circ \varphi_i^{-1}(v, x) = x$. In addition, in the intersection regions $U_i \cap U_j \neq \emptyset$, the diffeomorphisms are compatible in the sense that $\varphi_i^{-1} \circ \varphi_j = g \in G$ determines an automorphism of the fiber $V$ under the action of the structural group $G$. If the smooth manifold $V$ is a space of a linear representation of the group $G$, i.e., it is an $m$-dimensional vector space, the bundle $E$ is called vector bundle.
An example of the vector bundle is the $2n$-dimensional manifold defined as
\[ \bigcup_{x \in M} T_x M = TM \] which is called a tangent bundle.

When the fiber coincides with the structural group $V = G$ acting on itself by right shifts, the bundle is called a principal one. It is denoted by $P(M, G)$, where $P$ is a bundle space, $M$ is its base. In a certain sense, the notion of a principal bundle generalizes the notion of the Lie group, since formally it is nothing but the collection of the copies of the group $G$ “numbered” by the points from $M$. Its properties are described in detail in [3]. Here we mention just one important mathematical fact. It is known that the Lie algebra $G$ generates an action on the Lie group $G$: namely, every element $a \in G$ generates a one-parametric subgroup in $G$ by means of the exponential map $\exp a$. As a result, in the principal bundle $P(M, G)$ the right action of $G$ on $P$ induces a natural homomorphism of the Lie algebra $G$ into the Lie algebra of the vector fields on $P$: for $a \in G$ the corresponding induced vector field $a^*$ on $P$ is called a fundamental vector field.

By construction, $a^*$ is tangent to the fiber, i.e., it is a vertical vector field, $\pi^* (a^*) = 0$.

Let the action of the group $G$ be defined on a smooth manifold $V$. Then, for the principal bundle $P(M, G)$ one can introduce an associated bundle with $V$ as the typical fiber: $E(M, V, G, P) = (P \times V)/G$, where the factorization is considered with respect to the equivalence relation defined on $P \times V$ by the natural right action $G$, namely, $(p, v) g = (R_g p, v g^{-1})$, $\forall p \in P, v \in V, g \in G$.

The example of the principal bundle is the bundle of linear frames $L(M)$ with a structural group $GL(4, R)$ (see Sec. 5.2.), and the tangent bundle can be considered as an associated one with it.

The smooth mapping $\sigma : M \to E(M, V, G)$ such that $\pi \circ \sigma = \text{id}_M$ is called the cross-section of the bundle $E$. The cross-section of the principal bundle exists only if $P = M \times G$ is the trivial bundle [3, 4].

Let $H$ be a subgroup of the group $G$ and $\tilde{P}(M, H)$ is a principal bundle over $M$ with a structural group $H$. The mapping $f : \tilde{P}(M, H) \to P(M, G)$ is called the reduction of the structural group $G$ to the subgroup $H$, if the induced mapping $f : M \to M$ is identical, and $f : H \to G$ is a monomorphism. Then the subbundle $\tilde{P}(M, H)$ is called the reduced bundle. One can prove that the necessary and sufficient condition for the reduction of the structural group $G$ of the principal bundle $P(M, G)$ to a closed subgroup $H$ is the existence of a cross-section of the bundle $E(M, G/H, G, P)$ associated with $P$ with the typical fiber $G/H$, [3].

Connection

The gauge field in the geometrical approach in the field theory is identified with a connection in the bundle space. Let $P(M, G)$ be a principal bundle. The connection in $P$ is a smooth mapping $\rho_p$ of the tangent space to the base into the tangent space to the bundle, $\rho_p : T_x M \to T_p P$ (with $\pi(p) = x$) such that the horizontal subspace $H_p = \rho_p(T_x M)$ is invariant with respect to the action
of the group $G$ on $P$ and $\pi_* \circ \rho_p = id_{\pi_*^*\mathcal{M}}$. The smoothness is understood as a smooth dependence of $\rho_p$ on $p \in P$. Thus introduced connection defines the unique partition of the tangent space $T_pP = \mathcal{H}_p \oplus V_p$ into the direct sum of the horizontal $\mathcal{H}_p$ and the vertical $V_p$ subspaces. The latter subspace $V_p$ is formed by the tangent vectors to the fiber, and therefore, it is isomorphic to the Lie algebra of the group $G$. Observing this, one can give an equivalent definition of a connection. We say that the connection is defined in the bundle $P(M, G)$ if a 1-form of $\omega$ with the values in the Lie algebra of the structural group $G$ is given on $P$, such that $\omega(a^*) = a$ and $R_g^\ast \omega = g^{-1} \omega g$ for all $a \in G$ and $g \in G$. One can show that definition of the connection 1-form $\omega$ in $P$ is equivalent to the existence of a set of 1-forms $\omega_i$ on the base space $M$, which are determined for any $i$ in the corresponding chart $U_i$ (where $\{U_i\}$ is an open covering of $M$) by the local cross-sections $\sigma_i : U_i \to P$, so that $\omega_i = \sigma_i^\ast \omega$. Other properties of the connection are discussed in detail in [3, 4].

The gauge field strength coincides in the local coordinates with the components of the curvature of connection in the principal bundle. The form of the curvature $R$ of a connection $\omega$ on $P$ is 2-form with the values in the Lie algebra of the structural group $G$, which is determined as exterior covariant differential from $\omega$:

$$R(X, Y) = D\omega(X, Y) = d\omega(hX, hY), \quad X, Y \in T_pP,$$

where $h$ is a projection to the horizontal subspace $\mathcal{H}_p$. The curvature satisfies the structure equations [3, 4]

$$d\omega(X, Y) = -\frac{1}{2} [\omega(X), \omega(Y)] + R(X, Y), \quad X, Y \in T_pP.$$

Here the square bracket $[\ , \ ]$ denotes the Lie algebra commutator.

The definition of a connection $\omega$ determines the parallel transport of the vectors in the bundle associated with $P$, as well as introduces the covariant derivative in $E(M, V, G, P)$, [3].
A2. Spinor analysis on an arbitrary manifold

Let us consider the spinor bundle, the typical fiber of which is a two-dimensional complex vector space \( S^2 \), the structural group is \( GL(2, \mathbb{C}) \), and the base is a spacetime \( M_4 \). It is more convenient to work in the local chart, where the main object of our investigation – the contravariant spinor of the 1st rank is a two-component complex field

\[
\xi^A = \begin{pmatrix} \xi^1(x) \\ \xi^2(x) \end{pmatrix}.
\]

The spinor indices run the two values, \( A, B, C, \ldots = 1, 2 \). When changing the local basis, the spinor is transformed as

\[
\xi'^A = L^A_A \xi^A, \quad L^A_A \in GL(2, \mathbb{C}).
\]

As usual, the bilinear non-singular interior product \( (S^2 \times S^2 \to \mathbb{C}) \) is introduced in the space \( S^2 \), which in local basis is set by antisymmetric matrix \([81, 82]\)

\[
\varepsilon_{AB} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.
\]

The elements of the dual spinor space – the covariant spinors (\( \xi^A \) in local basis) are transformed under the action of \( L^{-1} \in GL(2, \mathbb{C}) \). Raising and lowering of the indices can be performed with the help of \( \varepsilon_{AB} \) and \( \varepsilon^{AB} \), where \( \varepsilon^{BA} \varepsilon_{BC} = \delta^A_C \).

In addition, we define the bundle with typical fiber \( S_2 \) over \( M_4 \), which is called the conjugate spinor space. The elements of the fiber \( (S_2)_x \) over the point \( x \) in the local chart are the complex-conjugate quantities to spinors from \( S^2 \). They are denoted by \( \Psi^\dagger_A = (\Psi_A)^\ast \), and they are transformed under the action of the complex-conjugate matrix \( L^A_B = (L^A_B)^\ast \). In \( S_2 \), the bilinear non-singular antisymmetric interior product is also introduced in local basis by the matrix \( \varepsilon^\dagger_{AB} = (\varepsilon_{AB})^\ast \).

Spinors of the higher ranks are defined as the elements of the tensor products \( S^2 \otimes \cdots \otimes S^2 \otimes \cdots S^2 \otimes \cdots S^2 \) and are transformed in an obvious way as the products of the 2-spinors of the first rank.

The linear connection in the spinor bundle is defined by the corresponding 1-form with the values in the Lie algebra \( gl(2, \mathbb{C}) \). For a local cross-section – the spinor field \( \xi(x) \) – this introduces the coefficients of the local connection \( \omega_\mu(x) \) which under the change of the local basis are transformed according to \( \omega_\mu \to \omega'_\mu = L \omega_\mu L^{-1} + L_\partial_\mu L^{-1} \). Accordingly, the covariant derivative of spinor fields is defined by

\[
D_\mu \xi^A_B = \partial_\mu \xi^A_B + \omega^A_{\ C\mu} \xi^C_B - \omega^D_{ B\mu} \xi^A_D.
\]

Such a general \( gl(2, \mathbb{C}) \)-connection in the spinor bundle is not consistent with the interior spinor product in the sense that

\[
D_\mu \varepsilon_{AC} = 2\omega_{[AC]\mu} = \varphi_\mu \varepsilon_{AC},
\]
\[ \varphi_\mu = \omega^A \mu \] is the trace of the coefficients of the spinor connection.

Until now, the properties of the spinor bundle were considered separately from the geometry of the base space, since its own structure (the interior product, connection, etc.) can be defined quite independently. However, developing a consistent gauge approach to the gravitational field, we now will establish a relation of the spinor structure with geometrical spacetime structures. This issue is the most important one in the construction of the spinor analysis on an arbitrary manifold.

Let us make use of the fact that there exists a fundamental isomorphism of the space \((S_2 \otimes \bar{S}_2)_x\) (the index \(H\) means the Hermitization) at a point \(x \in M_4\) and the tangent space \(T_x M_4\). For the typical fibers \(S_2 \otimes \bar{S}_2\) and \(R^4\), this isomorphism determines an arbitrary 4-vector as a linear combination of pairs \(\xi^A \mu^B\).

In the local chart, the aforementioned isomorphism is established with the help of the fundamental spin-tensor objects (the generalized Pauli matrices)

\[ g^{a \dot{A} B} = - (g^{a \dot{A} B})^\dagger \] (the Latin indices \(a, b, c, \cdots = 0, 1, 2, 3\) refer to an arbitrary Lorentz frame in \(T_x M\)). The matrices \(g^{a \dot{A} B}\) satisfy the postulate

\[ g^{a \dot{A} B} g^{b \dot{C} B} = \eta^{ab} \delta^A_C + \frac{i}{2} \epsilon^{abcd} g_c^{\dot{A} B} g_d^{\dot{C} B}, \] (A2.1)

and are arbitrary in other respects. The generalized Pauli matrices can be chosen at one’s convenience. We will use a particular representation:

\[ g^{0 \dot{A} \dot{B}} = -i \mathbf{1}, \quad g^{k \dot{A} \dot{B}} = -i \sigma^k, \quad k = 1, 2, 3, \]

where \(\mathbf{1}\) is the \(2 \times 2\) unit matrix, and \(\sigma^k\) are the standard \(2 \times 2\) Pauli matrices.

The important consequence of (A2.1) is the relation

\[ g^{(a \dot{A} B)} g^{b \dot{C} B} = \eta^{ab} \delta^A_C, \] (A2.2)

which provides a link between the world metric and the spinor interior product. As usual, their consistency is achieved with the help of the translational gauge gravitational field – the tetrads \(h^a_\mu\), which introduce the world Pauli matrices \(g^{\mu \dot{A} \dot{B}} = h^a_\mu g^{a \dot{A} B}\), and hence obviously modifies (A2.1), (A2.2).

One can make the connections \((\Gamma^a_{\mu \nu} \text{ and } \omega^{A}_{\mu \nu})\) compatible by imposing a natural requirement of coincidence of the covariant derivative of a 4-vector with respect to the connection \(\Gamma\) on \(M_4\) and the covariant derivative with respect to \(\omega\) of its spinor image under the canonical isomorphism. Then we postulate

\[ \nabla_\mu g^{\nu \dot{A} \dot{B}} = \frac{1}{2} K^{\lambda \nu}_{\mu \mu} g^{\lambda \dot{A} \dot{B}}, \] (A2.3)

which is consistent with (A2.2).

The equation (A2.3) is easily solved with respect to the spinor connection by multiplying it by \(g^{\nu \dot{C} B}\) and using the property (A2.1):

\[ \omega^A_{\dot{B} \mu} = \frac{1}{4} \Gamma_{[ab] \mu} (-)^{ab} S_{\dot{A} \dot{B}}, \] (A2.4)
where \((-\rangle S^{ab}_{AB} := g^{[a|AC|} g^{b]_{B\hat{C}}}\). The spin-tensor is symmetric \((-\rangle S^{ab}_{AB} = \langle -S^{ab}_{BA}\), and anti-self-dual, \(*S^{ab}_{AB} = -i S^{ab}_{BA}\). In a similar formula for \(\omega^{\hat{A}}_{\hat{B}\mu}\), instead of \((-\rangle S\) one has the self-dual spin-tensor \((+\rangle S^{ab}_{\hat{A}\hat{B}} := g^{[a|C\hat{A}|} g^{b]_{\hat{C}\hat{B}}}\).

The connection (A2.4) is determined up to a term \(i \Delta^\mu_{\delta A\delta B}\), where \(\Delta^\mu\) is an arbitrary real vector, usually related to electromagnetic field, which was assumed to vanish.

Computing the commutator of the covariant spinor derivatives, one can obtain the expression for the spinor curvature

\[
R^A_{\ B\mu\nu} = \partial_\mu \omega^A_{\ B\nu} - \partial_\nu \omega^A_{\ B\mu} + \omega^A_{\ C\mu} \omega^C_{\ B\nu} - \omega^A_{\ C\nu} \omega^C_{\ B\mu}
\]

in the form

\[
R^A_{\ B\mu\nu} = \frac{1}{4} R_{ab\mu\nu}(\Gamma) \langle -S^{ab}_{\ A\hat{B}}, \rangle
\]

and similarly for \(R^A_{\ B\mu\nu}\).

The equation (A2.4) shows that the Lorentz (the Riemann-Cartan) part of the connection \(\Gamma_{[ab]\mu}\) interacts with the spinors in a conventional minimal coupling manner, while the non-metricity is represented only by the Weyl vector \(K^\mu\). This agrees with the fact that the structural group of the spinor bundle \(GL(2,\mathbb{C})\) is isomorphic to the Weyl group (which is the direct product of the Poincaré group \(P_{10}\) by the 1-dimensional Abelian group of dilations \(R^1\)), multiplied by the group of the phase transformations.

The transition from the 2-spinors \(\xi^A\) to the bispinors (or the 4-spinors of Dirac) is straightforward:

\[
\Psi = \begin{pmatrix} \xi^A \\ \eta_{\hat{A}} \end{pmatrix}.
\]

The 4 \(\times\) 4 Dirac matrices then read

\[
\gamma^\mu = \begin{pmatrix} 0 & g^{\mu A\hat{B}} \\ g_{\mu A\hat{B}} & 0 \end{pmatrix}.
\]

With our choice of the generalized Pauli matrices, these Dirac matrices reproduce Weinberg’s [192] representation. The Dirac matrices satisfy

\[
\gamma^\mu \gamma^\nu = g^{\mu\nu} I + \frac{i}{2} \varepsilon^{\mu\nu\alpha\beta} \gamma^\alpha \gamma^\beta \gamma^5,
\]

where \(\gamma_5 = \frac{1}{4!} \varepsilon_{\alpha\beta\mu\nu} \gamma^\alpha \gamma^\beta \gamma^\mu \gamma^\nu = -i \gamma^0 \gamma^1 \gamma^2 \gamma^3\), and \(I\) is a 4 \(\times\) 4 unit matrix. From these fundamental relations, one straightforwardly derives

\[
\gamma^{[\alpha \gamma^\beta] \gamma^\mu} = 2 \left( \gamma^\alpha g^{\beta\mu} - \gamma^\beta g^{\alpha\mu} \right),
\]

\[
\gamma^{[\alpha \gamma^\beta] \gamma^\mu} + \gamma^\mu \gamma^{[\alpha \gamma^\beta]} = 2 i \varepsilon^{\beta\alpha\mu\nu} \gamma^\beta \gamma^5.
\]
Finally, we define the covariant derivative of 4-spinors: $D_\mu \Psi = \partial_\mu \Psi + \omega_\mu \Psi$. Similarly to (A2.3), we impose the compatibility condition

$$D_\mu \gamma^\nu = \partial_\mu \gamma^\nu + \Gamma^\nu_{\lambda\mu} \gamma^\lambda + \omega_\mu \gamma^\nu - \gamma^\nu \omega_\mu = -\frac{1}{2} K^{\nu}_{\mu\lambda} \gamma^\lambda,$$

which is consistent with the symmetrized relation $\gamma^{(\mu\nu)} = g^{\mu\nu} I$, see (A2.5). Note that $\nabla_\lambda g^{\mu\nu} = -g^{\mu\alpha} g^{\nu\beta} K_{\lambda\alpha\beta} = -K^{\mu\nu}_{\lambda}$. The solution of the compatibility condition is straightforward and represents a 4-spinor generalization of the connection (A2.4)

$$\omega_\mu = \frac{1}{4} \Gamma_{[ab]\mu} \gamma^{[a\lambda]} \gamma^{\lambda}.$$

Hence, the covariant derivative of 4-spinor reads

$$D_\mu \Psi = \partial_\mu \Psi + \omega_\mu \Psi = \partial_\mu \Psi + \frac{1}{4} \Gamma_{[ab]\mu} \gamma^{[a\lambda]} \gamma^{\lambda} \Psi. \quad \text{(A2.7)}$$

The covariant derivative of the Dirac conjugated spinor $\overline{\Psi} = \Psi^{\dagger} \beta$ has the following form:

$$D_\mu \overline{\Psi} = \partial_\mu \overline{\Psi} - \overline{\Psi} \omega_\mu = \partial_\mu \overline{\Psi} - \frac{1}{4} \Gamma_{[ab]\mu} \overline{\Psi} \gamma^{[a\lambda]} \gamma^{\lambda}. \quad \text{(A2.8)}$$

It is worthwhile to note that for the metric with the signature $(-, +, +, +)$ the Dirac conjugation is defined by the matrix $\beta = i \gamma^0$, and the Dirac matrices are anti-Hermitian in the sense [192]

$$\beta \gamma^a \beta = -\gamma^a.$$

In Weinberg’s representation we explicitly have

$$\beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and hence $\beta^2 = I$. 


A3. Basic notations

We summarize here the basic notations and conventions used in the book.

Indices

Greek indices $\alpha, \beta, \cdots = 0, 1, 2, 3$ refer to the local world coordinates on the spacetime manifold. Latin indices $a, b, c, \cdots = 0, 1, 2, 3$ label the local Lorentz (tetrad) components, except for the Chapters 2, 3, and 7, where the indices $a, b, c, \cdots = 1, 2, 3$ refer to the 3-dimensional (spatial) coordinates. Separate local Lorentz components are denoted with hats over the index, to distinguish them from the world components. E.g., $v^0$ is the 0-th component of a vector with respect to the orthonormal tetrad, in contrast to the 0-th component $v_0$ in the coordinate basis. The upper case Latin indices $A, B, C, \cdots = 1, 2$ label the spinor components.

Metric

The spacetime metric in 4 dimensions has the signature $(-, +, +, +)$, unless the other signature is assumed. Its determinant is $g \equiv |^4g| = - \det g_{\mu\nu}$. In Chapters 2, 3, and 7, the determinant of the 3-dimensional (spatial) metric is denoted as $g^{1/2} = (\det g_{ab})^{1/2}$. The totally antisymmetric Levi-Civita tensor density is $\varepsilon_{\mu\nu\alpha\beta} = \sqrt{|^4g|} \eta_{\mu\nu\alpha\beta}$, where the completely antisymmetric symbol has the only nontrivial component $\eta_{0123} = +1$. The scalar product for the spinors is defined by $\varepsilon_{AB} = - \varepsilon_{BA}$.

Derivatives

The partial derivatives with respect to the spacetime coordinates are denoted $\frac{\partial}{\partial x^\mu} = \partial_\mu = \partial_\mu$. The Riemannian covariant derivatives (defined by the Christoffel symbols) are denoted as $\nabla_\mu$ or $\{\}$ $\nabla_\mu$. The general covariant derivatives with respect to an arbitrary linear connection $\Gamma$ are denoted as $\nabla, \nabla^{\Gamma}, D^{\Gamma}$. For the non-Riemannian objects, constructed from $\Gamma$, the functional dependence is always explicitly specified, for example $R(\Gamma)$. 
A4. Comments on the literature and general remarks

Chapter 1

The equations of the general relativity theory of gravitation (GR) in the Riemannian spacetime were postulated by A. Einstein [22] in 1915, and simultaneously they were derived by D. Hilbert [23] on the basis of the principle of least action. In 1922, É. Cartan [20] introduced the torsion into the differential geometry. His studies of the manifolds with torsion were motivated by the interest in GR. Apparently, being under the influence of the results of the Cosserats brothers [25] on a continuous medium as a manifold, Cartan proposed [19] to couple the torsion tensor to an internal angular momentum of the matter. Later, the ideas of Cosserats-Cartan formed the basis of the continuous theory of dislocations [29].

In the gravity theory, these notions received a substantiation only after the gauge approach was formulated (H. Weyl [5], C.N. Yang and R. Mills [30], R. Utiyama [31]). The physical ideas that underlie the gauge theory about the existence of the fundamental matter fields and the fields mediating interactions became the basis of R. Utiyama’s attempt to consider the gravitational field as a gauge one for the Lorentz group $SO(3,1)$ (see also [32]). In his work, the Lorentz connection for the first time appeared as the corresponding gauge potential. However, the weak point in the approach of R. Utiyama was the introduction of the tetrads (metric) beyond the gauge scheme and an ad hoc requirement of the absence of the torsion. Eventually, these shortcomings were recognized. D. Sciama [33] dropped the condition of the vanishing torsion tensor. K. Hayashi [34] and B. De Witt [35] pointed out that the metric could be associated with a group of the spacetime translations $T_4$ (more precisely, with the general coordinate transformations, understood as translations), see also [13, 37]. T. Kibble [36] considered the Poincaré group $P_{10} = SO(3, 1) \rtimes T_4$ as a fundamental spacetime symmetry, which allowed him to give a gauge interpretation of the connection and the metric, cf. also [38]. Taking the Hilbert-Einstein linear in the curvature Lagrangian as a function of the independent metric and connection, D. Sciama and T. Kibble derived the field equations of the gauge $P_{10}$-theory of gravity, later named by A. Trautman [6] as the Einstein-Cartan theory (ECT). The thorough analysis of this theory is presented in the papers of F. Hehl [7] and A. Trautman [6] (see also [15, 16]).

Chapter 2

The canonical formalism of systems with singular Lagrangians was constructed by P. Dirac [40]. He was also the first who applied this formalism to the gravitational field and reduced the Hilbert-Einstein gravitational action of the canonical form [40]. The Hamiltonian formalism of the gravitational field was thoroughly studied in the papers of R. Arnowitt, S. Deser, and C. Misner [41]. The same program was realized by K. Kuchař [42, 43, 44] on the invariant
language making use of the system of arbitrary space-like hypersurfaces, also
extending the canonical formalism to the case of arbitrary matter field
minimally interacting with the gravitation. In our presentation of the canonical
formalism of gravitating systems, we followed the works of this author.

The proof of the degeneracy of the Lagrangian for the action invariant with
respect to the local group, was borrowed from the lecture of V. P. Frolov at the
Lebedev Physical Institute of the Russian Academy of Sciences.

The relation of the group of invariance of the action in the configuration
space with the corresponding system of canonical transformations in the phase
space, forming the “pseudogroup”, was studied in the papers of E. S. Fradkin,
G. A. Vilkovisky [45, 46], and I. A. Batalin [47]. The same was done in the paper
of C. Teitelboim [48] who discussed the general coordinate invariance group for
the action of the physical fields in the curved spacetime.

Chapter 3

The systematic method of the choice of dynamically independent degrees
of freedom of the gravitational field was constructed by R. Arnowitt, S. Deser,
and C. Misner [41], and subsequently generalized by K. Kuchař [44, 52] to the
case of arbitrary gravitating systems with the matter sources.

An important difference of the dynamics of the asymptotically flat open world
from the closed worlds was emphasized by De Witt [53], who attracted attention
to the significance of the surface term in the Hamiltonian of the asymptotically
flat gravitating world. The role of this surface integral in the Einstein theory
was also discussed by T. Regge and C. Teitelboim [51], who connected it with
the generator of the time translations at a spatial infinity.

The dynamics of the linearized gravitational field presented here has been con-
sidered by ADM [41], which identified the two transverse traceless gravitational
wave modes and constructed the Hamiltonian of the gravitational radiation.

Chapter 4

É. Cartan [79] was the founder of the spinor theory. B. Wan der Waerden
and L. Infeld [80] introduced the 2-spinors. H. Weyl [84], V. A. Fock, and
D. D. Ivanenko [85] developed a framework for the spinor analysis in the Rie-
mannian space $V_4$. The expressions obtained by them for the spin connections
entered the literature as the Fock-Ivanenko coefficients. The Dirac equation
in the Riemann-Cartan spacetime $U_4$ was first considered by H. Weyl [83].
V. I. Rodichev [55] demonstrated that the Dirac equation (with the zero mass)
in the Minkowski space $M_4$ with torsion becomes a non-linear spinor equation
of Ivanenko-Heisenberg [56] type (that underlies the unified theory of matter
developed by W. Heisenberg [57]). This result was generalized by T. Kibble [36]
to the spaces $U_4$. The spinor analysis in the space $G_4$ was constructed in [18].
Detailed presentations of the spinor formalism can be found in [81, 82].
The question, whether the torsion field interacts with vector fields is still open in the modern literature. A. Trautman [6] and F. Hehl [7] made a convincing observation that the massless vector field should not interact with the torsion since such a coupling would violate the gauge invariance (see also [89]). Nevertheless, in the book we decided to demonstrate that the opposite point of view has some interesting consequences, caused by this interaction. In particular, the non-linear nature of the vector field equations, arising due to the \((A-Q)\)-interaction, leads to the existence of the soliton solutions [63].

The conformal transformations introduced in Sec. 4.2., were also considered independently in [90, 91].

The first non-singular cosmological solutions in the ECT were obtained in [88, 123, 174]. The analysis of different types of spatially homogeneous cosmological models with the Weyssenhoff fluid is given in [175]. The analogy between the torsion and the cosmological \(\Lambda\)-term was indicated in [176, 177].

The first correct formulation of the problem of the particle production was developed for cosmology in GR by A. A. Grib, S. G. Mamaev, V. M. Mostepanenko [61], L. Parker [76], Ya. B. Zeldovich, A. A. Starobinsky [77]. We follow [61] in the presentation of the production of the scalar particles in the \(U_4\) space-time.

Chapter 5

The work of T. Kibble [36] is one of main contributions in the framework of so-called physical approach, where the direct localization of the global transformation of some group makes it necessary to “compensate” the corresponding non-covariance of the ordinary derivative by the introduction of the gauge field. Later works in this direction essentially were focused on the finding of the “true” gauge group for the theory of gravity, leaving the main structure of the physical approach unchanged (in the work [94], one can find an incomplete list of the groups and space-time structures). The criticism of this approach is contained in [96]. Along with this, the geometrical approach was developed on the basis of an idea of identifying the potential of the gauge field with a connection on some bundle over the spacetime. For the Yang-Mills theory, this fact was first noticed in [95] and later it was fully developed in investigations of A. Trautman [97].

A somewhat intermediate position between the physical and the geometrical approaches is held by the direction, developed by the group of F. Hehl [7, 98], which is based not so much on the gauge concept but on the fundamental link of the local invariance principle with the conservation laws. The latter studies established a remarkable fact (see also [38]): the source of the gauge field of the group \(G\) is the canonical Noether current corresponding to \(G\)-invariance of the theory.

The current status of the gauge theory of gravity is characterized by the application of the modern differential-geometric and topological methods for the study of their structure. The connection structure in the bundle of affine frames
presented in this book was first proposed in [99, 100], as well as independently developed in several different formulations in [101]-[105].

The idea to relate the existence of the gravitational field to the spontaneous breaking of a spacetime symmetry was in different forms expressed by many authors [97], [106]-[111]. The nonlinear realizations were introduced earlier in the study of the phenomenological Lagrangians in the theory of chiral symmetries [112, 113, 114]. The analysis of the dynamics of the non-linear gauge fields was considered in the paper of Y.M. Cho [116], and the authors followed this work in the description of spontaneous breaking in the Higgs mechanism.

Chapter 6

The Lagrangians and the theories based on them considered in Sec. 6.1. were studied in the following papers: (b) – in [128] (and more consistently in [129]); (d) see also [129]; (e) – in [130]. The general Lagrangians constructed as a sum of all possible terms quadratic in the curvature and the torsion (with arbitrary coefficients) were studied in [131, 132, 133, 138]. Some authors considered the possibility to restrict the choice of the coupling coefficients by attracting additional physical and geometrical ideas such as the requirement of the absence of the “ghost” and tachyon poles in the tree-level propagator [134, 135], the validity of Birkhoff’s theorem [136], the correspondence with GR [131], etc.

In Sec. 6.3., when studying the correspondence of the $\mathcal{P}_{10}$-theory with GR, we used the method applied in [137] to the Lagrangian of a particular form. For a review of the studies of the issue of the cosmological constant in the gravity and supergravity, see in [139].

The gravitational instantons (such as Euclidean Schwarzschild, Taub-NUT, $CP_{2}$) in the Riemannian gravity theory are studied in [142, 143, 144].

Chapter 7

The method of canonical quantization of the degenerate systems and, in particular, of the gravitational field was pioneered by P. A. M. Dirac [40], B. S. DeWitt [53] L. D. Fadeev and V. N. Popov [154], E. S. Fradkin and G. A. Vilkovisky [45]. The path integral technique was introduced into the quantum theory in the works of Feynman [162]. The method of the covariant quantization of the gravitational field is studied in detail in [45, 46, 154, 157, 158, 159]. Many papers are dedicated to the quantization of the cosmological systems in the ADM formalism. We mention just a few of them, which more clearly characterize the main features of this approach [162, 163, 165, 178]. In the majority of works, the study of the quantum dynamics was conducted in the framework of the model assumptions about cosmology. For example, it was common to assume homogeneity, or a certain symmetry, to take the majority of the modes of the physical fields as frozen, to neglect them on one or another stage of evolution, etc.

A. Peres [166] introduced the equations of the Einstein-Hamilton-Jacobi theory as a link between the classical GR and the quantum geometrodynamics.
Quantum geometrodynamics was developed in the works of P. A. M. Dirac [40], B. S. DeWitt [53], and J. A. Wheeler [167]. The attempt to construct the path integral in the quantum geometrodynamics was made by H. Leutwyler [169], who, in particular, was the first to propose the correct local measure for the quantum gravity. The method of the path integration in the quantum geometrodynamics was constructed in [168, 204]. Old application of the quantum geometrodynamical method in special models can be found in very many works of which we mention only a few: [165, 173, 179]-[182]. A special mention deserve the attempts of solving the Wheeler-DeWitt equation nonperturbatively [117]-[119] in the approximation of the so-called ultralocal state in the quantum gravitodynamics [117].

The Batalin-Fradkin-Vilkovisky (BFV) path integral and operator quantization of constrained systems was developed in [214, 215, 216, 217, 218, 219, 220] starting with the introduction of the relativistic phase space and the nilpotent BRST operator acting in their representation space [213, 46]. The Batalin-Marnelius procedure of gauging out the BRST symmetry in the physical subspace of this space was built in [221, 222, 223] and used for the construction of the inner product in quantum geometrodynamics in [204, 206, 208, 207, 210]. These works also convey the consistency between the reduced (physical) phase space quantization and the Dirac quantization in generic constrained systems and gravity theory, in particular. This consistency is demonstrated explicitly in the one-loop order of the semiclassical expansion in [205, 206, 207, 208, 209], where also the operator realization of quantum Dirac constraints is derived in the subleading (one-loop) order in the Planck constant.

Widely acclaimed prescriptions for the quantum state of the Universe in the form of the no-boundary or tunneling wavefunction can be found in [232, 233, 234, 235, 236, 237]. Recent applications of the path integral method in quantum geometrodynamics are the construction of the microcanonical density matrix of the early Universe [239], the calculation of its statistical sum for models dominated by conformal quantum fields, which suppresses to zero in the ensemble the contribution of the Hartle-Hawking no-boundary state [238] and imposes an upper subplanckian bound on the energy scale of the new type of (hill-top) inflationary scenario [240, 241].

Further reading

Two books by M. Blagojević and F. Hehl [200, 201] on the gauge gravity were published since the first Russian edition of our monograph. They give a broad overview of the history and of the recent developments, providing a complementary discussion of both mathematical and physical aspects of the subject. In an attempt to fill all the gaps and omissions that reflect our selection of the topics for the current book, we compiled an essentially complete bibliography on the gauge approach in gravity and related issues. The alphabetically ordered list can be found at the end of this book on page 201.
References


[56] Ivanenko D., Bemerkungen zur theorie der wechselwirkung, Physikalische Zeitschrift der Sowjetunion 13 (1938) 141-150.


References


[161] Frolov V. P. and Vilkovisky G. A., Quantum gravity removes classical
singularities and shortens life of black holes, in: Proc. of the 2nd Marcel
Grossmann Meeting on General Relativity, Trieste, Italy, 1979, Ed.

[162] Feynman R.P., Space-time approach to non-relativistic quantum mechan-

[163] Lapchinsky V.G. and Rubakov V.A., In: “Problems of the theory of
gravitation and elementary particles”, ed. K.P. Stanyukovich (Atomizdat:

[164] Barvinsky A.O. and Ponomarev V.N., Global effects of the production of

Rev. D8 (1973) 3271-3285.

[166] Peres A., On Cauchy’s problem in general relativity. II, Nuovo cimento


[168] Barvinsky A.O., Methods of the Quantum Theory of Gravitation and
Some of Their Applications, PhD thesis (Moscow State University:
Moscow, 1981).

[169] Leutwyler H., Gravitational field: Equivalence of Feynman quantization

[170] Dirac P.A.M., On the analogy between classical and quantum mechanics,
Revs Mod. Phys. 17 (1945) 195-199.

[171] Vilkovisky G.A., On the quantization of a particle in an external gravita-
359-368 (in Russian)].

[172] Cheng K.S., Quantization of a general dynamical system by Feynman’s


(1972) 219-220.

Pol. B6 (1975) 537-554.


References


Subject index

Action
  of gravitational field
    in GR, 11, 36
  of Yang-Mills type, 108, 122, 127

Affine
  connection structure, 1
  frame, 92
  tangent bundle, 92, 96
  transformations, 92

Anholonomity object, 7

Arnowitt-Deser-Misner energy, 51

Bianchi identity, 6

BRST
  equation, 149
  inner product, 144
  operator, 144

Bundle
  associated, 92, 96, 167
  of affine frames, 92
  of linear frames, 90, 92, 167
  of orthonormal frames, 6, 101
  principal, 6, 167
  spinor, 169
  tangent, 167
  vector, 166

Canonical (soldering) form, 90

Canonical momentum, 20
  gravitational, 37
  in geometrodynamics with
    matter sources, 41
  of matter field, 41

Christoffel symbol, 4

Commutation relations, 83, 129, 140

Conformally invariant gravity
  theory in $U_4$, 77

Connection, 2, 167
  defect tensor, 4
  form, 91
  generalized affine, 94, 96
  local Lorentz, 7
  semi-symmetric, 12
  spinor, 170

Conservation of total angular
  momentum, 16, 113

Constraint
198 Subject index

first kind, 22
primary, 20, 23
second kind, 22
Contortion tensor, 4
Coordinate frame, 6, 164
Cosmological term, 78, 116
Cotangent space, 91, 164
Covariant derivative, 2, 168
spinor, 172
Creation and annihilation operators, 83, 136
Cross-section, 90
Curvature
  form, 91
  local Lorentz, 7
tensor, 3
  decomposition, 5
  homothetic, 3, 127
De Sitter
  boost, 104
group, 103
universe, 123
  isotropic, 54, 75,
  123, 134
Density
  matrix, 137
number of particles, 85, 86
DeWitt supermetric, 37
Dilation current, 122, 126, 128
Dirac
  brackets, 25
  matrix, 171
  quantization, 140
Einstein’s
  λ-transformations, 72
  equation, 8
tensor, 6
universe, 78, 123
Einstein-Cartan theory, 13
Electromagnetic field, 20, 26
Energy-momentum tensor
  canonical, 14, 16
  conservation law in ECT, 16
  conservation law in gauge gravity, 113
  conservation law in GR, 8, 16
  effective, 16
  metrical, 8, 14
  of Dirac spinor, 63
  of fluid, 61
Exterior
differential form, 164
product, 165
Extrinsic curvature, 32
Faddeev-Popov
  ghost fields, 132
  operator, 131
Frenkel condition, 58
Friedman flat universe, 81
Gauge
  conditions, 25
  gravity theory, 87, 106
  group, 88
  theory, 87
Gauss-Bonnet identity, 115
Gauss-Codazzi identity, 34
Gauss-Weingarten formula, 32
General
  affine group, 92
  linear group, 90, 121
Ghost field, 122
Goldstone field, 89, 97, 99,
  102, 125
Gravitational
  constraints, 39, 44
  field Hamiltonian, 38, 50
Hamilton-Jacobi
  equation, 28, 138
  theory, 138
Higgs
  field, 99
  mechanism, 99
Hilbert action, 8
Hodge operator, 110, 166
Hypermomentum, 121
Hypersurface, 30
Ideal fluid, 58
Infinitesimal transformation
  of coordinates, 15, 18, 30, 88
Instantons, 117, 118
Subject index

Involution relations, 23
Ivanenko-Heisenberg equation, 64
Killing vector, 30, 49
Lapse and shift functions, 37, 53
Lie derivative, 15, 30, 41
Local Lorentz
  (anholonomic) frame, 7
  connection, 7, 101
Lorentz group, 4, 8

Manifold, 1, 163
Matter with microstructure, 59
Method of diagonalization
  of Hamiltonian, 84, 136
Metric structure, 1
Metric-affine gravity, 121
Metricity condition, 3

Nieh-Yan topological invariant, 118
Nonlinear gauge field, 98, 105
Nonmetricity tensor, 3
Normal basis on the hypersurface, 31

Palatini
  equation, 16
  principle, 11
Parallel transport, 2, 7, 168
Particle
  interpretation, 83, 136
  relativistic, 19, 26
Path integral, 130
Pauli matrix, 170
Penrose-Hawking energy condition, 75, 126
Poincaré group, 100, 107
Pontryagin index, 118
Regge-Teitelboim theorem, 51
Ricci identity, 5
Riemann-Cartan
  connection, 4
  space, 7
Riemannian space, 1
Scalar field with conformal
  coupling, 81
  Schrödinger equation, 130, 133
  sin-Gordon equation, 71
  Singular Lagrangian, 19
  Spin density tensor, 14
    of Dirac spinor, 63
    of fluid, 61
Spontaneous
  breaking of conformal
    symmetry, 124
    symmetry breaking, 97, 99
  SuperHamiltonian, 41, 135
  Supermomenta, 43
  Superspace, 140

Tangent
  bundle, 6, 89, 167
  space, 6, 90
  vector, 164
Tetrad coefficients (tetrads), 6
Third quantization, 161
Three-dimensional
  derivative, 31
  metric of hypersurface, 31
Topological Hausdorff space, 163
Torsion
  form, 91
  tensor, 3
  decompossition, 4
Translations, 89, 101
Truncated Lagrangian, 10

Unitarizing Hamiltonian, 145, 151
Vector
  bundle, 166
  field, 164
  non-linear, 68
Weyl
  group, 6, 171
  space, 7
  transformations, 71, 124
  vector, 4, 128
Weyssenhoff spinning fluid, 58
Wheeler-DeWitt equation, 141
Yang-Mills equation, 122
Bibliography on gauge gravity theory and related issues


13. Adak M. and Grumiller D., 


206 Bibliography on gauge gravity theory

B

Bibliography on gauge gravity theory


169. Baker W.M., Cosmic strings in Riemann-Cartan space-time, Class. Quantum Grav. 7 (1990) 717-730.


245. Berredo-Peixoto G. de, and Maică A.E., Form factors and non-local multiplicative anomaly for fermions with background torsion, Class. Quantum Grav. 31 (2014) 125002.


262. Blagojević M. and Cvetković B., Black hole entropy in 3D gravity with torsion, Class. Quantum Grav. 23 (2006) 4781.


Bibliography on gauge gravity theory


Bibliography on gauge gravity theory


432. Carlston S., Vignolo S., and Fabbrini L., The dynamics of Bianchi I universes in \( \mathbb{R}^n \) cosmologies with torsion, Class. Quantum Grav. 30 (2013) 205010.


224 Bibliography on gauge gravity theory


Bibliography on gauge gravity theory


Bibliography on gauge gravity theory

732. Drechsler W., Gauge theory for extended elementary objects, Class. Quantum Grav. 6 (1989) 623-657.


Bibliography on gauge gravity theory


Fabbri L. and Vignolo S., Dirac fields in f(R)-gravity with torsion, Class. Quantum Grav. 28 (2011) 125002.
<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
<th>Journal/Volume</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>815</td>
<td>Fariborz A.H. and McKeon D.G.C.</td>
<td>Spinors in Weyl geometry</td>
<td>Class. Quantum Grav. 14 (1997)</td>
<td>2517-2525</td>
</tr>
<tr>
<td>821</td>
<td>Fedorov F.I., Minkevich A.V., and Chuong Nguyen Hong</td>
<td>Spontaneous symmetry breaking by gravitational field and regular isotropic models with torsion</td>
<td>Class. Quantum Grav. 5 (1988)</td>
<td>515-519</td>
</tr>
<tr>
<td>831</td>
<td>Ferraris M., Francaviglia M., and Raiteri M.</td>
<td>Conserved quantities from the equations of motion: with applications to natural and gauge natural theories of gravitation</td>
<td>Class. Quantum Grav. 20 (2003)</td>
<td>4043</td>
</tr>
</tbody>
</table>


Bibliography on gauge gravity theory


903. Fukui M. and Masukawa J., Massless torsion fields. II. The case $\alpha + \frac{\omega}{3} = 0$, Progr. Theor. Phys. 73 (1985) 75-83.


G


934. Galiakhmetov A.M., Exact isotropic scalar field cosmologies in Einstein-Cartan theory, Class. Quantum Grav. 27 (2010) 055008.

935. Galiakhmetov A.M., Cosmology with polynomial potentials of the fourth degree in Einstein-Cartan theory, Class. Quantum Grav. 28 (2011) 105013.


994. Garcia De Andrade L.C., Cosmological magnetic fields as string dynamo seeds and axion fields in torsioned spacetime, JCAP 08 (2014) 023.


998. Garcia De Andrade L.C., Cosmological and astrophysical consequences from the magnetic dynamo equation in torsioned spacetime and teleparallel gravity, Class. Quantum Grav. 23 (2016) 025006.


1017. Garecki J., Gauge theory of gravitation with quadratic Lagrangian \(L_g = \alpha \Omega^{ij}_k \wedge \eta^{ij}_k + \beta \Omega^{ik}_k \wedge \Omega^{jk}_j + \gamma \Theta^{ij}_k \wedge \Theta^{jk}_j\) containing Einsteinian term \(\alpha \Omega^{ij}_k \wedge \eta^{ij}_k\) and spherically symmetric solutions to its field equations, in: “On Relativity Theory. Proc. of Sir A.Eddington Centenary Symp.,” Eds. Y. Choquet-Bruhat and T.M. Karade (World Scientific: Singapore, 1985) vol. 2, 232-260.

1018. Garecki J., New physical interpretation of the Bel-Robinson tensor, Class. Quantum Grav. 2 (1985) 403-408.


1020. Garecki J., Revised gauge theory of gravitation with quadratic Lagrangian \(L_g = a \Omega^{ij}_k \wedge \eta^{ij}_k + b \Omega^{ik}_k \wedge \Omega^{jk}_j + c \Theta^{ij}_k \wedge \Theta^{jk}_j\), Acta Phys. Pol. B20 (1989) 997-1014.


Bibliography on gauge gravity theory

1045. Germán G., Macías, and Obregón O., Kaluza-Klein approach in higher-dimensional theories of gravity with torsion, Class. Quantum Grav. 10 (1993) 1045-1053.


\( H \)


1156. Hammond R.T., Propagating torsion as a massive vector field, Class. Quantum Grav. 7 (1990) 2107-2112.


1171. Hammond R.T., Torsion gravity, the Brans - Dicke type generalization and string theory, Class. Quantum Grav. 13 (1996) L73-L79.


I


1356. Iorio L., Modified theories of gravity with nonminimal coupling and orbital particle dynamics, Class. Quantum Grav. 31 (2014) 085003.


Bibliography on gauge gravity theory


Bibliography on gauge gravity theory


Bibliography on gauge gravity theory


1667. Krisch J.P. and Smalley L.L., Spin fluids in van Stockum cylinders, Class. Quantum Grav. 7 (1990) 481-486.


Bibliography on gauge gravity theory


1733. Lavrov P.M. and Moshin P.Yu., Quantization of two-dimensional gravity with dynamical torsion, Class. Quantum Grav. 16 (1999) 2247-2258.


1823. Ma M.-S., Liu F., and Zhao R., Continuous phase transition and critical behaviors of 3D black hole with torsion, Class. Quantum Grav. 31 (2014) 095001.


<table>
<thead>
<tr>
<th>Year</th>
<th>Authors</th>
<th>Title</th>
<th>Journal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1917</td>
<td>McCrea J.D.</td>
<td>Irreducible decompositions of non-metricity, torsion, curvature and Bianchi identities in metric-affine spacetimes,</td>
<td>Class. Quantum Grav. 9 (1992) 553-568.</td>
</tr>
</tbody>
</table>
Bibliography on gauge gravity theory


294 Bibliography on gauge gravity theory


Bibliography on gauge gravity theory


Bibliography on gauge gravity theory


298 Bibliography on gauge gravity theory


Bibliography on gauge gravity theory


Bibliography on gauge gravity theory


O


Bibliography on gauge gravity theory


Bibliography on gauge gravity theory


308 Bibliography on gauge gravity theory


P


Pasic V. and Vassiliev D., pp-waves with torsion and metric-affine gravity, Class. Quantum Grav. 22 (2005) 3961.


Bibliography on gauge gravity theory


2419. Poplawski N., The energy and momentum of the Universe, Class. Quantum Grav. 31 (2014) 065005.
Bibliography on gauge gravity theory


Bibliography on gauge gravity theory


2511. Rizaoglu E., Two-dimensional nonlinear fermionic models in curved space with torsion, Class. Quantum Grav. 11 (1994) 2383-2394.


2545. Rumpf H., Creation of Dirac particles in general relativity with torsion and electromagnetism. II. The constant electric field - a pedagogical example, Gen. Relat. Grav. 10 (1979) 525-533.


Bibliography on gauge gravity theory


2630. Scipioni R., Isomorphism between non-Riemannian gravity and Einstein-Proca-Weyl theories extended to a class of scalar gravity theories, Class. Quantum Grav. 16 (1999) 2471-2478.


2634. Seeger A., Recent advances in the theory of defects in crystals, Physica status solidi (b) 1 (1961) 669-698.


2637. Seitz M., Proca field in space-time with curvature and torsion, Class. Quantum Grav. 3 (1986) 1265-1273.

2638. Seitz M., A quadratic Lagrangian for the Poincaré gauge field theory of gravity, Class. Quantum Grav. 3 (1986) 175-182.
2643. Sengupta S., Dark energy from the gravity vacuum, Class. Quantum Grav. 32 (2015) 195005.
2655. Shapiro I.L. and Teixeira P.M., Quantum Einstein-Cartan theory with the Holst term, Class. Quantum Grav. 31 (2014) 185002.
2661. Shestopal V.O., Feasibility of describing the gravitational field as a space with absolute parallelism, Sov. J. Phys. 16 (1973) 1692-1694 [Izvestia VUZov, Fizika n. 12 (1973) 83-85 (in Russian)].
324 Bibliography on gauge gravity theory


2680. Šijački Dj., \( SL(4, R) \) embedding for a 3D world spinor equation, Class. Quantum Grav. 21 (2004) 4575.


326 Bibliography on gauge gravity theory


Bibliography on gauge gravity theory


2730. Smalley L.L. and Ray J.R., Wegssenhoff spinning fluid in Riemann-Cartan spacetime: direct variation with respect to the spin density of the fluid, Class. Quantum Grav. 7 (1990) 1445-1461.


Bibliography on gauge gravity theory


2781. Spinosa R., Proca test field in a space-time with torsion, Class. Quantum Grav. 4 (1987) 473-484.


2808. Sur S. and Bhatia A.S., Constraining torsion in maximally symmetric (sub)spaces, Class. Quantum Grav. 31 (2014) 025020.


332 Bibliography on gauge gravity theory

2846. Tod K.P., Conical singularities and torsion, Class. Quantum Grav. 11 (1994) 1331-1339.
Bibliography on gauge gravity theory


334 Bibliography on gauge gravity theory


336 Bibliography on gauge gravity theory

U


V


Bibliography on gauge gravity theory


2957. Vasiliić M., String background fields and the Riemann-Cartan geometry, Class. Quantum Grav. 28 (2011) 075008.


W


3006. Wanas M.I., Youssef N.L., and Sid-Ahmed A.M., Teleparallel Lagrange geometry and a unified field theory, Class. Quantum Grav. 27 (2910) 045005.


Z


Bibliography on gauge gravity theory


Научное издание

Владимир Николаевич Пономарев,
Андрей Олегович Барвинский,
Юрий Николаевич Обухов

КАЛИБРОВОЧНЫЙ ПОДХОД
И МЕТОДЫ КВАНТОВАНИЯ
В ТЕОРИИ ГРАВИТАЦИИ

Издательство «Наука»
117997, Москва, Профсоюзная ул., 90
Главный редактор Т. Е. Филиппова
Редактор Л. В. Филиппова

Формат 70 х 100 1/16. Бумага офсетная 80 г/м²
Печать офсетная. Гарнитура «CMR»
Уч.-изд. л. 28,0. Усл.-печ. л. 29,03
Тираж 1000 экз. Заказ № 734517

Заказное
Отпечатано ФГУП Издательство «Наука»
Contraia non contradictria
red complementa sunt

Nietzsche

Physical law should have mathematical beauty
P.A.M. Drexel
3 Oct 1951

Nature is simple in its essence.
July 28, 59, Yulpana

自然は本質において
簡単である
湯川
(佐藤文隆)