Theoretical and Mathematical Physics

Yuri E. Gliklikh

Global and Stochastic Analysis with Applications to Mathematical Physics

Springer

Global and Stochastic Analysis with Applications to Mathematical Physics

Theoretical and Mathematical Physics

The series founded in 1975 and formerly (until 2005) entitled *Texts and Monographs in Physics* (TMP) publishes high-level monographs in theoretical and mathematical physics. The change of title to *Theoretical and Mathematical Physics* (TMP) signals that the series is a suitable publication platform for both the mathematical and the theoretical physicist. The wider scope of the series is reflected by the composition of the editorial board, comprising both physicists and mathematicians.

The books, written in a didactic style and containing a certain amount of elementary background material, bridge the gap between advanced textbooks and research monographs. They can thus serve as basis for advanced studies, not only for lectures and seminars at graduate level, but also for scientists entering a field of research.

For other titles published in this series, go to
www.springer.com/series/720

Yuri E. Gliklikh

Global and Stochastic Analysis with Applications to Mathematical Physics

Springer

Prof. Yuri E. Gliklikh
Voronezh State University
Mathematics Department
Universitetskaya pl. 1
394006 Voronezh
Russia
yeg@math.vsu.ru

ISSN 1864-5879 e-ISSN 1864-5887
ISBN 978-0-85729-162-2 e-ISBN 978-0-85729-163-9
DOI 10.1007/978-0-85729-163-9
Springer London Dordrecht Heidelberg New York

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Mathematics Subject Classification (2010): 58-02, 60-02, 58D30, 58J65, 60H30, 70G45, 76M35, 81P20, 82C31

Cover design: eStudio Calamar S.L.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To the memory of my parents and my wife Olga

Preface

Since the publication of my previous books [99, 106], many papers and monographs have appeared which develop similar themes. So, now is a good time for an "intermediate finish", a new book containing results obtained up to the present.

The themes of the book cross the borders of several mathematical subjects including the geometry of manifolds, stochastic analysis, set-valued analysis and some chapters of mathematical physics. Thus, an important feature of the book is that it includes a significant amount of preliminary material. However, my original intention to make the book completely self-contained yielded such an incredible volume of text that I had to reduce the preliminaries and assume the reader to be at least a little familiar with the aforementioned branches of mathematics. Nevertheless, I hope that the remaining preliminaries will be useful for an expert in one of the subjects who wishes to understand the others and will also familiarize non-experts with the general themes of this work.

I want to express my thanks to a lot of people who aided me in my research. First of all, I am obliged to my advisor Yu.G. Borisovich who introduced me to mathematics. In particular, I owe my interest in Global Analysis and its applications to Mathematical Physics to him. Another strong influence (this time in Stochastic Analysis) was exerted on my research by K.D. Elworthy. I am indebted to him for long and extremely useful discussions of the problems that interested me, as well as for repeated invitations to Warwick University, allowing me to ventilate my ideas with many mathematicians from all over the world. In addition, I should specially mention Yu.L. Daletskiĭ who was also a great help.

I express my gratitude to my colleagues and co-authors: Ya.I. Belopolskaya, B.D. Gelman, I.V. Fedorenko and T. Zastawniak and my former students and co-authors: S.V. Azarina, Yu.S. Baranov, L.A. Morozova, A.V. Obukhovskii, P.S. Ratiner and P.S. Zykov for their interest in our joint work. I am grateful to Z. Brzezniak, V.Ya. Gershkovich, A.I. Shnirelman, A. Truman, A.M. Vershik and many others for very useful discussions.

My research, the results of which are included in this book, was supported in part by INTAS Grants 94-00378 and 99-00559, RFBR Grants 96-01-00444, 99-01-00819, 03-01-00112, 04-01-00081, 07-01-00137 and 08-01-00155 as well as by Grants UR.04.01.003 and UR.04.01.008 of the program "Universities of Russia" and by Grant VZ-010-0 of the Russian Ministry of Education and CRDF. I am grateful to all these foundations.

Yuri E. Gliklikh
September, 2009

Contents

Introduction

The main aim of this book is to develop and combine the methods of Global Analysis and Stochastic Analysis allowing a more or less common treatment of areas of mathematical physics that are traditionally considered distant from one another and which have formerly required different methods of investigation. Among the areas we mention are classical mechanics on non-linear configuration spaces and some problems of statistical physics, quantum physics and hydrodynamics. The idea, yielding the unification of these topics, is based on the use of a geometrically invariant form of Newton's second law and its analogs (stochastic, set-valued, infinite-dimensional, etc.) as a fundamental equation of motion. The realization of this idea allows one to elaborate general approaches to the investigation of the above-mentioned theories whose modification in each concrete case permits us to create effective methods of investigation and to obtain new important results.

The principal project of the book incorporates a huge amount of mathematical machinery including, among other things, some branches of global analysis, stochastic analysis, set-valued analysis and analysis on infinite dimensional manifolds. The large amount of space devoted to preliminary material and recent results in the above-mentioned branches of mathematics is a result of the author's desire to make the book as self-contained as possible. Some of this preliminary material can be used as a first introduction to the subject. In those cases where a detailed description of the material would be too lengthy to include, we simply give a survey of notions and constructions without detailed proofs. Generally we limit ourselves to the material that is necessary for the applications to mathematical physics which appear in the last part of the book. Nevertheless, for the sake of completeness, we often describe notions, results and constructions that are not used in the book but that are close to its main theme.

The book consists of three parts (subdivided into chapters, sections and sometimes subsections), devoted, respectively, to global analysis, stochastic analysis and applications to mathematical physics. The contents of these parts are interrelated, however each part can be read independently. The first part contains an introduction to the geometry of manifolds from the very beginning to branches that are often absent in textbooks and monographs. In particular, it includes the machinery that is used in stochastic analysis on manifolds. Thus the first two parts together present a rather complete introduction to the latter. Some chapters of Part 3 are related only to Part 1 and can be read without reference to the material of Part 2. Other chapters use results and constructions from both Part 1 and Part 2.

The material of Part 1 is devoted to global analysis and forms the basis for the subsequent exposition. We begin with a glossary of commonly used definitions and formulae from the theory of manifolds. Following this we look at Lie groups and algebras, fiber bundles and related topics, and introduce Riemannian metrics, tensors, differential forms and Lie derivatives.

Most important for the forthcoming material is Chapter 2, devoted to the theory of connections. For vector bundles we use the approach to connections based on the so-called connectors (connection maps) first introduced by P. Dombrovski in the 1960s (see [56]). This is easily generalized to, say, manifolds of maps. The connections on manifolds are introduced as connections on their tangent bundles. We look at curvature and torsion tensors and Riemannian connections (in particular, the Levi-Civitá connection), connections on principal bundles and a connection on the total space of a vector bundle generated by a connection on the bundle and by a connection on the base. We conclude the chapter with the notion of second order tangent vectors that are transformed into ordinary (i.e., first order) tangent vectors by a connection.

Chapter 3 is devoted to ordinary differential equations on manifolds. The first topic addresses the necessary and sufficient conditions for completeness of flows of vector fields. We show how to modify the sufficient conditions for completeness in both one-sided and two-sided cases in order to obtain necessary and sufficient conditions. In both cases the necessary and sufficient conditions involve a transition to the extended phase space. One-sided conditions are formulated in terms of the existence of a proper function on the extended phase space such that its derivative in the direction of the natural extension of the right-hand side of the equation onto the extended phase space is uniformly bounded. For this type of necessary and sufficient condition we also find a natural infinite-dimensional generalization that is applicable also to some cases where the right-hand side is given only on an everywhere dense subset of the phase space.

Two-sided conditions are formulated in terms of the existence of a complete Riemannian metric on the extended phase space such that the norm of the natural extension of the right-hand side of the equation is uniformly bounded with respect to this metric.

We then describe the basic construction and properties of integral operators with parallel translation, elaborated by the author. These operators and their stochastic generalizations are applied to the investigation of various equations below.

In the remaining part of the chapter we describe second order differential equations on manifolds as special vector fields on tangent bundles and as given in terms of covariant derivatives. The latter, in particular, involve geodesic sprays of connections. For completeness we conclude the chapter with a brief account of Hamiltonian systems.

Chapter 4 is devoted to elements of set-valued analysis and forms the basis for all following set-valued problems. Besides the standard notions of upper and lower semi-continuous set-valued maps, differential inclusions, etc., here we also present two new results establishing the existence, in the finite-dimensional case, of a sequence of special ε-approximations for an upper semi-continuous set-valued map with convex values that point-wise converges to a Borel measurable selector of the set-valued map as $\varepsilon \to 0$. These results are important in the later study of various stochastic differential inclusion problems.

In Chapter 5 we describe analysis on groups of Sobolev diffeomorphisms of a compact manifold. These groups are natural configuration spaces for systems of hydrodynamics in the framework suggested by Arnold-Ebin-Marsden. The case of the group of diffeomorphisms of a flat torus is of special interest since a fluid motion on the torus is often considered in classical hydrodynamics.

The second part is devoted to the constructions and results from contemporary stochastic analysis with the main focus on stochastic analysis on manifolds. In Chapter 6 we recall some material that is not generally included in a standard university course. We describe, among other topics, conditional expectations, martingales, weak convergence of probability measures, stochastic integrals and stochastic differential equations in Euclidean spaces (both in Itô and in Stratonovich forms) and stochastic flows and their generators (forward and backward).

In the next chapter we pass to manifolds. It should be pointed out that topological and geometric constructions are used considerably in stochastic analysis on manifolds. Thus the material of this chapter is based on the material of Part 1 (first of all on the theory of connections) and, of course, on the material of the previous chapter. First we describe stochastic differential equations on manifolds in Stratonovich form. This formalism is widely used since the right-hand sides of such equations are transformed under coordinate changes like ordinary tangent vectors. Then we show that on each manifold there exists a Riemannian metric having the so-called uniform Riemannian atlas. If the coefficients of a Stratonovich equation are C^1-smooth and uniformly C^1 bounded with respect to such a metric, the flow is complete.

We then turn to Itô stochastic differential equations on manifolds. We describe an approach treating such equations as cross-sections of a special fiber bundle (the Itô bundle) with an interesting structure group, interrelated approaches elaborated by Belopolskaya and Daletskii and by Baxendale (based on the theory of connections), an approach based on integral operators with parallel translation along stochastic processes (due to the author), and some other contemporary constructions.

We should mention a necessary and sufficient condition for completeness of a stochastic flow, continuous at infinity. In some sense this is a generalization of the one-sided necessary and sufficient condition for completeness of the flow of a vector field described in section 3.1. There is also a result on criteria of weak compactness of measures on path spaces corresponding to solutions of stochastic differential equations on manifolds.

Besides the well-known Eells-Elworthy development we introduce another one, the so-called Itô development whose construction is based on Itô equations. On this basis we introduce the notion of Itô processes on manifolds (as Itô developments of Itô processes in tangent spaces), along which the parallel translation is well-defined. The use of Itô processes makes the construction of parallel translation simple and clear.

In addition, some general existence of solution theorems are proved, in particular for the so-called equations with unit diffusion coefficient on stochastically complete Riemannian manifolds.

The chapter concludes with the notion of a martingale with respect to a connection.

We then consider (in Chapter 8) a version of differential calculus for stochastic processes, Nelson's theory of mean derivatives. Later some equations of mathematical physics (in mechanics with random perturbation of forces or velocities, in quantum theory and in hydrodynamics) are given in terms of these derivatives.

The classical Nelson mean derivatives give information only about the drift of a stochastic process. By a slight modification of a certain idea of Nelson we introduce a new sort of mean derivative (called the quadratic mean derivative) that is responsible for the diffusion term. Considering the quadratic derivative together with Nelson's classical derivatives, we investigate first order differential equations and inclusions with mean derivatives: with forward mean derivatives, with backward mean derivatives and with the so-called current velocity, having a physical interpretation as a stochastic analog of ordinary physical velocity. In particular we prove some existence of solution theorems. We create a list of the first and second mean derivatives for a Wiener process, for solutions of Itô equations, for Itô diffusion type processes with unit diffusion coefficient, etc., that allow us in forthcoming chapters to prove the existence of solutions of higher order equations in mean derivatives.

The features of mean derivatives on manifolds are of special interest. The construction of forward and backward mean derivatives on a manifold involves a connection. It turns out that the forward and backward derivatives, determined with respect to a certain connection, are naturally related to Itô equations in Belopolskaya-Daletskii form, determined by the same connection. We show that forward and backward mean derivatives are related to forward and backward generators of a flow that is governed by the Itô equation (cross-section of the Itô bundle) corresponding to the Belopolskaya-Daletskii equation with respect to the given connection. This relationship is described in terms of the same connection as a fiber-wise linear mapping from the second order tangent bundle to the ordinary (i.e., first order) tangent bundle to the manifold.

Note that the quadratic mean derivative and current velocity are defined without using connections and have the form of a $(2,0)$-tensor field and a vector field, respectively.

Some existence of solutions theorems for equations and inclusions in mean derivatives on manifolds are proved.

We conclude Part 2 with elements of stochastic analysis on groups of diffeomorphisms. We consider right-invariant Itô stochastic differential equations in Belopolskaya-Daletskii form on general groups of diffeomorphisms and on volume preserving groups. The Wiener process, used in the construction of the equations, is finite-dimensional. In the general case it is taken from a Euclidean space in which the finite dimensional manifold is embedded by Nash's theorem. For the particular case of the group of diffeomorphisms of a flat n-dimensional torus a special n-dimensional Wiener process is constructed that allows one to apply the corresponding equations to the investigation of viscous hydrodynamics described below. Some existence of solution theorems are obtained.

Making use of the material of Parts 1 and 2, Part 3 is devoted to the description and investigation of various mechanical and physical systems. The exposition begins with a description of classical Newtonian mechanics in the language of invariant geometry and topology. Newton's second law is introduced in terms of the covariant derivative of the Levi-Civitá connection of a Riemannian metric that determines the kinetic energy on the configuration space. After introducing such mechanical systems in a very general form, we consider the special case of conservative systems, including Hamilton's principle of least action and Noether's theorem. We also consider systems with group structure, systems with discontinuous forces (where Newton's law is given in terms of differential inclusions), systems with delayed forces (described in terms of parallel translation), systems with constraints given in geometric form due to Vershik and Faddeev (including non-holonomic mechanics and the so-called vakonomic systems, i.e., variational problems with constraints), integral equations of geometric mechanics (involving parallel translations), velocity hodographs, and so on.

In Chapter 12 we apply the machinery developed above to the qualitative behavior of trajectories of mechanical systems. We consider the two-point boundary value problem for trajectories, i.e., whether it is possible to join two points of configuration space by a trajectory. It should be noted that on non-linear configuration spaces (i.e., on Riemannian manifolds), even for smooth bounded forces independent of velocities, this problem may not have a solution at all, unlike the case of linear configuration spaces. This may happen if the points are conjugate along all geodesics of the Levi-Civitá connection joining them (this is true for all types of forces, e.g., for forces depending on velocities with linear or quadratic growth). Besides, it is a well-known fact that for forces with quadratic growth, all trajectories starting at a certain point may be confined to some bounded domain for all time, unable to reach an exterior point. Examples of all these cases of non-solvability are described in the first section of the chapter.

A certain geometric condition is found for the geometry of a manifold, a pair of points and a force field, having quadratic growth, such that if the points are not conjugate along at least one geodesic and the condition is satisfied, the points can be connected by a trajectory at least in a small enough time interval. It is shown that if the force has less than quadratic growth, this condition is always satisfied and so the problem is solvable for all pairs of points, non-conjugate along at least one geodesic. If the manifold is flat (in particular, this means that conjugate points are absent) and the force is uniformly bounded, the construction yields the classical result that the problem is solvable for any pair of points in any time interval. For forces with quadratic growth an additional condition is found which, in combination with the first condition, ensures that the problem is solvable in any time interval. All results are proved for the general case of set-valued force fields with various continuity conditions (ordinary single-valued results follow as corollaries) and so they are connected with the problem of controllability for mechanical systems. A generalization to systems with non-holonomic constraints is also presented. In this case it is natural to investigate the problem of connecting a point with a certain submanifold.

A modification of the constructions of this chapter is later applied to a certain analogous problem on Lorentz manifolds.

In Chapter 13 we deal with the general theory of relativity. The material of the first section can be considered as an introduction to the subject, presented axiomatically, as is habitual for mathematicians. This part of the chapter is the basis for all following relativistic problems. We then investigate a certain two-point boundary value problem arising in A. Poltorak's concept of reference frame. In this concept the reference frame is a certain manifold equipped with a connection. On the basis of the machinery developed in Chapter 16, for two particular cases of reference frame some geometric conditions are found under which we can conclude from the fact that two events are connected by a time-like geodesic in the reference frame, that the same can be done in the space-time (i.e., if the second event belongs to proper future of the first one in the reference frame, the same is true in the space-time).

In the last section we describe the motion of a classical particle in a classical gauge field in terms of a special version of Newton's law on a fiber bundle with a connection (recall that mathematically the notion of a gauge field coincides with the notion of a connection on a fiber bundle). This section also contains a short introduction to the geometric theory of gauge fields.

In Chapter 14 we consider mechanical systems with random perturbation of either the force fields or of the velocities. Newton's laws for such systems are expressed in terms of forward mean derivatives and corresponding integral operators with stochastic parallel translation.

First we investigate the so-called Langevin equation on a Riemannian manifold. This is Newton's law for a mechanical system on a nonlinear configuration space whose force field takes the form $a(t, m(t), \dot{m}(t)) + A(t, m(t), \dot{m}(t))\dot{w}(t)$, where $a(t, m, X)$ is a deterministic vector force field, $A(t, m, X)$ is a $(1, 1)$-tensor field (i.e., a field of linear operators in tangent spaces), depending on velocity X, and $\dot{w}(t)$ is an Itô white noise in tangent space. In particular such an equation describes the motion of a physical Brownian particle in a non-linear configuration space. We present a well-posed mathematical description of this equation in terms of mean derivatives and of integral operators with parallel translation that avoids using distribution theory. Existence theorems for weak and strong solutions are proved. Natural analogs of Ornstein-Uhlenbeck processes on Riemannian manifolds are described. Generalizations to the case of the so-called Langevin differential inclusion (where both a and A are set-valued) are also considered.

We then consider the case where the velocity of a mechanical system trajectory is subjected to random perturbation. This situation is motivated by the motion of a particle, subjected to a deterministic force, that in addition moves in a medium under random influence. Such systems are described by Newton's law in terms of mean derivatives, whose form is different from that in the Langevin case. The stochastic integrals with stochastic parallel translation are applied to the investigation of such systems on manifolds. We also consider such systems in linear spaces (in particular, with set-valued forces) since some more general results can be obtained for them.

Another stochastic version of Newton's second law, namely the so-called Newton-Nelson equation, is considered in Chapter 15. It is given in terms of mixed second order mean derivatives and describes the motion of a quantum particle in the framework of Nelson's stochastic mechanics. The main result here is the existence of solution theorem where the force field is the sum of a vector field, independent of velocities, and a (1,1)-tensor field (i.e., a field of linear operators in tangent spaces), applied to the current velocity of the process. We investigate the non-relativistic case (in $\mathbb{R}^n$ and on a manifold) as well as the relativistic case (in Minkowski space and on a space-time of general relativity). In fact we obtain a revised version of stochastic mechanics that is free of the defects found by Nelson within his initial approach to this theory.

In Chapter 16 we describe hydrodynamics via the modern Lagrangian formalism suggested in the works of V.I. Arnold, D. Ebin and J. Marsden. This formalism arises from Newton's law on the group of Sobolev diffeomorphisms of a finite-dimensional manifold, formulated in terms of the covariant derivative of the Levi-Civitá connection of a weak Riemannian metric (determining the topology of the functional space L^2). The basic system here is the one of so-called diffuse matter. By considering a special force field we obtain the description of a perfect barotropic fluid and, by defining a special constraint, the description of a perfect incompressible fluid. If we collect the velocity vectors of a solution in the tangent space at the unit to the group by right translations, the curve that we obtain in that tangent space satisfies the Euler equation. The differential equation of motion (Newton's law) on the group usually has a smooth right-hand side. Passing to the Euler equation yields loss of derivatives. We prove, among other results, local existence of solutions theorems, regularity of solutions theorems (including the case of finite-dimensional manifolds with boundary) and a version of Noether's theorem.

For the description of viscous incompressible fluids we use stochastic analysis, particularly the machinery of mean derivatives. We mainly deal with the model problem of a fluid moving on an n-dimensional flat torus. A special second order equation is found with backward mean derivatives on the group of diffeomorphisms, subjected to a certain constraint (a special stochastic analog of Newton's law), such that the expectations of its solutions are flows of a viscous incompressible fluid on the torus. (It should be noted that the stochastic Newton law here is expressed in terms of backward mean derivatives and so it has a form different to that appearing in Chapters 14 and 15.) Passing to the Euler description yields a Navier-Stokes equation in the tangent space at the unit in complete analogy to the appearance of the Euler equation in the case of a perfect incompressible fluid.

In complete form this construction is realized under the assumption that the backward mean derivative of the process satisfying the stochastic Newton law, mentioned above, is generated by a right-invariant vector field. If this is not the case, in the "algebra", after passing to the Euler approach, some other types of hydrodynamical equations may arise.

We finish the chapter by introducing a special stochastic perturbation of a flow of diffuse matter on the group of diffeomorphisms such that the perturbed flow satisfies the stochastic Newton law and the corresponding curve in the tangent space at the unit satisfies the Burgers equation. The same perturbation of a perfect incompressible flow without external force satisfies the stochastic Newton law with zero force, but yields a curve in the tangent space at the unit that is a solution of a Reynolds type equation. Nevertheless, under the action of a certain special external force on the flow, this curve becomes a solution of a Navier-Stokes equation without external force. As above, we consider a fluid motion on the flat n-dimensional torus $\mathcal{T}^n$.

Everywhere in the book we use Einstein's summation convention:

Einstein's Convention.

A monomial with a shared upper and lower index represents the summation where the common index ranges from 1 to n, n being the dimension of the manifold under consideration.

For example, $a_i^{ik} = \sum_{i=1}^{n} a_i^{ik}$. Further to this convention, we treat upper indices appearing in a denominator as lower indices and lower indices appearing in a denominator as upper indices. Examples of such expressions are: $X = X^i \frac{\partial}{\partial q^i}$ and $P = P_i \frac{\partial}{\partial p_i}$.

Part I
Global Analysis

Chapter 1
Manifolds and Related Objects

1.1 Manifolds, Vectors and Covectors. A Glossary

A manifold will generally be denoted by the symbol M. Recall that a finite-dimensional manifold, as a topological space, is assumed to be Hausdorff and satisfy the second countability axiom. Unless otherwise stated, $\dim M = n$ for a finite-dimensional manifold M.

We denote the charts of a maximal atlas by the symbols $(\mathcal{V}_\alpha, \varphi_\alpha)$, where $\mathcal{V}_\alpha$ is an open ball in $\mathbb{R}^n$, and the corresponding neighborhoods in M by $\mathcal{U}_\alpha = \varphi_\alpha \mathcal{V}_\alpha$. Usually we do not distinguish between $\mathcal{V}_\alpha$ and $\mathcal{U}_\alpha$ and so the latter are also called charts. Recall that in this case $\mathbb{R}^n$ is called the *model space*.

The change of coordinates between $\mathcal{U}_\alpha$ and $\mathcal{U}_\beta$ (i.e., between $\mathcal{V}_\alpha$ and $\mathcal{V}_\beta$) is denoted by $\varphi_{\beta\alpha} = \varphi_\beta^{-1}\varphi_\alpha$. If the changes of coordinates are homeomorphisms, the manifold is called *topological*, and if C^k-smooth, a C^k*-manifold*. If $k = \infty$, the manifold is said to be *smooth*.

Convention 1.1 *Everywhere below, unless stated to the contrary, all manifolds are assumed to be smooth, i.e., C^∞-smooth.*

Theorem 1.2 (Whitney) *Every smooth manifold with dimension n can be embedded as a smooth surface into a linear space with dimension $2n + 1$.*

Finite dimensional manifolds are clearly locally compact. It is a well-known fact that all locally compact spaces satisfying the second countability axiom are paracompact (see, e.g., [30]). Thus every finite-dimensional manifold is paracompact.

We introduce coordinates $q^1, \dots, q^n$ in a chart $\mathcal{V}_\alpha$ as in a domain in $\mathbb{R}^n$ and analogously coordinates $q^{1'}, \dots, q^{n'}$ in another chart $\mathcal{V}_\beta$. As there is no chance of confusion, the corresponding coordinates in $\mathcal{U}_\alpha$ and $\mathcal{U}_\beta$, obtained by the homeomorphisms φ_α and φ_β, are denoted by the same symbols. They are called *local* or *curvilinear coordinates*. The change of coordinates $\varphi_{\beta\alpha}$ is

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_1,

given by expressing the coordinates $q^{i'}$ of every point $m \in \mathcal{U}_{\alpha\beta}$ in terms of its coordinates $q^j : q^{i'} = q^{i'}(q^1, \ldots, q^n)$. Of course the inverse mapping $\varphi_{\alpha\beta}$ is determined by expressing the coordinates q^i in terms of $q^{i'}$. The change of coordinates is *smooth* if all $q^{i'}(q^1, \ldots, q^n)$ and $q^j(q^{1'}, \ldots, q^{n'})$ are jointly smooth functions of all their variables.

On smooth manifolds the notion of a smooth mapping is well-defined. A mapping $f : M \to N$ is determined not only by a chart $\mathcal{U}_\alpha$ on M but also by a chart $\mathcal{U}_\gamma$ on N and is denoted by $f_{\gamma\alpha}$. For other charts $\mathcal{U}_\beta$ on M and $\mathcal{U}_\delta$ on N we obtain $f_{\delta\beta} = \varphi_{\delta\gamma} \circ f_{\gamma\alpha} \circ \varphi_{\alpha\beta}$. Recall that the changes of coordinates $\varphi_{\delta\gamma}$ and $\varphi_{\alpha\beta}$ are smooth. Thus, if $f_{\gamma\alpha}$ is smooth at the point $\varphi_\alpha^{-1}(m)$, $f_{\delta\beta}$ is smooth at the point $\varphi_\beta^{-1}(m)$.

A particular case of the above definition is a mapping from an interval of $\mathbb{R}$ to a manifold, called a *curve* (or *path*) in the manifold. Thus, the notion of a smooth curve is well-defined.

A *submanifold* M' of a manifold M is a subset of M that is a manifold with the property that each point of M' belongs to a chart $\mathcal{U}$ of M such that for the corresponding ball $\mathcal{V}$ of $\mathcal{U}$ (where $\mathcal{V}$ is a ball in a space E, say) the intersection $\mathcal{U} \bigcap M'$ corresponds to an open ball of a linear subspace of E.

For certain manifolds M the charts for some points are not open balls but "half-balls", i.e., intersections of open balls with a closed half-space. Such points form the *boundary* of M, which is usually denoted by the symbol ∂M.

If two copies of a manifold M with boundary ∂M are pasted together in such a way that identical points of ∂M are pairwise identified and so that the resulting manifold has no boundary, the latter manifold is called the *double* of the manifold M. In the double of M the boundary ∂M is a submanifold of codimension 1.

The case where the model space of a manifold is infinite-dimensional has some special features. First of all the model space is assumed to be a Hilbert or a Banach space since in more general infinite-dimensional spaces the notion of differentiability is not completely well-defined. Manifolds with Hilbert or Banach model spaces are respectively called *Hilbert* or *Banach manifolds*.

In addition, an infinite-dimensional manifold is not assumed to satisfy the second countability axiom: it is a well-known fact that a topological space satisfying the second countability axiom is separable, while many manifolds arising in applications have non-separable model spaces. Usually (but not necessarily) the manifolds are assumed to be paracompact.

A *scalar field* f on a manifold M is a mapping from M to a vector space $\mathbb{R}^k$, $f : M \to \mathbb{R}^k$. If this map is continuous (smooth), it is called a *continuous* (*smooth*, respectively) *scalar field*. A *scalar* is a value of a certain scalar field at some point $m \in M$, but usually (when it does not lead to confusion) a scalar field also is called a scalar. Note that according to this convention a scalar may have dimension greater than 1 (this corresponds to the use of the term scalar by physicists).

The characteristic feature of any scalar $f(m)$ is that its presentation in a chart does not transform under coordinate changes, unlike all other fields (see below).

For the sake of future applications we should also give another presentation of scalar fields. Consider the direct product $M \times \mathbb{R}^k$ and denote by π the projection onto M. Then a scalar field can be considered as a map $f : M \to M \times \mathbb{R}^k$ such that for each $m \in M$ the relation $\pi(f(m)) = m$ holds. This relation is often expressed in operator form: $\pi \circ f = \mathrm{id}$ where id is the identity map.

The tangent space to M at $m \in M$ is denoted by $T_m M$ and the total tangent bundle by TM. By $\pi : TM \to M$ we denote the natural projection.

In every $T_m M$ there is a basis generated by coordinates $(q^1, \dots, q^n)$ in a chart $\mathcal{U}_\alpha \ni m$. This basis is denoted by $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}$, where the vector $\frac{\partial}{\partial q^i}$ is the derivative of the i-th coordinate axis passing through m, with respect to the corresponding parameter q^i. Thus every tangent vector $X \in T_m M$ is represented in $\mathcal{U}_\alpha \ni m$ with coordinates $(q^1, \dots, q^n)$ as $X_\alpha = X^i \frac{\partial}{\partial q^i}$ and in $\mathcal{U}_\beta \ni m$ with coordinates $(q^{1'}, \dots, q^{n'})$ as $X_\beta = X^{i'} \frac{\partial}{\partial q^{i'}}$. Note that the indices of vectors are subscripts while those of vector coordinates are superscripts.

The relation between X_α and X_β is described by the formula

$$X_\beta = \varphi'_{\beta\alpha} X_\alpha, \text{ i.e., in coordinates } X^{i'} = \frac{\mathrm{d}q^{i'}}{\mathrm{d}q^j} X^j \tag{1.1}$$

where $\varphi'_{\beta\alpha}$ is the Jacobi matrix of $\varphi_{\beta\alpha}$. Formula (1.1) is the transformation rule for vectors under a change of coordinates.

Given the basis $\frac{\partial}{\partial q^i}$ in $T_m M$, we can create a coordinate system in $T_m M$ with respect to this basis. Denote by $\dot{q}^i$ the coordinate corresponding to $\frac{\partial}{\partial q^i}$ and consider its coordinate axis passing through a certain point $X \in T_m M$. Consider the tangent space $T_X T_m M$ to $T_m M$ at X. The coordinate system $(\dot{q}^1, \dots, \dot{q}^n)$ in $T_m M$ generates in $T_X T_m M$ the basis $\frac{\partial}{\partial \dot{q}^1}, \dots, \frac{\partial}{\partial \dot{q}^n}$ by complete analogy with the procedure that creates the basis $\frac{\partial}{\partial q^i}$ in $T_m M$ from the coordinate system $(q^1, \dots, q^n)$ in $\mathcal{U}_\alpha$.

Obviously the linear space $T_X T_m M$ is naturally isomorphic to the linear space $T_m M$ (visually they differ only by the location of the origin: the origin of $T_X T_m M$ is located at the point $X \in T_m M$). Denote by $\mathbf{p} : T_X T_m M \to T_m M$ the linear isomorphism between them which is defined by the rule

$$\mathbf{p}\left(\frac{\partial}{\partial \dot{q}^i}\right) = \frac{\partial}{\partial q^i}. \tag{1.2}$$

Note that the construction of $\mathbf{p}$ is valid for any linear vector space and its tangent. For example, consider the space $\mathbb{R}^k$ and (using the previous notation) denote a basis in it by $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}$. Then at a given point $X \in \mathbb{R}^k$ create the basis $\frac{\partial}{\partial \dot{q}^1}, \ldots, \frac{\partial}{\partial \dot{q}^n}$ in $T_X\mathbb{R}^k$ as described above. Now $\mathbf{p} : T_X\mathbb{R}^k \to \mathbb{R}^k$ is well-defined by formula (1.2).

Consider a chart $\mathcal{U}_\alpha$ on M and denote by $T\mathcal{U}_\alpha$ the restriction of the tangent bundle to $\mathcal{U}_\alpha$. Note that $T\mathcal{U}_\alpha$ can be presented as the direct product $T\mathcal{U}_\alpha = \mathcal{U}_\alpha \times \mathbb{R}^n$ since every point of $T\mathcal{U}_\alpha$ can be described by $2n$ coordinates: n coordinates of the point in $\mathcal{U}_\alpha$ and n coordinates of the vector with respect to the basis $\frac{\partial}{\partial q^i}$. Such a presentation as a direct product is called a *trivialization*, in the case under consideration, with respect to coordinates $(q^1, \ldots, q^n)$. If we choose another system of coordinates in $\mathcal{U}_\alpha$, we obtain another trivialization. There exist some other types of trivializations, for example, by a field of orthonormal frames in a chart of a Riemannian manifold (see below).

Consider $T\mathcal{U}_\alpha$ as a chart in TM. The change of coordinates between $T\mathcal{U}_\alpha$ and $T\mathcal{U}_\beta$ is given as a pair:

$$(\varphi_{\beta\alpha}, \varphi'_{\beta\alpha}) : \mathcal{U}_\alpha \times \mathbb{R}^n \to \mathcal{U}_\beta \times \mathbb{R}^n \tag{1.3}$$

where $\varphi_{\beta\alpha}$ sends the coordinates of points with respect to $\mathcal{U}_\alpha$ into those of the same points with respect to $\mathcal{U}_\beta$ and $\varphi'_{\beta\alpha}$ sends coordinate columns of a vector with respect to the basis $\frac{\partial}{\partial q^i}$ into those of the same vector with respect to the basis $\frac{\partial}{\partial q^{i'}}$ according to formula (1.1).

Note that if $\varphi_{\beta\alpha}$ is a C^k-map, $\varphi'_{\beta\alpha}$ is only a C^{k-1}-map, i.e. TM is a C^{k-1}-manifold if M is C^k. However, if M is C^∞, TM is also C^∞ (recall that, unless the contrary is stated, we assume our manifolds to be C^∞).

Convention 1.3 *In what follows we denote the points of a tangent bundle in two different ways: (m, X) as a point in TM and X_m as a tangent vector to M at m. Both symbols have the same meaning and can be used interchangeably; we will favor one when it is more suitable that the other. Strictly speaking, this notation makes proper sense only in charts but for the sake of convenience we also use it in invariant language.*

Definition 1.4. A *vector field* on M is a map $X : M \to TM$ such that $\pi X = \mathrm{id}$, i.e., $\pi X(m) = m$ for each $m \in M$.

Note that for a general mapping $X : M \to TM$ the value $X(m)$ could belong to any fiber but from the relation $\pi X(m) = m$ it follows that $X(m) \in T_m M$.

Definition 1.5. We say that a vector field X is *continuous* (*smooth*) if the map $X : M \to TM$ is continuous (smooth, respectively).

Clearly the derivative of a smooth curve is a tangent vector. Let a smooth vector field X be given on M. The curve $m(t)$ described by equation

$$\dot{m}(t) = X(m(t)). \tag{1.4}$$

is called the *integral curve* of X.

In a given chart, (1.4) can be considered as a differential equation in vector space. Hence, we can apply the theory of ordinary differential equations to (1.4). In particular, since the right-hand side of (1.4) is smooth, we obtain from the classical existence theorem that for any point $m_0 \in M$ there exist a real number $\varepsilon_{m_0} > 0$ and a unique solution $m(t)$ of (1.4) with the initial condition $m(0) = m_0$ that is well-defined for $t \in [0, \varepsilon_{m_0})$. In order to investigate all solutions of (1.4) we denote by $g_t(m_0)$ the solution such that $g_0(m_0) = m_0$. From the theorem on smooth dependence of a solution on initial values and parameters it follows that $g_t(m_0) : U \to M$ is jointly smooth in t and $m_0 \in M$, where U is some neighborhood of $0 \times M$ in $\mathbb{R} \times M$. From the uniqueness theorem for solutions of ordinary differential equations it follows that $g_t(\cdot)$ is a diffeomorphism for all t such that $g_t(m)$ exists for all $m \in M$.

Definition 1.6. $g_t(\cdot)$ is called the *general solution* of (1.4) or the *flow* of vecotr field X .

Let $f : M \to R$ be a real-valued function on M. Specify a point $m \in M$ and denote by $m(t)$ the integral curve of X such that $m(0) = m$. Consider the restriction of f onto $m(t)$, i.e. the function $f(m(t))$. Note that $f(m(t))$ is a real-valued function of a real argument.

Definition 1.7. $\frac{\mathrm{d}}{\mathrm{d}t} f(m(t))_{|t=0}$ is called the *derivative of f in the direction of X at m*. Having found the derivative of f along X at all points of M, we obtain a new function on M, denoted by Xf, that is called the *derivative of f in the direction of X*.

Note that applying the above construction to the derivative of f in the direction of a vector field $\frac{\partial}{\partial q^i}$, we obtain the partial derivative $\frac{\partial f}{\partial q^i}$. This is the reason for denoting this vector field by the partial derivative symbol.

Let X be a vector field on M and in a chart $\mathfrak{U}_\alpha$ let it be presented in coordinate form as $X = X^i \frac{\partial}{\partial q^i}$. Then we easily obtain the following formula for Xf in $\mathfrak{U}_\alpha$:

$$Xf = X^i \frac{\partial f}{\partial q^i} \tag{1.5}$$

Let X and Y be smooth vector fields on M.

Definition 1.8. The *Lie bracket* $[X, Y]$ of X and Y is the vector field on M such that for any smooth real-valued function f on M its derivative along $[X, Y]$ is given by the formula: $[X, Y]f = X(Yf) - Y(Xf)$.

Usually the definition of $[X, Y]$ is given in operator form as follows:

$$[X, Y] = X \circ Y - Y \circ X. \tag{1.6}$$

By direct calculation one can easily show that the vector field $[X, Y]$ exists, is unique and, in local coordinates, is described by the formula

$$[X,Y] = \left(X^l \frac{\partial Y^k}{\partial q^l} - Y^l \frac{\partial X^k}{\partial q^l} \right) \frac{\partial}{\partial q^k}. \tag{1.7}$$

The Lie bracket is evidently skew-symmetric: $[X,Y] = -[Y,X]$.

Proposition 1.9 *The Lie bracket satisfies the so-called* Jacobi identity*:*

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0. \tag{1.8}$$

Consider a smooth mapping of manifolds $F : M \to N$. Specify a point $m \in M$ and its image $F(m) \in N$. The mapping F generates a mapping $\mathrm{d}_m F : T_m M \to T_{F(m)} N$ that in a chart at m and in a chart at $F(m)$ is described by the Jacobi matrix of F at m. We call $\mathrm{d}_m F$ the *differential of F at $m \in M$*.

Definition 1.10. The tangent mapping $TF : TM \to TN$ is defined by the formula

$$TF(m,X) = (F(m), \mathrm{d}_m F(X)).$$

So, the tangent mapping TF is defined globally as a mapping of tangent bundles, and the differential $\mathrm{d}_m F$ of F at m is a restriction of TF to $T_m M$.

Note that for a smooth curve $m(t)$ in M we have the relation

$$\frac{\mathrm{d}}{\mathrm{d}t} F(m(t)) = TF \left(\frac{\mathrm{d}}{\mathrm{d}t} m(t) \right), \tag{1.9}$$

which can be considered as a coordinate-free definition of TF.

If F sends M into a linear space, say, $F : M \to \mathbb{R}^k$, the construction of the differential can be modified so that it becomes a map from TM to $\mathbb{R}^k$. We introduce it as the composition

$$\mathrm{d}F = \mathbf{p} \circ \mathrm{d}_m F : T_m M \to \mathbb{R}^k \tag{1.10}$$

where $\mathbf{p}$ is defined in (1.2). Note that $\mathrm{d}F$ can be applied to vectors from any $T_m M,\ m \in M$, so it is well-defined on the entire tangent bundle TM.

Definition 1.11. A *cotangent vector* (which we also call a *covector* or *1-form*) b at $m \in M$ is a linear functional on the tangent space $T_m M$, i.e. a linear map $b : T_m M \to \mathbb{R}$. The set of all covectors at m is called the *cotangent space* at m and is denoted by $T_m^* M$.

By definition, $T_m^* M$ is a linear space, dual to $T_m M$. The total cotangent bundle is denoted by $T^* M$.

In every chart $\mathcal{U}_\alpha$ the coordinates $(q^1, \ldots, q^n)$ generate the basis $\mathrm{d}q^1, \ldots, \mathrm{d}q^n$ in every $T_m^* M$, dual to the basis $\frac{\partial}{\partial q^1}, \ldots, \frac{\partial}{\partial q^n}$ (i.e., $\mathrm{d}q^i(\frac{\partial}{\partial q^j}) = \delta_j^i$ where δ_j^i equals 1 if $i = j$ and 0 otherwise). Thus every covector a in $\mathcal{U}_\alpha$ has coordinate representation $a^\alpha = a_i \mathrm{d}q^i$ (note that the indices of covectors are superscripts while the indices for their coordinates are subscripts). The relation between such presentations a^α and a^β in the charts $\mathcal{U}_\alpha$ and $\mathcal{U}_\beta$ is given by the formula

$$a^\beta = a^\alpha \circ (\varphi'_{\beta\alpha})^{-1}, \text{ i.e., in coordinates } a_{j'} = \frac{\partial q^i}{\partial q^{j'}} a_i. \tag{1.11}$$

Note that in (1.11) the Jacobi matrix $\varphi'_{\beta\alpha}$ is applied to the row a^α from the right while in (1.1) it is applied to the column X_α from the left. Having transposed both sides of (1.11) we obtain (1.11) in the form similar to (1.1):

$$a^T_\beta = [(\varphi'_{\beta\alpha})^{-1}]^T \circ a^T_\alpha. \tag{1.12}$$

Proposition 1.12 *The transformation rules* (1.1) *and* (1.11) *coincide if and only if* $\varphi'_{\beta\alpha}$ *is an orthogonal matrix.*

Indeed, if and only if $\varphi'_{\beta\alpha}$ is orthogonal, we have $[(\varphi'_{\beta\alpha})^{-1}]^T = \varphi'_{\beta\alpha}$ by a characteristic property of orthogonal operators.

Recall that the velocity of a curve, in particular of the trajectory of a mechanical particle, is a tangent vector. The momentum p of the trajectory is introduced in elementary text books by the property that the inner product of p and the velocity v is equal to the kinetic energy multiplied by 2:

$$p \cdot v = 2\mathcal{K}. \tag{1.13}$$

Note that in fact p is a covector. Indeed the kinetic energy $\mathcal{K}$ is a scalar, i.e. it takes the same value in all charts and coordinate systems. So, since v is a vector (transforming under changes of coordinates by (1.1)), $p \cdot v$ in (1.13) can take the same value in all coordinate systems if and only if p transforms by (1.11), i.e., if p is a covector.

Another physical example of a covector is *force*. Indeed, the inner product of a force f and velocity v is a scalar known as the *power* $\mathcal{N}$:

$$f \cdot v = \mathcal{N}.$$

As in (1.13), since v transforms according to (1.1) and $\mathcal{N}$ has the same value in all coordinate systems, f must transform according to (1.11), i.e., it is a covector.

Note that in elementary text books only motion in $\mathbb{R}^3$ with orthonormal coordinate systems is considered. Thus, only orthogonal coordinate changes are used and so the vectors v and covectors p and f have the same transformation rules (see above). This is not the case for systems given on manifolds with arbitrary coordinate systems.

As for TM, T^*M inherits a manifold structure from M: the charts on T^*M are of the form $T^*\mathcal{U}_\alpha$, which can be represented as $\mathcal{U}_\alpha \times \mathbb{R}^n$, and changes of coordinates are of the form $(\varphi_{\beta\alpha}, ((\varphi'_{\beta\alpha})^{-1})^T)$ (formula (1.12) is in use instead of (1.1)). The points of T^*M we shall denote either by (m, b) (as a point of T^*M) or, equivalently, as b_m (covector b at the point m).

As for TM we denote by the symbol $\pi : T^*M \to M$ the projection which sends $(m, b) \in T^*M$ into $m \in M$.

Definition 1.13. A *covector field* on M is a map $b : M \to T^*M$ such that $\pi \circ b = \mathrm{id}$ (cf. Definition 1.4).

A covector field is called *continuous* (*smooth*) if b as a map of manifolds is continuous (smooth, respectively).

Since $\mathbb{R}$ is a linear space, the construction of the differential $\mathrm{d}f$ according to (1.10) is well-defined for any given real-valued function $f : M \to R$. Here we are to emphasize the following:

Proposition 1.14 *$\mathrm{d}f$ is a covector field on M.*

Indeed, by the construction, $\mathrm{d}f$ is a map from TM into $\mathbb{R}$ which is linear (coinciding with $\mathrm{d}_m f$) on any tangent space $T_m M$. This means that at any $m \in M$ the map $\mathrm{d}f$ is a linear functional on $T_m M$, i.e. a covector.

In a given chart $\mathcal{U}_\alpha$, $\mathrm{d}f$ can be expressed in terms of coordinates with respect to a basis $\mathrm{d}q^1, \dots \mathrm{d}q^n$. To do this we should calculate the Jacobi matrix for f in coordinates $q^1, \dots q^n$. We then find that it takes the form $(\frac{\partial f}{\partial q^1}, \dots, \frac{\partial f}{\partial q^n})$ and hence

$$\mathrm{d}f = \frac{\partial f}{\partial q^i} \mathrm{d}q^i. \tag{1.14}$$

Remark 1.15. Note that in classical vector analysis $(\frac{\partial f}{\partial q^1}, \dots, \frac{\partial f}{\partial q^n})$ is known as the gradient of the function f while (1.14) has the form of a total differential. We should point out that both formulas describe (in different forms) the object we have called the differential of f. The geometrically well-defined definition of gradient is given in the next chapter in such a way that in Euclidean n-space it turns out to be a vector (i.e., a coordinate column, not a row) with coordinates $\frac{\partial f}{\partial q^i}$.

Taking into account (1.14), we obtain for $X = X^i \frac{\partial}{\partial q^i}$ that

$$\mathrm{d}f(X) = X^i \frac{\partial f}{\partial q^i}. \tag{1.15}$$

Comparing (1.5) with (1.15) we obtain that the equality

$$Xf = \mathrm{d}f(X) \tag{1.16}$$

holds for any X and any f.

Note that (1.16) is sometimes used as the definition of $\mathrm{d}f$.

Above we have described the notion of a tangent mapping of tangent bundles generated by a smooth mapping of manifolds. A natural mapping of a cotangent bundle is also generated but, unlike the tangent mapping, it sends T^*N to T^*M. Let M and N be smooth finite-dimensional manifolds and let

$F : M \to N$ be a smooth map. Specify a point $m \in M$ and consider its image $F(m) \in N$. Let $a \in T^*_{F(m)}N$ be a covector at $F(m)$. It can be mapped into $T^*_m M$ by the following procedure. Denote by $T^*F(a)$ the covector in $T^*_m M$ such that its value on any vector $X \in T_m M$ is defined by the relation

$$T^*F(a)(X) = a(TF(X)) \tag{1.17}$$

where TF is the tangent mapping. By the construction of TF (see Definition 1.10), we get that $TF(X) \in T_{F(m)}N$ and so the value $a(TF(X))$ is well-defined.

Definition 1.16. The map $T^*F : T^*N \to T^*M$ is called the *cotangent mapping* of $F : M \to N$ or the *pull-back*.

1.2 Lie Groups and Lie Algebras

Definition 1.17. A manifold G is called a *Lie group* if there exists an algebraic operation $\bullet$ on G such that G is a group with respect to $\bullet$ and $g_1 \bullet g_2$ is jointly smooth in $g_1, g_2 \in G$ as a map from $G \times G$ to G.

The first examples of Lie groups are the space $\mathbb{R}^n$ (which is obviously a commutative Lie group with respect to addition) and the circle S^1, i.e. the set of complex numbers with unit modulus, with respect to multiplication (also commutative).

It is clear that the n-dimensional torus $\mathcal{T}^n = S^1 \times \cdots \times S^1$ (the cartesian product of n copies of S^1) is a Lie group. The group operation on $\mathcal{T}^n$ is given by coordinate-wise multiplication.

Remark 1.18. $\mathcal{T}^n$ may also be described as the quotient space $\mathbb{R}^n/\mathbb{Z}^n$ of $\mathbb{R}^n$ with respect to the integral lattice $\mathbb{Z}^n$. This means that in $\mathbb{R}^n$ the points whose corresponding coordinates differ from one other by an integer are considered as equivalent and are pasted onto each other. It is clear that S^1 is obtained from the corresponding coordinate axis in $\mathbb{R}^n$ by factorization with respect to integer points.

Let us turn to non-commutative groups. First we mention the three-dimensional sphere S^3 (the set of points with unit modulus in the space of quaternions) and the sphere S^7 (the set of points with unit modulus in the space of octaves).

Consider the group of real invertible $n \times n$ matrices with respect to matrix multiplication. We denote this group, called the *general linear group*, by $GL(n, \mathbb{R})$.

Theorem 1.19 *$GL(n, \mathbb{R})$ is a Lie group.*

The closed set $\det^{-1}(0) \in L(n,\mathbb{R})$ subdivides $GL(n,\mathbb{R})$ into two connected components: the matrices with positive determinants and those with negative determinants. Since the determinant of a product equals the product of the determinants, the product of two matrices with positive determinants has positive determinant, i.e., this connected component is a Lie sub-group of $GL(n,\mathbb{R})$.

Another Lie sub-group of $GL(n,\mathbb{R})$ is the group $O(n)$ of orthogonal $n \times n$ matrices. Recall that a matrix $\mathcal{A}$ is *orthogonal* if its action on vectors preserves the inner product in $\mathbb{R}^n$ or, equivalently, if $\mathcal{A}^t = \mathcal{A}^{-1}$ where $\mathcal{A}^t$ is the transposed matrix. $O(n)$ is a regular surface (i.e., a submanifold) in $GL(n,\mathbb{R})$.

Like $GL(n,\mathbb{R})$, $O(n)$ has two connected components: the orthogonal matrices with determinant $+1$ and those with determinant -1. The former is a Lie sub-group of $O(n)$ (the product of matrices with unit determinant has unit determinant). This group is called the *special orthogonal group* and is denoted by $SO(n)$. The matrices in $SO(n)$ preserve the space orientation while the matrices with determinant -1 reverse it. The latter set of matrices is not a subgroup of $O(n)$.

Definition 1.20. A *left action* of a Lie group G on a manifold M is defined if a certain C^∞-map $G \times M \to M$, denoted for $g \in G$ and $m \in M$ by gm, is given such that the following hypotheses hold:

(i) for any $g \in G$ the map $g : M \to M$ that sends m to gm is a diffeomorphism;

(ii) $(g \bullet h)m = g(hm)$ for $g, h \in G$, $m \in M$.

A *right action* of a Lie group G on a manifold M is the specification of a certain C^∞-map from $M \times G$ to G, for $g \in G$ and $m \in M$, which satisfies (i) and the following replacement of (ii):

(iii) $(g \bullet h)m = h(gm)$ for $g, h \in G$, $m \in M$.

When a right action is given, the notation mg for $g \in G$, $m \in M$ is used so that $m(g \bullet h) = (mg)h$.

In what follows we shall denote the unit of a Lie group G by e.

For $g \in G$ two special maps, the left translation $L_g : G \to G$ and the right translation $R_g : G \to G$, are defined by the formulae $L_g h = g \bullet h$ and $R_g h = h \bullet g$, respectively, for any $h \in G$. From Definition 1.17 it follows that both L_g and R_g are smooth maps of G.

Note that the tangent bundle TG is trivial, i.e., it can be presented as direct product. Indeed, having taken a certain basis in T_eG we can translate it to T_gG at each point $g \in G$ by TL_g (or TR_g) and so obtain the presentation of TG as $G \times \mathbb{R}^n$, where $n = \dim G$.

Definition 1.21. The vector field on G obtained by left (right) translations of a vector $X \in T_eG$ at all points of G is called the *left-invariant* (*right-invariant*, respectively) *vector field, generated by X*.

It should be noted that one can imitate the construction of left- and right-invariant vector fields and define left- and right-invariant Riemannian metrics and more general tensors (see the definitions of these object below).

Proposition 1.22 *Let $\bar{X}$ and $\bar{Y}$ be left-invariant (right-invariant) vector fields on G generated by $X, Y \in T_eG$. Then the vector field $[\bar{X}, \bar{Y}]$ is left-invariant (right-invariant, respectively).*

See, e.g., [26] for the proof of Proposition 1.22.

Denote by $[X, Y]$ the vector in T_eG which generates $[\bar{X}, \bar{Y}]$.

Definition 1.23. $[X, Y]$ is called the *bracket* of X, Y.

Proposition 1.24 *The bracket in T_eG introduced in Definition 1.23 satisfies the Jacobi identity (1.8).*

The assertion of Proposition 1.24 follows from Propositions 1.9 and 1.22.

Definition 1.25. A linear space on which an additional operation $[\cdot, \cdot]$ satisfying the Jacobi identity is given is called a *Lie algebra*.

Thus, by Proposition 1.24, T_eG has the structure of a Lie algebra.

Definition 1.26. The vector space T_eG, equipped with the bracket defined in Definition 1.23, is called the *Lie algebra of the Lie Group G*.

We generally denote Lie groups by Latin capitals (say, G) and their Lie algebras by the corresponding lower case Fraktur characters (say, $\mathfrak{g}$).

It is known that every finite dimensional Lie algebra is the Lie algebra of a certain Lie group, however different Lie groups may have the same Lie algebra. For example, a group and a subgroup of the same dimension have the same algebra.

Let us present some examples of Lie algebras.

On every linear space one can introduce the *trivial bracket* defined by the equality $[X, Y] = 0$ for every X and Y. Thus every linear space is a (trivial) Lie algebra. It is clear that the algebras of the groups $\mathbb{R}^n$, S^1 and $\mathcal{T}^n$ are trivial, as well as the algebras of all commutative groups.

Consider Euclidean space $\mathbb{R}^3$ together with the skew-symmetric vector product operation. This operation is usually denoted by the bracket symbol, a notation which we retain, i.e., for two vectors $X, Y \in \mathbb{R}^3$ the symbol $[X, Y]$ denotes their vector product. By direct calculation one can easily prove the following:

Proposition 1.27 *The vector product operation satisfies the Jacobi identity.*

Thus, $\mathbb{R}^3$ with the vector product is a Lie algebra.

The Lie algebra $\mathfrak{gl}(n, R)$ is the set of all $n \times n$ matrices with linear addition in the underlying vector space and with the bracket $[A, B] = AB - BA$ (commutator of matrices).

The groups $O(n)$ and $SO(n)$ generate the same Lie algebra (indeed, the unit of $O(n)$ is contained in the subgroup $SO(n)$). This algebra is denoted by $\mathfrak{so}(n)$. It consists of skew-symmetric $n \times n$ matrices with the same bracket as in $\mathfrak{gl}(n, R)$ ($\mathfrak{so}(n)$ is a Lie subalgebra of $\mathfrak{gl}(n, R)$).

A particular case, the algebra $\mathfrak{so}(3)$, is of special interest to us because of its use, below, in some examples of mechanical systems. The matrices from $\mathfrak{so}(3)$ have the form

$$\begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}. \tag{1.18}$$

Denote by $\Psi : \mathfrak{so}(3) \to \mathbb{R}^3$ the mapping that sends the matrix (1.18) to the vector $\begin{pmatrix} -a \\ -b \\ -c \end{pmatrix}$. The next statement is obvious.

Proposition 1.28 *Ψ is a linear isomorphism of $\mathfrak{so}(3)$ to $\mathbb{R}^3$.*

Proposition 1.29 *For every $A, B \in \mathfrak{so}(3)$ the operator Ψ sends the commutator $AB - BA$ to the vector product of $\Psi(A)$ and $\Psi(B)$ in $\mathbb{R}^3$.*

We introduce on $\mathfrak{so}(3)$ a bilinear form by the formula $(A, B) = -\frac{1}{2}\mathrm{tr}(AB)$. This bilinear form is called the *Killing form*.

Proposition 1.30 *$-\frac{1}{2}\mathrm{tr}(AB)$ equals the inner product of the vectors $\Psi(A)$ and $\Psi(B)$ in $\mathbb{R}^3$.*

Propositions 1.29 and 1.30 are proved by direct calculation with matrices.

Thus $-\frac{1}{2}\mathrm{tr}(AB)$ is an inner product in $\mathfrak{so}(3)$ and so we have introduced the structure of Euclidean space in $\mathfrak{so}(3)$. From Propositions 1.29 and 1.30 we obtain:

Proposition 1.31 *The linear isomorphism $\Psi : \mathfrak{so}(3) \to \mathbb{R}^3$ is an isomorphism of Lie algebras and of Euclidean spaces.*

1.3 Fiber Bundles

Definition 1.32. A *fiber bundle* comprises the following five objects:

(i) a manifold M, called the *base* of the bundle;
(ii) a manifold E, called the *total space* of the bundle;
(iii) a manifold F, called the *standard fiber* of the bundle;
(iv) a Lie group G, called the *structure group* of the bundle;
(v) a smooth projection $\pi : E \to M$, called the *projection* of the bundle,

and the following interrelations between these objects hold:

1) a left action of G on F is given (see Definition 1.20);
2) for any chart $\mathcal{U}_\alpha$ in M the set $\pi^{-1}\mathcal{U}_\alpha$ is homeomorphic to $\mathcal{U}_\alpha \times F$;
3) for $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \neq \emptyset$ the "change of coordinates" from $\mathcal{U}_\alpha \times F$ to $\mathcal{U}_\beta \times F$ is given as the pair $(\varphi_{\beta\alpha}, g_{\beta\alpha}(m))$ where $\varphi_{\beta\alpha} : \mathcal{U}_{\alpha\beta} \to \mathcal{U}_{\alpha\beta}$ is defined above and $g_{\beta\alpha}(m) : F \to F$ is an element of G, smooth in $m \in \mathcal{U}_{\alpha\beta}$, which is a diffeomorphism of F according to the definition of the left action of G.

Usually we denote a fiber bundle by the symbol of its total space, say E, and in this case E_m denotes the fiber at the base point m (note that $E_m = \pi^{-1}(m)$ is homeomorphic to F). In order to indicate the details of a certain fiber bundle we may say that there is a bundle E over M or that $\pi : E \to M$ is a fiber bundle. To denote the bundle with fiber F, a convenient notation is $\underline{F}$.

Strictly speaking $\mathcal{U}_\alpha \times F$ and $\mathcal{U}_\beta \times F$ are not charts, but $\mathcal{U}_\alpha \times F$ and $\mathcal{U}_\beta \times F$ contain the charts in E of the form $\mathcal{U}_\alpha \times V_\kappa$ and $\mathcal{U}_\beta \times V_\lambda$ (where V_κ and V_λ are charts in F). Here, the restrictions of $(\varphi_{\beta\alpha}, g_{\beta\alpha})$ become real changes of coordinates.

The simplest example of a fiber bundle is a trivial (or product) bundle $E = M \times F$. Here all $g_{\beta\alpha}$ are equal to the unique element $e = \mathrm{id}$ (unit) in G (and so the group G may be reduced to its subgroup consisting of the unique element e). The elements of any product bundle are scalars (see above). Recall that under changes of coordinates in M, scalars are not transformed at all.

Note that according to Definition 1.32 every bundle over each chart $\mathcal{U}_\alpha$ is presented as a trivial one by means of a certain diffeomorphism $\mathcal{F}_\alpha$ that is called a *trivialization* (it is often said that over each chart the bundles are trivial or that all bundles by means of Definition 1.32 are locally trivial). It is important to understand from the very beginning that many different trivializations of a bundle may exist even over a specified chart.

Definition 1.33. A *vector bundle* is a fiber bundle where $F = \mathbb{R}^k$ for a certain k, G is the group $GL(k, \mathbb{R})$ of non-degenerate linear operators in $\mathbb{R}^k$, or a subgroup thereof, and the elements of G act on $\mathbb{R}^k$ as linear automorphisms.

The product bundles with $F = \mathbb{R}^k$ are examples of vector bundles. Two more examples are the tangent and the cotangent bundles of a manifold. Indeed, for the tangent bundle, M is the base, TM is the total space, the fiber is $\mathbb{R}^n$ (assuming that $\dim M = n$), $G = GL(n, \mathbb{R})$ with the natural action on $\mathbb{R}^n$ and $g_{\alpha\beta}(m) = \varphi'_{\alpha\beta}(m)$ (see (1.3)). For the cotangent bundle, M is also the base, T^*M is the total space, the fiber is $\mathbb{R}^n$ and G is also $GL(n, \mathbb{R})$ with the same natural action on $\mathbb{R}^n$ but $g_{\alpha\beta}$ now takes the form $[(\varphi'_{\alpha\beta})^{-1}]^T$ (see (1.12)).

Of course, in the infinite-dimensional case Definition 1.33 is modified by replacing $\mathbb{R}^k$ with some Hilbert or Banach space and $G = GL(n, \mathbb{R})$ by the group of bounded invertible linear operators.

Definition 1.34. A *principal bundle* is a fiber bundle where $F = G$.

It should be pointed out that on a principal bundle E both the left and the right actions of its structure group G on fibers are well-defined, since any fiber of E is isomorphic to G. In particular, having specified $g \in G$, we can consider the right translation $\eta \circ g$ of each point $\eta \in E$. Thus we obtain the right action of G on E (see Definition 1.20).

Every principal bundle generates the so-called associated bundle as follows. Let E be a principal bundle with structure group G, and let F be a manifold on which a left action of G is given. Consider the quotient of the direct product $E \times F$ with respect to the right action of G defined by the formula $(\eta, f)g = (\eta \circ g, g^{-1}f)$ (note that this is indeed a right action since $(gh)^{-1} = h^{-1}g^{-1}$). Consider a chart $\mathfrak{U}_\alpha$ on M. The restriction of $E \times F$ on $\mathfrak{U}_\alpha$ has the form $\mathfrak{U}_\alpha \times G \times F$. The above-mentioned quotient space consists of the orbits of elements of this product: for $(m, h, f) \in \mathfrak{U}_\alpha \times G \times F$ the orbit is the set $(m, hg, g^{-1}f)$ for all $g \in G$. Such an orbit is associated with the point $(m, hf) \in \mathfrak{U}_\alpha \times F$. Indeed, for any point of the orbit we have $(m, (hg)(g^{-1}f)) = (m, hf)$ and so this association is well-defined. Thus, over $\mathfrak{U}_\alpha$, the above quotient space is presented in the form $\mathfrak{U}_\alpha \times F$ and the total quotient space is a bundle over M with fiber F and structure group G where the $g_{\alpha\beta}$ are the same as in E.

Definition 1.35. The above quotient is called the *bundle associated to E with fiber F*.

Notation 1.36 *Denote by λ the mapping of $E \times F$ onto the total space of the associated bundle that is the factorization sending the orbits to the corresponding points in the quotient.*

It should be noted that every non-principal fiber bundle is associated to some principal bundle.

Let Q be a vector bundle with fiber $\mathbb{R}^k$ and let it be an associated bundle to a principal bundle E with $G = GL(k, \mathbb{R})$. Observe that, clearly, any $b \in E$ can be considered as a frame $b = e_1, \dots, e_k$ in the fiber $Q_{\pi b}$ of Q through πb and, consequently, as a linear map $b : \mathbb{R}^k \to Q_{\pi b}$ sending $x = (x^1, \dots, x^k)$ to $bx = x^1 e_1 + \dots + x^k e_k$.

Definition 1.37. The *frame bundle BM* of a manifold M is a principal bundle over M with $G = GL(n, \mathbb{R})$ ($n = \dim M$) and $g_{\beta\alpha} = \varphi'_{\beta\alpha}$.

Note that TM and T^*M are bundles associated to BM, and so a point $b \in BM$ may be regarded as a frame $b = e_1, \dots, e_n$ in the tangent space $T_{\pi b}M$. Thus b can also be considered as a linear mapping $b : \mathbb{R}^n \to T_{\pi b}M$ which sends a vector $x = (x^1, \dots, x^n) \in \mathbb{R}^n$ to the vector $bx = x^1 e_1 + \dots + x^n e_n \in T_{\pi b}M$.

Definition 1.38. A *cross-section X* of a fiber bundle E is a map $X : M \to E$ such that $\pi \circ X = \mathrm{id} : M \to M$, i.e. $X(m) \in E_m$ for any $m \in M$.

Examples of cross-sections are vector fields (cross-sections of a tangent bundle) and covector fields (cross-sections of a cotangent bundle). This immediately follows from Definitions 1.38, 1.4 and 1.13. Every vector bundle E

has a smooth cross-section, called the *zero-section*, which sends $m \in M$ into the origin of E_m. Unlike vector bundles, a principal bundle has a global (i.e., defined on the entire manifold M) continuous cross-section if and only if the bundle is trivial.

Definition 1.39. If the tangent bundle of a manifold M is trivial, M is called *parallelizable* or *trivializable*.

As mentioned above, all Lie groups are parallelizable by left- or right-invariant vector fields. In particular, the n-dimensional torus is parallelizable.

Remark 1.40. A natural trivialization Φ of the tangent bundle $T\mathcal{T}^n$ can be described as a trivialization by coordinate frames as follows. Introduce on S^1_i the angle coordinate $q^i = \frac{\varphi}{2\pi}$ where φ is an angle. We then obtain the system of angle coordinates $(q^1, \dots, q^n)$ on $\mathcal{T}^n$. At each point $m \in \mathcal{T}^n$ the vectors $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}$ form an orthonormal frame in $T_m\mathcal{T}^n$. For $(X^i \frac{\partial}{\partial q^i})_m$ we set $\Phi(X^i \frac{\partial}{\partial q^i})_m = (m, (X^1, \dots, X^n)) \in \mathcal{T}^n \times \mathbb{R}^n$.

Definition 1.41. If in each tangent space T_mM to a manifold M a k-dimensional linear subspace β_m is chosen that smoothly depends on $m \in M$ (in the sense described below), we say that a *k-dimensional distribution* β is given on M.

Smooth dependence of β_m on m means that *in some neighborhood of each point $m \in M$ there are k smooth linearly independent vector fields $X_1, \dots, X_k$ such that at each m' in this neighborhood the space $\beta_{m'}$ is the linear span of the vectors $X_1(m'), \dots, X_k(m')$.*

If on M there exists a smooth vector field nowhere equal to zero, the straight lines spanned on the vectors of this field form a 1-dimensional distribution.

Definition 1.42. A k-dimensional distribution β is said to be *integrable* if for each point m there exists a k-dimensional submanifold $M' \ni m$ such that at every point $m' \in M'$ the property $T_{m'}M' = \beta_{m'}$ is fulfilled. In this case the manifold M' is called an *integral manifold* of the distribution β.

Every 1-dimensional distribution is integrable. Its integral manifolds are integral curves of the vector field on which the distribution is spanned. Distributions of dimension greater than 1 may not be integrable. There is a necessary and sufficient condition for integrability that involves the notion of involutory distributions.

Definition 1.43. A distribution β is called *involutory* if for every two vector fields X and Y on M such that $X_m, Y_m \in \beta_m$ at each $m \in M$, their Lie bracket $[X, Y]$ also belongs to the space of distributions at each point of M.

Theorem 1.44 (Frobenius' Theorem) *A distribution is integrable if and only if it is involutory.*

A proof of Theorem 1.44 can be found, for example, in [26, 172, 212].

1.4 Riemannian and Semi-Riemannian Metrics

Definition 1.45. We say that a *Riemannian metric* is given on a manifold M if in each tangent space $T_m M$, $m \in M$, a symmetric positive-definite bilinear form $\langle \cdot, \cdot \rangle_m$ is specified which depends smoothly on m (in the sense defined below). A manifold with a Riemannian metric is called a *Riemannian manifold.*

When we say that $\langle \cdot, \cdot \rangle_m$ *depends smoothly* on m we mean that for any two smooth vector fields X and Y on M the real-valued function $\langle X_m, Y_m \rangle_m$ on M is smooth in m.

For a Riemannian manifold every tangent space $T_m M$ becomes a Euclidean space with inner product $\langle \cdot, \cdot \rangle_m$.

In what follows, if it does not lead to confusion, we shall omit the point m in the inner product notation. Indeed, $\langle X, Y \rangle$ obviously has only one interpretation: X and Y lie in the same tangent space and they are to be substituted into the inner product of that space.

The first example of a Riemannian metric is the so-called *first fundamental form* of a surface in Euclidean space. Recall that for $X, Y \in T_m M$, where M is a surface in a Euclidean space E, the number $I(X, Y)$ (the value of the first fundamental form I at X and Y) is defined by the equality: $I(X, Y) = (X, Y)$ where (X, Y) is the inner product in E. Note that, in spite of the fact that the unique inner product in E is used to determine the inner products in tangent spaces, the latter are still specific to their own spaces: we still cannot identify different tangent spaces and so cannot say whether the inner products are the same.

An embedding $i : M \to N$ of a Riemannian manifold M into another Riemannian manifold N is *isometrical* if $\langle X, Y \rangle_M = \langle iX, iY \rangle_N$ for every pair X, Y of vectors belonging to the same tangent space of M. Note that an embedded manifold with the first fundamental form is isometrically embedded into Euclidean space.

Theorem 1.46 (Nash [186]) *Every n-dimensional Riemannian manifold can be isometrically embedded into a Euclidean space $\mathbb{R}^N$ with $N = \frac{1}{2}n(n+1)(3n+11)$.*

In a chart $\mathcal{U}_\alpha$ the form $\langle \cdot, \cdot \rangle_m$ can be described in terms of its matrix. As is typical in linear algebra, we consider the coefficients of $\langle \cdot, \cdot \rangle_m$ with respect to the basis $\frac{\partial}{\partial q^k}$ defined by the formula $g_{ij} = \langle \frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j} \rangle$. Note that the coefficients g_{ij} depend on $m \in \mathcal{U}_\alpha$ and so they are real-valued functions on $\mathcal{U}_\alpha$.

The linearity of $\langle \cdot, \cdot \rangle$ allows us to derive a coordinate formula for the inner product of vectors involving the coefficients g_{ij}. For vectors $X = X^i \frac{\partial}{\partial q^i}$ and $Y = Y^i \frac{\partial}{\partial q^j}$ we obviously have $\langle X, Y \rangle = g_{ij} X^i Y^j$. Note that this notation can be applied both for vectors X and Y at a given point and for vector fields on

$\mathcal{U}_\alpha$. In the latter case both the coefficients g_{ij} and the coordinates X^i and Y^j are functions on $\mathcal{U}_\alpha$.

One can easily prove that $\langle\cdot,\cdot\rangle_m$ is smooth in $m \in \mathcal{U}_\alpha$ if and only if all of the functions $g_{ij}(m)$ are smooth.

Using a Riemannian metric one can create natural analogs of many notions of ordinary geometry in Euclidean space.

For a vector $X \in T_m M$ we define its norm $\|X\|$ by the formula $\|X\| = \sqrt{\langle X, X\rangle}$. For two vectors $X, Y \in T_m M$ denote by θ the angle between them. Then $\cos\theta = \frac{\langle X,Y\rangle}{\|X\|\cdot\|Y\|}$. Let $m(t)$, $t \in [a,b]$, be a curve in M. Its length s is defined by the formula

$$s|_a^b = \int_a^b \|\dot{m}(t)\| \mathrm{d}t. \tag{1.19}$$

Note that all the notions introduced above, after embedding M isometrically into a Euclidean space, coincide with the usual definitions of norm, angle and length, respectively, in that space.

Let $m_0, m_1 \in M$. Denote by $\aleph$ the total set of piecewise smooth curves in M connecting m_0 and m_1.

Definition 1.47. The infimum of the lengths of the curves in $\aleph$ is called the *Riemannian* (or *internal*) *distance* in M between m_0 and m_1 and is denoted by $\rho(m_0, m_1)$.

Proposition 1.48 *The distance $\rho(\cdot,\cdot)$ satisfies the axioms of a metric:*

(i) $\rho(m,m) = 0$ *and from* $\rho(m_0,m_1) = 0$ *it follows that* $m_0 = m_1$*;*
(ii) $\rho(m_0,m_1) = \rho(m_1,m_0)$ *for every pair* $m_0, m_1 \in M$*;*
(iii) $\rho(m_0,m_1) + \rho(m_1,m_2) \geq \rho(m_0,m_2)$ *for any* $m_0, m_1, m_2 \in M$*.*

The proof of Proposition 1.48 can be found, for example, in [26].

Hence, from Proposition 1.48 it follows that a Riemannian manifold is a metric space with respect to its Riemannian distance ρ.

Definition 1.49. The Riemannian metric is called *complete* if the metric space M with distance ρ generated by the Riemannian metric is complete. In this case one says that M is a *complete Riemannian manifold* .

Recall that a bilinear form is called *non-degenerate* if its matrix (g_{ij}) is not degenerate, i.e., it is invertible. Since the bilinear form of a Riemannian metric is positive-definite it follows that the form is non-degenerate.

Definition 1.50. We say that a *semi-Riemannian metric* is given on a manifold M if in each tangent space $T_m M$, $m \in M$, a symmetric non-degenerate bilinear form $\langle\cdot,\cdot\rangle_m$ is specified which depends smoothly on m. A manifold with a semi-Riemannian metric is called a *semi-Riemannian manifold.*

So, a Riemannian metric is a particular case of a semi-Riemannian metric. For a semi-Riemannian metric a scalar square $\langle X, X\rangle$ may be negative or even zero for non-zero X, hence neither norm nor distance are well-defined on semi-Riemannian manifolds. Nevertheless, semi-Riemannian metrics appear naturally in many physical theories. The main example for us is the appearance of semi-Riemannian metrics in relativity theory.

Notation 1.51 *For a Riemannian or semi-Riemannian metric with matrix (g_{ij}) in a certain chart, we denote the inverse matrix $(g_{ij})^{-1}$ by (g^{ij}), i.e., its coefficients are denoted by g^{ij}.*

Remark 1.52. One can easily show that (g^{ij}) is the matrix of an inner product in the cotangent space at the corresponding point. Both the latter field of inner products in cotangent spaces and the Riemannian metric (i.e., the field of inner products in tangent spaces described by matrices (g_{ij})) are traditionally called "metric tensors".

Real mechanical and physical objects can be described both as vectors and as covectors which are said to be physically equivalent. The set up of a physical problem usually involves a certain Riemannian or semi-Riemannian metric on a manifold. These metrics determine the physical equivalence as follows.

Definition 1.53. For $X \in T_m M$ the *physically equivalent covector* $b_X \in T^*_m M$ is defined by the relation $b_X(Y) = \langle X, Y \rangle$ for any $Y \in T_m M$.

For $b \in T^*_m M$ the *physically equivalent vector* $X_b \in T_m M$ is defined by the relation $b(Y) = \langle X_b, Y \rangle$ for any $Y \in T_m M$.

From Definition 1.53 it follows that for any $X \in T_m M$ the physically equivalent covector b_X exists and is unique. Moreover, one can easily calculate its coordinates. Let $X = X^i \frac{\partial}{\partial \mathrm{d}^i}$ and $Y = Y^j \frac{\partial}{\partial \mathrm{d}^j}$. Denote by X_j the coordinates of b_X, i.e., $b_X = X_j \mathrm{d}q^j$. Then, since $b(Y) = X_j Y^j$ and $\langle X, Y \rangle = g_{ij} X^i Y^j$ for any coordinate column Y^j, we obtain from the definition

$$X_j = g_{ij} X^i. \tag{1.20}$$

Applying (1.20) to the matrix (g^{ij}), the inverse of the matrix of the Riemannian metric, one can easily see that

$$X^i = g^{ij} X_j. \tag{1.21}$$

Remark 1.54. Note that here we use the existence of the matrix $(g^{ij}) = (g_{ij})^{-1}$, and do not use positive-definiteness. Thus physical equivalence is well-defined for the general case of a semi-Riemannian metric, not only for a Riemannian one.

Definition 1.55. The vector physically equivalent to the differential $\mathrm{d}f$ of a function f (see (1.14)) is called the *gradient* of f and is denoted by $\mathrm{grad} f$.

So, the gradient is determined by the relation $\mathrm{d}f(X) = \langle \mathrm{grad} f, X\rangle$ for any X. Comparing this formula with (1.16) we obtain the following equality which holds for all X and f:

$$Xf = \mathrm{d}f(X) = \langle \mathrm{grad} f, X\rangle. \tag{1.22}$$

Convention 1.56 *In what follows we shall denote the coordinates of physically equivalent vectors and covectors by the same character, using upper indices for vector coordinates and lower indices for covector coordinates as in formulae* (1.20) *and* (1.21)*.*

In mechanics and physics (1.20) and (1.21) are usually called 'formulae of lifting and lowering indices'. The term "physical equivalence" is suggested in [200].

Let M be a Riemannian manifold. If in some $T_m M$, $m \in \mathcal{U}_\alpha$, we have

$$g_{ij} = \delta_{ij} = \begin{cases} 1 & \text{if } \ i = j \\ 0 & \text{if } \ i \neq j \end{cases},$$

i.e., the basis $\frac{\partial}{\partial q^i}$ consists of orthonormal vectors, the coordinates of physically equivalent vectors and covectors coincide. In particular, and this is true only in this case, $\mathrm{d}q^i$ and $\frac{\partial}{\partial q^i}$ are physically equivalent. However, a coordinate system satisfying $g_{ij} = \delta_{ij}$ can be created only in a Euclidean space. In a general Riemannian manifold one can easily create a coordinate system satisfying the above equality at a single point, but the relation will generally not be satisfied at other points in its neighborhood. For semi-Riemannian manifolds the situation is quite analogous, the only modification being: if $\frac{\partial}{\partial q^i}$ is an orthonormal basis, $g_{ii} = \pm 1$.

1.5 Tensors

Main definitions

In order to describe many physical and mechanical notions one needs to use mathematical objects more general than vectors and covectors. These objects are called tensors and, in common with vectors and covectors, they are characterized by the rules for transformation of their components under changes of coordinates. Moreover, vectors and covectors turn out to be trivial particular cases of tensors.

Let $m \in M$ and consider the Cartesian product of r copies of $T^*_m M$ and s copies of $T_m M$.

Definition 1.57. A *tensor of type* (r, s) (or (r, s)-*tensor*) at the point m is a polylinear form on the above Cartesian product.

This means that a tensor T of type (r, s) is a real-valued function, linear in all its arguments, such that for any collection of r covectors $a, b, \dots, c$ and s vectors $X, Y, \dots, Z$ the real value $\mathsf{T}(a, b, \dots, c, X, Y, \dots, Z)$ is defined.

A *tensor field* on M is defined if we associate to each point of M a tensor.

Definition 1.58. The total number of arguments in a tensor T is called its *valency*. The number of covector arguments is called the *contravariant rank* of T and the number of vector arguments is called the *covariant rank* of T.

Scalars are tensors of valency 0. Tensors of valency 1 are vectors and covectors. By definition, a vector is a tensor of contravariant rank 1 and a covector is a tensor of covariant rank 1.

There exists a construction allowing one to create tensors of higher valency from vectors and covectors. Consider vectors $E_1, \dots, E_r$ and covectors $p^1, \dots, p^s$. Define the tensor $\mathsf{T} = E_1 \otimes \dots \otimes E_r \otimes p^1 \otimes \dots \otimes p^s$ of type (r, s) by the equality

$$\begin{aligned} &E_1 \otimes \dots \otimes E_r \otimes p^1 \otimes \dots \otimes p^s (a^1, \dots, a^r, X_1, \dots, X_s) \\ &= E_1(a^1) \dots E_r(a^r) p^1(X_1) \dots p^s(X_s) \end{aligned} \tag{1.23}$$

for every multiplicity of covectors $a^1, \dots, a^r$ and vectors $X_1, \dots, X_s$.

Definition 1.59. A tensor of the form (1.23) is called the *tensor product* of vectors $E_1, \dots,\ E_r$ and covectors $p^1, \dots, p^s$. Tensors of this type are called *elementary.*

The space of tensors of type (r, s) at a point m is obviously a linear space. It is clear that the elementary tensors

$$\frac{\partial}{\partial q^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial q^{i_r}} \otimes \mathrm{d}q^{j_1} \otimes \dots \otimes \mathrm{d}q^{j_s}, \tag{1.24}$$

where the indices $i_1, \dots, i_r, j_1, \dots, j_s$ take all values from 1 to n (we assume an n-dimensional manifold M), form a basis for this space.

For any (r, s)-tensor T, its coordinates with respect to the basis (1.24) are denoted by $\mathsf{T}^{i_1 \dots i_r}_{j_1 \dots j_s}$ so that

$$\mathsf{T} = \mathsf{T}^{i_1 \dots i_r}_{j_1 \dots j_s} \frac{\partial}{\partial q^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial q^{i_r}} \otimes \mathrm{d}q^{j_1} \otimes \dots \otimes \mathrm{d}q^{j_s}. \tag{1.25}$$

In order to avoid any confusion (we usually assume that coordinates appear with respect to a basis in $T_m^* M$ or in $T_m M$), we call $\mathsf{T}^{i_1 \dots i_r}_{j_1 \dots j_s}$ the *components* of the tensor T.

Note that, for the sake of simplicity, the summands of (1.25) are ordered so that all $\frac{\partial}{\partial q^{i_r}}$ precede all $\mathrm{d}q^{j_s}$ in the tensor products (and consequently, in the components all terms with upper indices precede all terms with lower indices). In arbitrary tensors the factors of types $\frac{\partial}{\partial q^{i_r}}$ and $\mathrm{d}q^{j_s}$ (and component terms with upper and lower indices) may appear in any order.

Taking into account formulae (1.1) and (1.11), one can easily derive the following formula for the transformation of components under a change of coordinates:

$$\mathsf{T}^{i'_1\dots i'_r}_{j'_1\dots j'_s} = \frac{\mathrm{d}q^{i'_1}}{\mathrm{d}q^{i_1}}\cdots\frac{\mathrm{d}q^{i'_r}}{\mathrm{d}q^{i_r}}\,\frac{\mathrm{d}q^{j_1}}{\mathrm{d}q^{j'_1}}\cdots\frac{\mathrm{d}q^{j_s}}{\mathrm{d}q^{j'_s}}\mathsf{T}^{i_1\dots i_r}_{j_1\dots j_s} \tag{1.26}$$

Note that all upper indices are transformed as in (1.1) and all lower indices as in (1.11).

Formula (1.26) is the main characteristic of tensors. The above definition of a tensor as a polylinear form is convenient only from the general point of view. For concrete tensors, the fact that they can be presented as polylinear forms may not be important for a given physical problem. On the other hand, transformation rule (1.26) distinguishes tensors from all other objects.

In analogy with the constructions of tangent and cotangent bundles we define the (r, s)-*tensor bundle* over M as the set of all (r, s)-tensors at all points of M and define a smooth manifold structure on it as follows. The sets of tensors over the charts $\mathcal{U}_\alpha$ and $\mathcal{U}_\beta$ in M play the roles of charts in our bundle (note that using (1.25) one can easily see that these new charts are presented as direct products) and the changes of coordinates are given in the form $(\varphi_{\beta\alpha}, g_{\beta\alpha})$ where $\varphi_{\beta\alpha}$ is the change of coordinate in M and $g_{\beta\alpha}$ is defined by (1.26). Now one can easily give the definition of an (r, s)-*tensor field* as a cross-section of a tensor bundle according to Definition 1.38 (i.e., in analogy with Definitions 1.4 and 1.13).

It is clear that a Riemannian metric is an example of a $(0, 2)$-tensor, i.e., in a chart, $\langle\cdot,\cdot\rangle = g_{ij}\mathrm{d}q^i \otimes \mathrm{d}q^j$. The matrix (g^{ij}), the inverse of the matrix (g_{ij}) of a Riemannian metric in a certain coordinate system, is an example of a $(2, 0)$-tensor, i.e., in a chart this metric tensor takes the form $g^{ij}\frac{\partial}{\partial q^i} \otimes \frac{\partial}{\partial q^i}$. Thus, in any $T^*_m M$ this matrix describes the inner product of covectors, dual to the inner product on vectors in $T_m M$ (the Riemannian inner product) with respect to physical equivalence.

It appears that $(1, 1)$-tensors are linear operators. Indeed, consider an elementary tensor $E \otimes p$ at some $m \in M$ and substitute into p a certain vector $X \in T_m M$. Then $E \otimes p(\cdot, X) = p(X)E \in T_m M$ is vector linearly dependent on X, i.e. $E \otimes p : T_m M \to T_m M$ is a linear operator. Thus the tensor $\mathsf{T} = a^i_j \frac{\partial}{\partial q^i} \otimes \mathrm{d}q^j$ is the linear operator in $T_m M$ with matrix (a^i_j).

Note that we can substitute a covector $a \in T^*_m M$ into E in $E \otimes p$ so that $E \otimes p(a, \cdot) = E(a)p \in T^*_m M$, i.e. $E \otimes p$ can be considered as a linear operator acting on $T^*_m M$. Hence $a^i_j \frac{\partial}{\partial q^i} \otimes \mathrm{d}q^j$ can also be considered as a linear operator on $T^*_m M$. This operator is known as the *dual* (or *conjugate*) operator to the operator on $T_m M$, mentioned above, with the same matrix (a^i_j).

Operations with tensors

The set of all (r, s)-tensors at a point $m \in M$ form a linear space, i.e., addition and multiplication (by a real number) are well-defined. From this it follows that for tensor fields the operations of addition and of multiplication by a real-valued function are also well-defined.

Further operations will be first determined for elementary tensors. Since every tensor is a sum of elementary tensors (see expansion (1.25)), we can easily extend the operations from elementary to all tensors.

The *tensor product* of an (r, s)-tensor A and a (k, l)-tensor B creates a new $(r+k, s+l)$-tensor $\mathsf{A} \otimes \mathsf{B}$ as follows. For $\mathsf{A} = E_1 \otimes \cdots \otimes E_r \otimes p^1 \otimes \cdots \otimes p^s$ and $\mathsf{B} = \bar{E}_1 \otimes \cdots \otimes \bar{E}_k \otimes \bar{p}^1 \otimes \cdots \otimes \bar{p}^l$ we have $\mathsf{A} \otimes \mathsf{B} = E_1 \otimes \cdots \otimes E_r \otimes p^1 \otimes \cdots \otimes p^S \otimes \bar{E}_1 \otimes \cdots \otimes \bar{E}_k \otimes \bar{p}^1 \otimes \cdots \otimes \bar{p}^l$. For arbitrary tensors T and $\bar{\mathsf{T}}$, expanded according to (1.25), we multiply by this rule each summand of T with each summand of $\bar{\mathsf{T}}$ and sum up all the products to get the expansion of $\mathsf{T} \otimes \bar{\mathsf{T}}$. This means that the set of components for $\mathsf{T} \otimes \bar{\mathsf{T}}$ is obtained by taking the product of all components of T with all components of $\bar{\mathsf{T}}$.

A *contraction* (or *trace*) transforms an (r, s)-tensor into an $(r-1, s-1)$-tensor as follows. For $\mathsf{A} = E_1 \otimes \cdots \otimes E_r \otimes p^1 \otimes \cdots \otimes p^s$ its contraction by $l-th$ lower and $v-th$ upper factors is the tensor of the form

$$p^v(E_l)E_1 \otimes \cdots \otimes E_{l-1} \otimes E_{l+1} \ldots E_r \otimes p^1 \otimes \cdots \otimes p^{v-1} \otimes p^{v+1} \otimes \cdots \otimes p^s.$$

Applying this rule to a general tensor T with expansion (1.25) and taking into account that $\mathrm{d}q^v(\frac{\partial}{\partial q^l}) = \delta^v_l$, we see that the contraction causes a lot of summands in (1.25) to vanish. The remaining non-zero summands have the form $\mathsf{T}^{i_1 \ldots i_{l-1}\, k\, i_{l+1} \ldots i_r}_{j_1 \ldots j_{v-1}\, k\, j_{v+1} \ldots j_s}$ (the sum with respect to $k = 1, \ldots, n$).

For example, consider the contraction of a $(1, 1)$-tensor (linear operator) $\mathsf{T} = a^i_j \frac{\partial}{\partial q^i} \otimes \mathrm{d}q^j$ (see above). Evidently we get $\mathrm{tr}\mathsf{T} = a^k_k$, the ordinary trace of the matrix of the linear operator T.

Physically equivalent tensors

Let M be a Riemannian manifold. For an elementary tensor we define a physically equivalent tensor by replacing a certain vector (or covector) in the corresponding tensor product by the physically equivalent covector (vector, respectively).

When transforming an arbitrary tensor with expansion (1.25) into a physically equivalent one, we should remember that generally speaking $\frac{\partial}{\partial q^i}$ is not physically equivalent to $\mathrm{d}q^i$. This is why, on replacing a certain $\frac{\partial}{\partial q^i}$ (of $\mathrm{d}q^i$) by the physically equivalent object, we should create the expansion (1.25) for the obtained tensor. Taking into account formulae (1.20) and (1.21), one can easily derive the corresponding formulae for arbitrary tensors. For the sake of simplicity we present them for tensors with three indices; the general formulae are analogous. Usually the components of physically equivalent tensors are denoted by the same character, changing only the location of indices. So,

$$\begin{aligned} \mathsf{A}^i_{jk} &= g^{il}\mathsf{A}_{ljk}; \\ \mathsf{B}^{ik}_j &= g_{lj}\mathsf{B}^{ilk}; \\ \mathsf{C}^{ij}_k &= g^{iu}g^{jv}\mathsf{C}_{uvk}, \end{aligned}$$

and so on.

Remark 1.60. Note that the two versions of a metric tensor, i.e., matrices (g_{ij}) and (g^{ij}) (see Remark 1.52), are physically equivalent to each other.

Symmetric tensors

We say that an $(r, 0)$- or $(0, s)$-tensor is *symmetric* or *skew-symmetric* if the polylinear form describing the tensor is respectively symmetric or skew-symmetric. A symmetric form is a form whose value does not depend on the order of its arguments while the sign of a skew-symmetric form changes whenever two of its arguments are interchanged.

We consider $(0, s)$-tensors since the case of $(r, 0)$-tensors is analogous. One can easily see that for a symmetric tensor, if the basis tensors in expansion (1.25) differ only by the order of their factors, the corresponding components are equal. In order to simplify the notation we introduce the notion of a *symmetric tensor product*:

$$\mathrm{d}q^{i_1} \odot \cdots \odot \mathrm{d}q^{i_s} = \frac{1}{s!}\sum_{1}^{s!} \mathrm{d}q^{j_1} \otimes \cdots \otimes \mathrm{d}q^{j_s}$$

where the summands $\mathrm{d}q^{j_1} \otimes \cdots \otimes \mathrm{d}q^{j_s}$ differ only by the order of factors (it is obvious that the number of such summands is $s!$). Thus, any symmetric tensor has the expansion:

$$\mathsf{T} = \mathsf{T}_{i_1 \ldots i_s} \mathrm{d}q^{i_1} \odot \cdots \odot \mathrm{d}q^{i_s}$$

where the indices $i_1, \ldots, i_s$ are written in increasing order. One can also easily show that the tensors $\mathrm{d}q^{i_1} \odot \cdots \odot \mathrm{d}q^{i_s}$ form a basis in the space of all symmetric $(0, s)$-tensors.

1.6 Differential Forms and Polyvectors

A skew-symmetric $(0, s)$-tensor given at a point on a manifold is called an *exterior s-form* and the corresponding tensor field is called a *differential s-form*; the valency of these tensors is called the *degree* of the form.

Skew-symmetric $(r, 0)$-tensors are called *r-vectors*. A vector field in which all the tensors are r-vectors is sometimes called an *r-vector field*. If one needn't indicate the valency, we simply refer to an *exterior* (*differential*) *form* or *polyvector* (*polyvector field*). Often one also omits the words "exterior" or "differential", saying "s-form".

Real-valued functions are included in this terminology as 0-forms, covectors as 1-forms and vectors as 1-vectors. Note that the term "1-form" is far more commonly used than the word "covector". In what follows we shall generally refer to 1-forms, only rarely using the term "covector". Similarly, the term "1-vector" is rarely used, so we shall generally refer to vectors.

Clearly there are two parallel theories for s-forms and r-vectors. Our main focus shall be on s-forms with the understanding that analogous results hold for r-vectors. Only when we need to use both will polyvectors appear.

Exterior product

As for symmetric tensors, as described above, there exists a certain special reduction of the tensor product which is very useful when dealing with s-forms. It is called the *exterior product*.

Let $\tilde{a}^1, \tilde{a}^2 \dots \tilde{a}^l$ be 1-forms. We introduce the following formal matrix with 1-form coefficients

$$A = \begin{pmatrix} \tilde{a}^1\tilde{a}^2 \dots \tilde{a}^l \\ \tilde{a}^1\tilde{a}^2 \dots \tilde{a}^l \\ \dots\dots\dots\dots \\ \tilde{a}^1\tilde{a}^2 \dots \tilde{a}^l \end{pmatrix}. \tag{1.27}$$

Just as for a matrix with real-valued coefficients, for (1.27) one can consider its determinant $\det A$ where the ordinary product of real numbers is replaced by the tensor product of 1-forms. Since the tensor product is not commutative (unlike the ordinary numerical product) this yields a non-trivial expression. For a permutation σ of $\{1, \dots, n\}$, the coefficient of the summand $\tilde{a}^{\sigma(1)}\tilde{a}^{\sigma(2)} \cdots \tilde{a}^{\sigma(n)}$ in this determinant is equal to the sign of σ.

Definition 1.61.

$$\tilde{a}^1 \wedge \tilde{a}^2 \wedge \cdots \wedge \tilde{a}^l = \det A.$$

Immediately from Definition 1.61 we see that $\tilde{a}^1 \wedge \tilde{a}^2 \wedge \cdots \wedge \tilde{a}^l$ is a skew-symmetric $(0, l)$-tensor; changing the order of any two forms in the product results in a change of the sign. It follows that if there are at least two equal factors in the product, the product is equal to zero (since the product remains unchanged if we swap two such identical factors).

For two 1-forms a and b, by Definition 1.61, we obtain $a \wedge b = (a \otimes b - b \otimes a)$.

Remark 1.62. Note that there are $l!$ summands in the determinant of (1.27). Thus it is tempting to insert the factor $\frac{1}{l!}$ before $\det A$ in the definition of the exterior product, imitating the definition of the symmetric tensor product in Section 1.5. We emphasize that we do not do so since it is rather convenient to omit it in the formulae below. It should be pointed out that that the factor $\frac{1}{l!}$ does appear in some text-books. This leads to some changes in consequent formulae. So, if you use formulae from different books, you should check what definitions of exterior product the authors use in order to avoid running into absurdities of the sort $1 = 0$.

From the definition of the exterior product we can easily see that any s-form can be presented as a linear combination of exterior products of basis 1-forms (covectors) $\mathrm{d}q^i$. The problem is that these products are not linearly independent: for example $\mathrm{d}q^1 \wedge \mathrm{d}q^2 \wedge \mathrm{d}q^3$, $\mathrm{d}q^2 \wedge \mathrm{d}q^1 \wedge \mathrm{d}q^3$ and $\mathrm{d}q^3 \wedge \mathrm{d}q^2 \wedge \mathrm{d}q^1$ differ from one another only by sign. Moreover, a product having at least two common factors is equal to zero (see the previous section). In order to avoid this problem we shall choose the order of factors in the exterior products of basis 1-forms according to increasing order of indices and exclude the products with equal factors. The resulting products become linearly independent and so they form the basis of the space of forms of specified degree at a point of M.

Thus any s-form at a given point of M has the expansion

$$\omega = \omega_{i_1 \dots i_s} \mathrm{d}q^{i_1} \wedge \dots \wedge \mathrm{d}q^{i_s} \tag{1.28}$$

where $i_1 < \dots < i_s$.

The space of all s-forms at a point $m \in M$ is denoted by $\wedge^s T_m^* M$ and the bundle of all s-forms by $\wedge^s T^* M$. We can now calculate the dimension of $\wedge^s T_m^* M$ for the n-dimensional manifold M. By the construction it is equal to the binomial coefficient C_n^s. Note that $C_n^s = C_n^{n-s}$. In particular this means that $\wedge^s T_m^* M$ is isomorphic to $\wedge^{n-s} T_m^* M$ since the dimensions of these finite-dimensional linear spaces coincide. Unfortunately, as usual, there are many isomorphisms between these spaces but none can be considered as canonical – the best one for all problems. Later we shall find a candidate for the 'best' isomorphism on a Riemannian manifold with some additional property.

Theorem 1.63 $\dim \wedge^n T_m^* M = 1$, *and hence* $\wedge^n T_m^* M$ *is isomorphic to the space of* 0*-forms at* m*, i.e., to* $\mathbb{R}^1$.

Indeed, $C_n^1 = C_n^n = 1$. The unique basis form in $\wedge^n T_m^* M$ is $\mathrm{d}q^1 \wedge \mathrm{d}q^2 \wedge \dots \wedge \mathrm{d}q^n$.

We say that an n-form ω at some point m is *identically zero* if its value is zero on any multiplicity of n vectors from $T_m M$.

Corollary 1.64

(i) *Let* $\omega_1,\ \omega_2 \in \wedge^n T_m^* M$ *and* ω_1 *be not identically zero. Then* $\omega_2 = \lambda \omega_1$ *for some real number* λ.

(ii) *Let* $\omega_1,\ \omega_2$ *be two differential* n*-forms on* M *and* ω_1 *be nowhere identically zero on* M*. Then* $\omega_2 = \lambda(m) \omega_1$ *where* $\lambda(m)$ *is a real-valued function on* M.

Theorem 1.65 *The forms of degree greater than* n *on a manifold with dimension* n *are identically zero.*

Indeed, if the degree is greater than n and there are only n basis 1-forms $\mathrm{d}q^i$, there will be equal factors in the exterior products of expansion (1.28). By the properties of the exterior product, all such products are equal to zero.

The exterior differential d

Recall that in Section 1.1 we defined the notion of the differential $\mathrm{d}F$ of a map F sending a certain manifold M, $\dim M = n$, to a linear space $\mathbb{R}^k$. This operation d was very useful for dealing with a real-valued function f on M (here $k = 1$). It transforms f into the differential 1-form $\mathrm{d}f$ on M (see Proposition 1.14).

Here we extend the operation d to differential forms of higher degree so that it transforms an s-form α into an $(s+1)$-form $\mathrm{d}\alpha$. This extension is called the *exterior differential*. For the sake of simplicity we introduce the exterior differential in terms of coordinates. This is probably the only exception to our general idea to first introduce a new object in invariant form and then to describe it in terms of coordinates. The invariant construction of the exterior differential is more complicated than the coordinate construction.

We define the exterior differential d on 0-forms as the operation coinciding with the above-mentioned d. For an s-form α, $s > 0$, with the expansion (1.28), i.e.,

$$\alpha = \alpha_{i_1\dots i_s}\mathrm{d}q^{i_1} \wedge \cdots \wedge \mathrm{d}q^{i_s},$$

we define the action of d by the formula

$$\mathrm{d}\alpha = \mathrm{d}\alpha_{i_1\dots i_s} \wedge \mathrm{d}q^{i_1} \wedge \cdots \wedge \mathrm{d}q^{i_s} \tag{1.29}$$

where $\mathrm{d}\alpha_{i_1\dots i_s}$ is the differential of the function $\alpha_{i_1\dots i_s}$, i.e., a 1-form that can be expanded by general formula (1.14): $\mathrm{d}\alpha_{i_1\dots i_s} = \frac{\partial\alpha_{i_1\dots i_s}}{\partial q^i}\mathrm{d}q^i$ so that (1.29) is transformed into

$$\mathrm{d}\alpha = \frac{\partial\alpha_{i_1\dots i_s}}{\partial q^i}\mathrm{d}q^i \wedge \mathrm{d}q^{i_1} \wedge \cdots \wedge \mathrm{d}q^{i_s}. \tag{1.30}$$

Note that (1.30) does not satisfy the above convention that the factors in the exterior products $\mathrm{d}q^i \wedge \mathrm{d}q^{i_1} \wedge \cdots \wedge \mathrm{d}q^{i_s}$ should be ordered according to increasing size of indices, so for each form α it needs to be rewritten accordingly. Note also that $\mathrm{d}q^i \wedge \mathrm{d}q^{i_1} \wedge \cdots \wedge \mathrm{d}q^{i_s}$ may have equal factors so that some summands in (1.30) may vanish.

Definition 1.66. If $\mathrm{d}\alpha = 0$ the form α is called *closed*. If $\alpha = \mathrm{d}\beta$ for some form β, α is called *exact*.

Theorem 1.67 $\mathrm{d}^2 = \mathrm{d} \circ \mathrm{d} = 0$, *i.e., all exact forms are closed.*

Remark 1.68. Note that in $\mathbb{R}^k$ all closed forms are exact. The same is true of any contractible manifold. In the general case the difference between closed and exact forms indicates the topological structure of the manifold. This difference is described by means of the so-called de Rham cohomologies.

Physically equivalent r-forms and r-vectors

Let M be a Riemannian manifold, i.e. a manifold M on which a Riemannian metric $\langle\cdot,\cdot\rangle = g_{ij}\mathrm{d}q^i \otimes \mathrm{d}q^j$ is given. As for ordinary tensors, in this case the notion of physical equivalence is well-defined for skew-symmetric tensors of types $(s,0)$ and $(0,s)$. For the sake of convenience we shall denote physically equivalent forms and polyvectors by the same letter putting a tilde over the form and a bar over the polyvector. For instance, $\tilde{X}$ and $\bar{X}$ denote a form and polyvector, respectively, each physically equivalent to the other.

Theorem 1.69 *Let $\tilde{\alpha} = \tilde{a}^1 \wedge \cdots \wedge \tilde{a}^k$ be a k-form where a^i are 1-forms. Then $\bar{\alpha} = \bar{a}^1 \wedge \cdots \wedge \bar{a}^k$ where $\bar{a}^i$ are the vectors physically equivalent to $\tilde{a}^i$, $i = 1,\ldots,k$.*

This theorem follows directly from the construction.

Recall that $\mathrm{d}q^i$ and $\frac{\partial}{\partial q^i}$ are generally not physically equivalent to each other. This means that the transition to the physically equivalent object should be done by means of the general formulae (1.20) and (1.21) in the same manner as in Section 1.5 for general tensors.

The interior product

This is a version of a contraction (or trace, see Section 1.5) adapted to the language of differential forms.

Consider an r-vector X and an s-form α, $r < s$. Their *interior product* is denoted by $X \rfloor \alpha$ and is defined as follows. Create the tensor product of X and α, both expanded as in (1.25). This yields an (r,s)-tensor. Now contract it with respect to all r contravariant factors and the first r covariant factors in each summand of the expansion (1.25) so that the k-th contravariant factor is contracted with the k-th covariant factor. $X \rfloor \alpha$ is the resulting tensor of type $(0, s-r)$. Since X and α are skew-symmetric, so too is $X \rfloor \alpha$, i.e., it is a differential (exterior) form. Its expansion should be found in the form (1.28).

Volume forms. Orientable manifolds

Recall (see Corollary 1.64) that on an n-dimensional manifold M for two n-forms α and β at some point $m \in M$, where β is not identically zero, we have $\alpha = \lambda\beta$ where λ is a real number. Thus for two differential forms α and β, such that β is identically zero nowhere on M, we have $\alpha = \lambda(m)\beta$ where λ is a real-valued function on M.

The problem is whether there exist a differential n-form on M which is nowhere identically zero. There are manifolds on which any n-form is identically zero at at least one of its points. Such manifolds are called *non-orientable.*

If there exist an n-form that is nowhere identically zero on M, M is called *orientable.* Obviously if a single such form exists, there must be infinitely many such forms: for example one can multiply the form by any real number or by a non-zero real-valued function on M. The set of all such n-forms on an orientable manifold M is naturally divided into two classes as follows: specify a certain set of n linearly independent vectors at a given point $m \in M$, then two n-forms belong to the same subclass if their values on the above set have the same sign. We summarize this as follows: there exist two possible orientations on an orientable manifold.

On an orientable manifold M it is convenient to specify a certain nowhere identically zero n-form. In this case we say that an *orientation* has been chosen on M and call M an *oriented manifold*. This form is called the *volume form* on the oriented manifold M.

The latter term originates from the following. Let M be an oriented Riemannian (or semi-Riemannian) manifold (as is commonly found in many physical problems). In this case there is a canonical volume form Ω constructed as follows. Choose an orientation on M, i.e., specify a certain n-form ω that is nowhere identically zero. Note that any tangent space T_mM, $m \in M$, is a Euclidean (or semi-Euclidean, respectively) space. For any set of vectors $X_1, \ldots, X_n \in T_mM$ define the value $\Omega(X_1, \ldots, X_n)$ to be the volume of the parallelepiped spanning $X_1, \ldots, X_n$, with the sign $+$ if $\omega(X_1, \ldots, X_n) > 0$ and with the sign $-$ if $\omega(X_1, \ldots, X_n) < 0$. Note that if $X_1, \ldots, X_n$ are not linearly independent, the volume of the spanning parallelepiped is equal to zero, i.e. the above construction is well-defined. Obviously Ω is skew-symmetric, i.e., it is an n-form at m. Doing this for all $m \in M$ we obtain a differential n-form that is clearly nowhere identical zero.

Definition 1.70. The above form Ω is called the *Riemannian volume form.*

In a Euclidean space with orthonormal basis $\frac{\partial}{\partial q^1}, \ldots, \frac{\partial}{\partial q^n}$ (i.e., $\mathrm{d}q^i(\frac{\partial}{\partial q^j}) = \delta^i_j$ here) we have $\Omega = \mathrm{d}q^1 \wedge, \ldots, \wedge \mathrm{d}q^n$. In a chart of a Riemannian manifold the Riemannian volume form is described by the formula (see [202])

$$\Omega = \sqrt{\det(g_{ij})}\mathrm{d}q^1 \wedge, \ldots, \wedge \mathrm{d}q^n. \tag{1.31}$$

On a manifold the integrals of real-valued functions are well-defined if the integrator is a volume form. We refer the reader, say, to [212] for details.

The operations $*$, δ and Δ

Let M be a Riemannian (or semi-Riemannian) oriented manifold with Riemannian volume form Ω.

The operation $*$ is an isomorphism of the space of k-forms onto the space of $(n-k)$-forms at any point $m \in M$. For a certain k-form $\tilde{\alpha}$ the $(n-k)$-form $*\tilde{\alpha}$ is defined by the formula:

$$* \tilde{\alpha} = \frac{1}{k!} \bar{\alpha} \rfloor \Omega \tag{1.32}$$

where $\bar{\alpha}$ is the k-vector physically equivalent to $\tilde{\alpha}$. Note that the coefficient $\frac{1}{k!}$ is involved in order to naturally coordinate the formulas with each other. For instance, one can easily see that in the Euclidean space $\mathbb{R}^3$, by formula (1.32), $*(\mathrm{d}q^1 \wedge \mathrm{d}q^2) = \mathrm{d}q^3$ while it would be equal to $2\mathrm{d}q^3$ without the coefficient.

Since for all $k = 0, \ldots, n$, $* : \wedge^k T^*M \to \wedge^{n-k} T^*M$ is an isomorphism, there exists an inverse $*^{-1} : \wedge^s T^*M \to \wedge^{n-s} T^*M$ for all $s = 0, \ldots, n$. (Note: if $s = n - k$ then $k = n - s$.)

Definition 1.71. The operator $\delta = *^{-1} \mathrm{d} *$ is called the *codifferential.*

Let α be a k-form. Then $*\alpha$ is an $(n-k)$-form, $\mathrm{d} * \alpha$ is an $(n-k+1)$-form and $\delta\alpha = *^{-1}\mathrm{d} * \alpha$ is a $(k-1)$-form.

Theorem 1.72 $\delta^2 = 0$.

Indeed, $\delta^2 = (*^{-1}\mathrm{d}*)(*^{-1}\mathrm{d}*) = *^{-1}\mathrm{d}^2* = 0$ since $\mathrm{d}^2 = 0$ by Theorem 1.67.

Definition 1.73. $\Delta = (\mathrm{d} + \delta)^2$ is called the *Laplace-de Rham operator* (or *Kodaira-Hodge Laplacian*).

Since $\mathrm{d}^2 = 0$ and $\delta^2 = 0$, one can easily calculate that

$$\Delta = \mathrm{d}\delta + \delta\mathrm{d} \tag{1.33}$$

(note that d and δ do not commute and so the summands on the right-hand side of (1.33) are not equal to each other).

If α is a k-form, then since d increases the degree by one and δ decreases the degree by one, $\Delta\alpha$ is also a k-form, i.e., Δ preserves the degree.

Relations to operators of classical vector analysis

Making use of the operations with differential forms introduced above, one can generalize the operators $\mathrm{grad} f$, $\mathrm{div} X$ and $\mathrm{rot} X$ of vector analysis in three-dimensional Euclidean space to spaces of higher dimension and to Riemannian (semi-Riemannian) manifolds.

In the remainder of this section we consider a Riemannian (or semi-Riemannian) manifold M, a smooth vector field $\bar{X}$ and a smooth real-valued function f on M. As introduced above, we use the symbols *bar* and *tilde* over a certain expression to denote physically equivalent polyvectors and forms, respectively. In particular, by $\tilde{X}$ we denote the 1-form physically equivalent to $\bar{X}$. Recall that the coordinates of $\tilde{X}$ are denoted by X_i and those of $\bar{X}$ by X^i.

Recall also the definition of the gradient of a function given in Definition 1.55. In our new notation $\mathrm{grad} f = \overline{\mathrm{d}f}$. One can easily see that in the Euclidean space $\mathbb{R}^3$, i.e., where $\overline{\mathrm{d}q^i} = \frac{\partial}{\partial q^i}$, this formula yields the usual definition of gradient from vector analysis.

Definition 1.74. The *divergence* of the vector field $\bar{X}$ is the function $\mathrm{div}\bar{X} = \delta\tilde{X}$.

By direct calculation one can easily show that in the Euclidean space $\mathbb{R}^3$ the above definition gives the ordinary divergence.

Remark 1.75. Below (see Definition 1.86) we give an alternative definition of divergence in terms of the so-called Lie derivative. It will be applicable in even broader settings (for instance, the manifold M may not be Riemannian).

Definition 1.76. The *rotation* of the vector field X is the $(n-2)$-vector field $\mathrm{rot}\bar{X} = \overline{*\mathrm{d}\tilde{X}}$.

Notice that only for $n = 3$ do we have $n - 2 = 1$, i.e., only in this case is $\mathrm{rot}\bar{X}$ a 1-vector field. As in the case of divergence, by direct calculations one can easily show that in the Euclidean space $\mathbb{R}^3$ the above definition gives the ordinary rotation.

1.7 The Lie Derivative

Here we present a method of differentiation of various objects along a vector field on a manifold. Note that even in Euclidean spaces this method, generally speaking, does not coincide with the ordinary derivative.

First we should describe the general notion of differentiating something along a smooth curve. Let Q be a vector space. Associate a copy of Q to each point m of a manifold M and denote it by Q_m. Let $q(m) \in Q_m$ be chosen so that the "field" $q(m)$ is given on M. Consider a smooth curve $m(t)$, $t \in [0, l]$. Suppose that there exists a set of linear operators $A(t) : Q_{m(t)} \to Q_{m(0)}$. Consider the curve $q(t) = A(t)q(m(t))$ in the vector space $Q_{m(0)}$. Calculate $\frac{\mathrm{d}}{\mathrm{d}t}q(t)$ at $t = 0$. This derivative (if it exists) is a vector of $Q_{m(0)}$, i.e., an object of the same sort as the field $q(m)$. Note that this manner of differentiating depends on the operator $A(t)$; having determined $A(t)$ we obtain a specific type of differentiation.

For example, if $q(m)$ is a smooth vector field on M and $A(t)$ is the operator of parallel translation, the above derivative coincides with the covariant derivative $\frac{\mathrm{D}}{\mathrm{d}t}$ (see Section 2.2).

The Lie derivative is defined according to the above scheme where the operator $A(t)$ is determined by a particular vector field on M, as we describe below.

Let a smooth vector field X be given on M. Consider its flow $g_t(m)$ (see Definition 1.6). Let for a point $m_0 \in M$ the integral curve $g_t(m_0)$ exist for $t \in [0, \varepsilon_{m_0}]$ and so for any $t \in (0, \varepsilon_{m_0})$ the neighboring integral curves also exist. Thus we can consider the map g_t on a neighborhood of m_0 and since it is jointly smooth in t and m, we can consider its tangent Tg_t and its cotangent

T^*g_t mappings. These mappings will be used as operators $A(t)$ for translating vectors and forms along $g_t(m_0)$ in the construction of Lie derivatives. Note that since g_t is a diffeomorphism, Tg_t is also one-to-one and there exists an inverse map Tg_t^{-1}.

We begin the construction of the Lie derivative with that for real-valued functions. Consider such a function $f : M \to R$. Let $m_0 \in M$ and consider the integral curve $g_t(m_0)$. Note that here the general construction is reduced since by definition all values $f(g_t(m_0))$ belong to the same real line $\mathbb{R}$ and so we needn't move them along. Thus, according to the general idea, we should create the function $f(g_t(m_0))$ and differentiate it at $t = 0$. The resulting number $\mathcal{L}_X f(m_0)$ is called the *Lie derivative of f along X at m_0*. Having done this at all points of M we obtain the function $\mathcal{L}_X f$ called the *Lie derivative of f along X*.

Proposition 1.77 *$\mathcal{L}_X f$ coincides with the ordinary derivative Xf of f along X.*

Indeed, the statement is obvious since the constructions of $\mathcal{L}_X f$ and Xf coincide (cf. Definition 1.7).

Now we generalize the construction to differential forms of higher degree. Let α be a k-form, $0 < k \leq n$. Consider the forms $\alpha_{g_t(m_0)}$ along $g_t(m_0)$. In order to move them to m_0 along $g_t(m_0)$ we should generalize the construction of the cotangent mapping (see Definition 1.16) so that it applies to k-forms. We do this for the map g_t (the general construction is quite analogous).

Denote by $T^*g_t(\alpha_{g_t(m_0)})$ the k-form at the point m_0 whose value on the k vectors $\bar{X}_1, \ldots, \bar{X}_k \in T_m M$ is given by:

$$T^*g_t(\alpha_{g_t(m_0)})(\bar{X}_1, \ldots, \bar{X}_k) = \alpha_{g_t(m_0)}(Tg_t\bar{X}_1, \ldots, Tg_t\bar{X}_k). \tag{1.34}$$

Definition 1.78. The k-form $\frac{\mathrm{d}}{\mathrm{d}t} T^*g_t(\alpha_{g_t(m_0)})_{|t=0}$ at m_0 is called the *Lie derivative of α along X at m_0* and is denoted by $\mathcal{L}_X\alpha(m_0)$. Having carried out this procedure at all points of M we obtain the k-form $\mathcal{L}_X\alpha$ on M which is called the *Lie derivative of α along X*.

Theorem 1.79 *$\mathcal{L}_X\alpha = X \rfloor \mathrm{d}\alpha + \mathrm{d}(X \rfloor \alpha)$.*

Note that both summands on the right-hand side of the last formula have degree k (this follows from the properties of d and the interior product). We leave the proof to the reader as a (not so simple) exercise (for a proof, see [212]).

The last modification of the construction deals with vector fields. Let Y be a smooth vector field on M. Consider its restriction $Y(g_t(m_0))$ to the curve $g_t(m_0)$. Now we translate those vectors into $T_{m_0}M$ by Tg_t^{-1} which by its construction sends $T_{g_t(m_0)}M$ into $T_{m_0}M$.

Definition 1.80. The vector $\frac{\mathrm{d}}{\mathrm{d}t} Tg_t^{-1}(Y(g_t(m_0))_{|t=0} \in T_{m_0}M$ is called the *Lie derivative of Y along X at m_0* and is denoted by $\mathcal{L}_X Y(m_0)$. Having carried out this procedure at all points of M we obtain the vector field $\mathcal{L}_X Y$, called the *Lie derivative of* Y *along* X.

Theorem 1.81 $\mathcal{L}_X Y = [X, Y]$, *the Lie bracket of* X *and* Y.

It is important to understand what is meant if the Lie derivative of an object along X is equal to zero. Directly from the construction of the Lie derivative we obtain:

Theorem 1.82

(i) *If* $\mathcal{L}_X f = 0$, f *is constant along integral curves of* X.
(ii) *If* $\mathcal{L}_X \alpha = 0$, $T^* g_t(\alpha_{g_t(m)}) = \alpha_m$ *for all* $m \in M$.
(iii) *If* $\mathcal{L}_X Y = 0$, $Tg_t Y_m = Y_{g_t(m)}$ *for all* $m \in M$.

Definition 1.83. If $T^* g_t(\alpha_{g_t(m)}) = \alpha_m$ $(Tg_t Y_m = Y_{g_t m})$ for all $m \in M$ we say that α (Y, respectively) is *constant along the flow of* X.

So, a zero value of a Lie derivative means that the corresponding object is constant along the flow of X. Hence, in particular, if we apply g_t to an integral curve of Y such that $[X, Y] = 0$, the image is also an integral curve of Y. Since $[X, Y] = -[Y, X]$ (see above), the same is true of the image of an integral curve of X under the action of the flow of Y.

Definition 1.84. If $[X, Y] = 0$ we say that X and Y *commute.*

An obvious example of commuting vector fields is the pair of coordinate fields $\frac{\partial}{\partial q^i}$ and $\frac{\partial}{\partial q^j}$.

Theorem 1.85 (see, e.g., [26]) *If* $[X, Y] = 0$ *in some chart, in this chart there exists a coordinate system* $(q^1, \dots, q^n)$ *such that* $X = \frac{\partial}{\partial q^1}$ *and* $Y = \frac{\partial}{\partial q^2}$.

Using the Lie derivative one can extend the notion of divergence of a vector field. Let Ω be a volume form on an orientable manifold M and let X be a smooth vector field. The Lie derivative $\mathcal{L}_X \Omega$ is also an n-form and so it is presented as the product of a real-valued function and Ω.

Definition 1.86. The function $\mathrm{div} X$ in the equality $\mathcal{L}_X \Omega = \mathrm{div} X \cdot \Omega$ is called the *divergence* of the vector field X on the manifold M with specified volume form Ω.

Let M be a Riemannian manifold.

Proposition 1.87 *If the form* Ω *in Definition* 1.86 *is the Riemannian volume form,* $\mathrm{div} X$ *from Definition* 1.86 *coincides with the divergence in Definition* 1.74.

Proof. By Theorem 1.79 $\mathcal{L}_X \Omega = X \rfloor \mathrm{d}\Omega + \mathrm{d}(X \rfloor \Omega)$. Since Ω is an n-form, $\mathrm{d}\Omega = 0$ and so $\mathcal{L}_X \Omega = \mathrm{d}(X \rfloor \Omega)$. But $\mathrm{d}(X \rfloor \Omega) = \mathrm{d}(*\tilde{X})$ (see (1.32)). On the other hand, by construction $\mathrm{div} X$ from Definition 1.86 equals $*^{-1} \mathcal{L}_X \Omega$. Hence $\mathrm{div} X = *^{-1} \mathrm{d} * \tilde{X} = \delta \tilde{X}$ (see Definition 1.71). □

Chapter 2
Connections

2.1 The Structure of a Tangent Bundle to a Vector Bundle

Let $\pi : \Theta \to M$ be a vector bundle with standard fiber $\mathbb{R}^d$, $\dim M = n$. Denote by Θ_m the fiber at $m \in M$ and by $(m, \vartheta) = \vartheta_m$ the points of this fiber. Consider a chart $\mathcal{U}_\alpha$ on a manifold M with local coordinates $(q^1, \dots, q^n)$ and a trivialization $\mathcal{F}_\alpha$ of the bundle over that chart. Let $e_1, \dots, e_d$ be the standard basis in $\mathbb{R}^d$. Since $\mathcal{F}_\alpha(\pi^{-1}\mathcal{U}_\alpha) = \mathcal{U}_\alpha \times \mathbb{R}^d$, this basis generates a basis in every fiber Θ_m, $m \in \mathcal{U}_\alpha$. We obtain a smooth field of bases that will also be denoted by $e_1, \dots, e_d$. Thus every cross-section ϑ of the bundle Θ can be represented in terms of coordinates with respect to these bases in the form $\vartheta = \vartheta^i e_i$, $i = 1, \dots, d$. In $\mathcal{U}_\alpha \times \mathbb{R}^d$ the set of vectors $\mathcal{U}_\alpha \times \{X_0\}$ for some $X_0 \in \mathbb{R}^d$ corresponds to the vectors ϑ from Θ_m, $m \in \mathcal{U}_\alpha$ that have the same coordinates with respect to $e_1, \dots, e_d$ as X_0. Another trivialization of the bundle over $\mathcal{U}_\alpha$ would generate another set of vectors equivalent to X_0 that is different from the former.

In the vector bundle the set $\mathcal{F}_\alpha(\pi^{-1}\mathcal{U}_\alpha) = \mathcal{U}_\alpha \times \mathbb{R}^d$ can be considered as a chart on the total space Θ. Denote by ϑ^i the coordinates in the fibers Θ_m, $m \in \mathcal{U}_\alpha$, whose coordinate axes are spanned by the basis vectors e_i in the fibers. We obtain the coordinate system $(q^1, ..., q^n, \vartheta^1, ..., \vartheta^d)$ in the chart $\mathcal{U}_\alpha \times \mathbb{R}^d$ on Θ. By a general scheme (see Section 1.1) this system generates a basis $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}, \frac{\partial}{\partial \vartheta^1}, \dots, \frac{\partial}{\partial \vartheta^d}$ in the tangent space $T_{(m,\vartheta)}\Theta$ to the total space Θ of the bundle at every point (m, ϑ), $m \in \mathcal{U}_\alpha$.

The symbols $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}$ for the first "half" of the basis vectors coincide with the symbols for the basis vectors $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}$ in the tangent space $T_m M$, $m \in \mathcal{U}_\alpha$, to M generated by coordinates $(q^1, \dots, q^n)$. This is natural since the former vectors are tangent to a submanifold $\mathcal{U}_\alpha \times \{V\}$ in $\mathcal{F}_\alpha(\pi^{-1}\mathcal{U}_\alpha)$ where the coordinates $(q^1, \dots, q^n)$ are generated by the projection of the same coordinates from $\mathcal{U}_\alpha$, and that projection is an isomorphism. On the

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_2, © Springer-Verlag London Limited 2011

other hand, by construction the vectors $\frac{\partial}{\partial\vartheta^1}, \dots, \frac{\partial}{\partial\vartheta^d}$ are tangent to the fiber Θ_m, i.e., they belong to the tangent space $T_\vartheta\Theta_m$.

Definition 2.1. The tangent space $T_\vartheta\Theta_m$ to the fiber of a bundle Θ at m is called the *vertical subspace* in the tangent space $T_{(m,\vartheta)}\Theta$ to the total space of the bundle and is denoted by $\mathsf{V}_{(m,\vartheta)}$. The vectors of $\mathsf{V}_{(m,\vartheta)}$ are said to be *vertical*.

For every vector $Y_{(m,\vartheta)}$ at a point $(m,\vartheta) \in \Theta$ we can find its coordinate decomposition with respect to the basis mentioned above: $Y = Y^i\frac{\partial}{\partial q^i} + \dot{Y}^j\frac{\partial}{\partial\vartheta^j}$, $i = 1, \dots, n$, $j = 1, \dots, d$. By introducing vectors $Y_1 = Y^i\frac{\partial}{\partial q^i} \in T_mM$ and $Y_2 = \dot{Y}^j\frac{\partial}{\partial\vartheta^j} \in T_\vartheta\Theta_m$ the vectors $Y_{(m,\vartheta)} \in T_{(m,V)}\Theta$ are represented as quadruples (m, ϑ, Y_1, Y_2). This notation is compatible with that of Convention 1.3 for tangent vectors as points of the tangent bundle: here (m,ϑ) is a point of the manifold Θ and (Y_1, Y_2) is a tangent vector to Θ.

Let us find the formula of transformation of Y_1 and Y_2 under standard changes of coordinates of the form $(\varphi_{\beta\alpha}, g_{\beta\alpha}(m))$ (see Definition 1.32) on the total space Θ. Recall that, since Θ is a vector bundle, $g_{\beta\alpha}(m)$ is a linear operator in $\mathbb{R}^d$ and so it is equal to its derivative. On the other hand the derivative in $m \in M$ of the linear operator $g_{\beta\alpha}(m)$, depending on m, is a bilinear operator. Denote it by $g'_{\beta\alpha}(m)(\cdot,\cdot)$. The first argument of this operator is a vector from the fiber and the second argument is a vector tangent to the base M. In particular, the derivative $g_{\beta\alpha}(m)$ in m at the point $(m,\vartheta) \in \Theta$ takes the form $g'_{\beta\alpha}(m)(\vartheta,\cdot)$. Since $\varphi_{\beta\alpha}$ does not depend on the points of fiber, the derivative of $\varphi_{\beta\alpha}$ in $\mathbb{R}^d$ equals zero. Taking this into account it is easy to see that the derivative of the change of coordinates $(\varphi_{\beta\alpha}, g_{\beta\alpha}(m))$ at the point $(m,\vartheta) \in \Theta$ is represented in the form

$$(\varphi_{\beta\alpha}, g_{\beta\alpha}(m))' = \begin{pmatrix} \varphi'_{\beta\alpha} & 0 \\ g'_{\beta\alpha}(m)(\vartheta,\cdot) & g_{\beta\alpha}(m) \end{pmatrix}.$$

This means that under the above-mentioned changes of coordinates the column (Y_1, Y_2) transforms by the formula

$$\begin{aligned}(Y_1^\beta, Y_2^\beta)_{(m^\beta,\vartheta^\beta)} &= \begin{pmatrix} \varphi'_{\beta\alpha} & 0 \\ g'_{\beta\alpha}(m^\alpha)(\vartheta^\alpha,\cdot) & g_{\beta\alpha}(m^\alpha) \end{pmatrix} \begin{pmatrix} Y_1^\alpha \\ Y_2^\alpha \end{pmatrix} \\ &= \begin{pmatrix} \varphi'_{\beta\alpha}Y_1^\alpha \\ g'_{\beta\alpha}(m^\alpha)(\vartheta^\alpha, Y_1^\alpha) + g_{\beta\alpha}(m^\alpha)(Y_2^\alpha) \end{pmatrix}. \end{aligned} \tag{2.1}$$

In terms of quadruples formula (2.1) takes the form

$$\begin{aligned}&(m^\beta, \vartheta^\beta, Y_1^\beta, Y_2^\beta) \\ &= (\varphi_{\beta\alpha}m^\alpha, g_{\beta\alpha}(m^\alpha)\vartheta^\alpha, \varphi'_{\beta\alpha}Y_1^\alpha, g'_{\beta\alpha}(m^\alpha)(\vartheta^\alpha, Y_1^\alpha) + g_{\beta\alpha}(m^\alpha)(Y_2^\alpha)).\end{aligned} \tag{2.2}$$

By the definition of the projection π we have $\pi(q^1, \dots, q^n, \vartheta^1, \dots, \vartheta^d) = (q^1, \dots, q^n)$. Hence, the Jacobi matrix (presentation of the differential $d_{(m,\vartheta)}\pi$ in the given coordinate system) takes the form $(I \quad 0)$ where I and 0 are the unit $n \times n$ matrix and zero $k \times n$ matrix, respectively. As a consequence we obtain the formula for $T\pi : T\Theta \to TM$ in the form

$$T\pi(m, \vartheta, Y_1, Y_2) = (m, Y_1) \tag{2.3}$$

(by definition the tangent mapping $T\pi$ acts as π on the points (m, ϑ) and as $\mathrm{d}_{(m,\vartheta)}\pi$ on the vectors (Y_1, Y_2)). Recall that by construction Y_1 on the left-hand side of (2.3) belongs to $T_{(m,\vartheta)}\Theta$ while Y_1 on the right-hand side of (2.3) belongs to $T_m M$ but both vectors have the same coordinates with respect to $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}$ (the first "half"of the basis in $T_{(m,\vartheta)}\Theta$ and the entire basis in $T_m M$, respectively, are isomorphic to each other, see above). We do not distinguish between these two vectors or the frames in this notation.

Formula (2.3) means that

$$T\pi \left(Y^i \frac{\partial}{\partial q^i} + \dot{Y}^j \frac{\partial}{\partial \vartheta^j} \right) = Y^i \frac{\partial}{\partial q^i}. \tag{2.4}$$

Remark 2.2. As on every manifold, there is a natural projection of $T\Theta$ onto Θ. Denote it by $\pi_1 : T\Theta \to \Theta$. In coordinates it is represented in the form

$$\pi_1(m, \vartheta, Y_1, Y_2) = (m, \vartheta). \tag{2.5}$$

We emphasize the difference between (2.3) and (2.5).

Let Y be a cross-section of the bundle Θ. Over the chart $\mathcal{U}_\alpha$ we have the decomposition $Y = Y^i e_i$ (see above). Recall that we consider the cross-section Y as a mapping $Y : M \to \Theta$ such that $\pi Y = \mathrm{id}$ (see Definition 1.38). Consider also its tangent mapping $TY : TM \to T\Theta$. Since Y has the form

$$Y(q^1, \dots, q^n) = (q^1, \dots, q^n, Y^1(q^1, \dots, q^n), \dots, Y^k(q^1, \dots, q^n)),$$

its Jacobi matrix takes the form

$$\mathrm{d}_m Y = \begin{pmatrix} I \\ \left(\frac{\partial Y^i}{\partial q^j} \right) \end{pmatrix},$$

where I is the unit $n \times n$ matrix and $\left(\frac{\partial Y^i}{\partial q^j} \right)$ is the $n \times d$ Jacobi matrix of Y. Thus for $(m, X) \in TM$ we obtain

$$TY(m, X) = \left(m, Y, X, \left(\frac{\partial Y^i}{\partial q^j} \right) X \right) \tag{2.6}$$

(recall that TY acts as Y on points m and as $\mathrm{d}_m X$ on X).

On the vector bundle Θ the so-called *action of the real line* is given as follows. For every $a \in \mathbb{R}$ defined $a : \Theta \to \Theta$ (we denote the number and the corresponding mapping by the same symbol a) by:

$$a(m, \vartheta) = (m, a\vartheta), \tag{2.7}$$

where $(m, \vartheta) \in \Theta$, i.e. the action consists of multiplying all vectors from all fibers of Θ by a. Thus $a(q^1, \dots, q^n, \vartheta^1, \dots, \vartheta^d) = (q^1, \dots, q^n, a\vartheta^1, \dots, a\vartheta^d)$ and evidently $\mathrm{d}_{(m,\vartheta)}a = \begin{pmatrix} I & 0 \\ 0 & aI \end{pmatrix}$, where I and 0 are the unit and zero matrices, respectively, of corresponding dimensions. Hence

$$Ta(m, \vartheta, Y_1, Y_2) = (m, a\vartheta, Y_1, aY_2). \tag{2.8}$$

Below we shall often use the constructions described in this section on tangent and cotangent bundles. For these cases we have to define the previous formulae and notation more precisely.

Since the fibers of a tangent bundle are tangent spaces, they already have the standard frames $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}$. For the case of a tangent bundle we most often use such frames and the trivialization generated by them in a tangent bundle over charts (the construction of this trivialization is described in Section 1.1). Sometimes we shall also use alternative trivializations but those cases will be mentioned explicitly.

In this case the notation $\dot{q}^i$ for coordinates in fibers is compatible with the interpretation of a tangent vector as a velocity of some curve. We replace ϑ^i by this notation. Thus the frames in tangent spaces to TM have the form $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}, \frac{\partial}{\partial \dot{q}^1}, \dots, \frac{\partial}{\partial \dot{q}^n}$.

For an analogous trivialization in a cotangent bundle we use the basis $\mathrm{d}q^1, \dots, \mathrm{d}q^n$ and coordinates in fibers with respect to those frames are denoted by p_i (here we take into account the interpretation of cotangent vectors as momenta). Respectively, the frame in a cotangent space to T^*M takes the form $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}, \frac{\partial}{\partial p_1}, \dots, \frac{\partial}{\partial p_n}$.

The frame in a cotangent space to TM is denoted by $\mathrm{d}q^1, \dots, \mathrm{d}q^n$, $\mathrm{d}\dot{q}^1, \dots, \mathrm{d}\dot{q}^n$ and in a cotangent space to T^*M by $\mathrm{d}q^1, \dots, \mathrm{d}q^n, \mathrm{d}p_1, \dots, \mathrm{d}p_n$.

We consider in detail the case of tangent bundles. The constructions on cotangent bundles are analogous.

Definition 2.3. The tangent bundle to a tangent bundle TM is called the *second tangent bundle* to the manifold M and is denoted by TTM or T^2M.

The vectors of the second tangent bundle, i.e., tangent vectors to TM, are described as quadruples of the form (m, X, Y_1, Y_2) where X and Y_1 belong to T_mM while Y_2 is a vector tangent to T_mM.

The Jacobi matrix of the natural projection $\pi : TM \to M$ has the form $(I, 0)$ where I is the unit matrix and 0 is the zero matrix, both $n \times n$. Thus $T\pi\left(Y^i \frac{\partial}{\partial q^i} + \tilde{Y}^i \frac{\partial}{\partial \dot{q}^i}\right) = Y^i \frac{\partial}{\partial q^i}$.

The action of the real line on TM is defined as a particular case of the general definition. For $a \in \mathbb{R}$ the Jacobi matrix of the corresponding mapping $a : TM \to TM$ has the same form as above.

The transformation rule for quadruples describing vectors of the second tangent bundle under changes of coordinates also has to be specified. First of all on TM the transformation $g_{\beta\alpha}(m)$ in fibers takes the form $g_{\beta\alpha}(m) = \varphi'_{\beta\alpha}$. Hence $g'_{\beta\alpha}(m)(\cdot,\cdot) = \varphi''_{\beta\alpha}(\cdot,\cdot)$ where $''$ denotes the second derivative of the change of coordinates $\varphi_{\beta\alpha}$. Since the fiber of TM at m is the tangent space T_mM, both arguments in $\varphi''_{\beta\alpha}(\cdot,\cdot)$ have the same nature: they are vectors tangent to M. This is why we replace the symbol ϑ in the notation for an element in the fiber of Θ by the symbol X of a tangent vector to M.

Thus formula (2.1) is transformed into

$$
\begin{aligned}
(Y_1^\beta, Y_2^\beta)_{(m^\beta, X^\beta)} &= \begin{pmatrix} \varphi'_{\beta\alpha} & 0 \\ \varphi''_{\beta\alpha}(X^\alpha, \cdot) & \varphi'_{\beta\alpha} \end{pmatrix} \begin{pmatrix} Y_1^\alpha \\ Y_2^\alpha \end{pmatrix} \\
&= \begin{pmatrix} \varphi'_{\beta\alpha} Y_1^\alpha \\ \varphi''_{\beta\alpha}(m^\alpha)(X^\alpha, Y_1^\alpha) + \varphi'_{\beta\alpha}(m^\alpha)(Y_2^\alpha) \end{pmatrix}.
\end{aligned}
\tag{2.9}
$$

So, by formula (2.9) the transformation of quadruples as vectors tangent to TTM under the change of coordinates $\varphi_{\beta\alpha}$ on M has the form

$$
\begin{aligned}
&(m^\beta, X^\beta, Y_1^\beta, Y_2^\beta) \\
&= (\varphi_{\beta\alpha} m^\alpha, \varphi'_{\beta\alpha} X^\alpha, \varphi'_{\beta\alpha} Y_1^\alpha, \varphi''_{\beta\alpha}(X^\alpha, Y_1^\alpha) + \varphi'_{\beta\alpha}(Y_2^\alpha)).
\end{aligned}
\tag{2.10}
$$

We return to the general case.

Recall that by Definition 2.1 the space $T_\vartheta \Theta_m$ is called the vertical subspace in $T_{(m,\vartheta)}\Theta$ and is denoted by $\mathsf{V}_{(m,\vartheta)}$. The vectors belonging to $\mathsf{V}_{(m,\vartheta)}$ are said to be vertical.

As a direct consequence of the construction we obtain the following:

Proposition 2.4 *The space $\mathsf{V}_{(m,\vartheta)}$ does not depend on the choice of the chart $\mathfrak{U}_\alpha$, its coordinate system $(q^1, \ldots, q^n)$ in a neighborhood of $m \in M$, or on the choice of the trivialization of $\pi^{-1}\mathfrak{U}_\alpha$.*

Indeed, the fiber Θ_m and hence the tangent space $T_\vartheta \Theta_m = V_{(m,\vartheta)}$ are determined without use of any coordinate system. The system $(q^1, \ldots, q^n)$ is involved only in representing $\mathsf{V}_{(m,\vartheta)}$ as the linear span of $\frac{\partial}{\partial \vartheta^1}, \ldots, \frac{\partial}{\partial \vartheta^d}$. Notice that these vectors do depend on the trivialization of $\pi^{-1}\mathfrak{U}_\alpha$.

Recall that by formula (1.2) we introduced the linear isomorphism $\mathbf{p}$ of a tangent space to a vector space onto the vector space. Thus here $\mathbf{p} : V_{(m,\vartheta)} \to \Theta_m$ is well-defined and takes the coordinate representation

$$
\mathbf{p}\left(\frac{\partial}{\partial \vartheta^i}\right) = e_i. \tag{2.11}
$$

Denote by $\mathsf{H}^E_{(m,\vartheta)}$ the linear span of the vectors $\frac{\partial}{\partial q^1}, \ldots, \frac{\partial}{\partial q^n}$ in $T_{(m,\vartheta)}\Theta$. By construction, $\mathsf{H}^E_{(m,\vartheta)}$ is the tangent space to the submanifold $\mathcal{U}_\alpha \times \vartheta$ in $\pi^{-1}\mathcal{U}_\alpha$ with respect to the given trivialization of $\pi^{-1}\mathcal{U}_\alpha$. Immediately from the definition we get $T_{(m,\vartheta)}\Theta = \mathsf{H}^E_{(m,\vartheta)} \oplus \mathsf{V}_{(m,\vartheta)}$ where $\oplus$ is the direct sum. Notice that for the vector $Y_{(m,\vartheta)} = (m, \vartheta, Y_1, Y_2) \in T_{(m,\vartheta)}\Theta$ by definition $Y_1 \in \mathsf{H}^E_{(m,\vartheta)}$ and $Y_2 \in \mathsf{V}_{(m,\vartheta)}$ so that $Y_{(m,\vartheta)} = Y_1 \oplus Y_2$.

Proposition 2.5 *The subspace $\mathsf{H}^E_{(m,\vartheta)}$ depends on the choice of trivialization of $\pi^{-1}\mathcal{U}_\alpha$ and hence on the chart $\mathcal{U}_\alpha$.*

Proof. Indeed, consider another chart $\mathcal{U}_\beta$ with non-empty intersection $\mathcal{U}_{\alpha\beta}$ with $\mathcal{U}_\alpha$. Let a trivialization of $\pi^{-1}\mathcal{U}_\beta$ be given so that the standard basis in $\mathbb{R}^d$ generates another field of bases $e'_1, \ldots, e'_d$ different from the field $e_1, \ldots, e_d$ generated by the above trivialization of $\pi^{-1}\mathcal{U}_\alpha$. The layers in $\mathcal{U}_\alpha \times \mathbb{R}^d$ of the vectors $\mathcal{U}_\alpha \times \vartheta$ and in $\mathcal{U}_\beta \times \mathbb{R}^d$ of the vectors $\mathcal{U}_\beta \times \vartheta$ for a specified vector $\vartheta \in \mathbb{R}^d$ are different since the former is generated by those ϑ''s from $\Theta_{m'}$, $m' \in \mathcal{U}_\alpha$ whose coordinates with respect to $e_1, \ldots, e_d$ are the same as the coordinates of ϑ and the latter by those whose coordinates with respect to $e'_1 \ldots, e'_d$ are the same as those of ϑ. Since the layers going through (m, ϑ) are different, their tangent spaces at (m, ϑ) are also different. □

Remark 2.6. From the definitions it immediately follows that the quadruple for a vector from $\mathsf{H}^E_{(m,\vartheta)}$ takes the form $(m, \vartheta, Y_1, 0)$ and, for a vector from $\mathsf{V}_{(m,\vartheta)}$, the form $(m, \vartheta, 0, Y_2)$.

Proposition 2.7 *$T\pi$ sends any $\mathsf{H}^E_{(m,\vartheta)}$ isomorphically onto $T_m M$ and $\mathsf{V}_{(m,\vartheta)}$ is the kernel of $T\pi$ at any $T_{(m,\vartheta)}\Theta$.*

Indeed, a vector from $\mathsf{H}^E_{(m,\vartheta)}$ takes the form $(m, \vartheta, Y_1, 0)$ and from $\mathsf{V}^E_{(m,\vartheta)}$ the form $(m, \vartheta, 0, Y_2)$ (see Remark 2.6). So, by (2.3) $T\pi(m, \vartheta, Y_1, 0) = (m, Y_1)$ and $T\pi(m, \vartheta, 0, Y_2) = (m, 0)$.

2.2 Connections on Vector Bundles

Connection and connector

Definition 2.8. Let Θ be a vector bundle and suppose that in every tangent space $T_{(m,\vartheta)}\Theta$ a subspace $\mathsf{H}_{(m,\vartheta)}$, complementary to $\mathsf{V}_{(m,\vartheta)}$ (i.e. $T_{(m,\vartheta)}\Theta = \mathsf{H}_{(m,\vartheta)} \oplus \mathsf{V}_{(m,\vartheta)}$ at any $(m, \vartheta) \in \Theta$), is specified such that the total family of subspaces $\mathsf{H} = \{\mathsf{H}_{(m,\vartheta)} | (m, \vartheta) \in \Theta\}$ satisfies the following two properties:

(i) the space $\mathsf{H}_{(m,\vartheta)}$ depends smoothly on $(m, \vartheta) \in \Theta$ (in the sense described below);
(ii) the family H is invariant with respect to the action of the real line on Θ, i.e., $Ta\mathsf{H}_{(m,\vartheta)} = \mathsf{H}_{(m,a\vartheta)}$ for every $a \in \mathbb{R}$ and $(m, \vartheta) \in \Theta$.

Then H is said to be a *connection* on Θ.

The subspaces $\mathsf{H}_{(m,\vartheta)}$ of a connection H are called *horizontal*, as are the vectors of $T_{(m,\vartheta)}\Theta$ belonging to $\mathsf{H}_{(m,\vartheta)}$.

The precise meaning of the statement that $\mathsf{H}_{(m,\vartheta)}$ is smooth in (m,ϑ) is as follows. In a neighborhood of any point $(m,\vartheta) \in \Theta$ there are n smooth linearly independent vector fields such that, for any (m',ϑ') in the neighborhood, the subspace $\mathsf{H}_{(m',\vartheta')}$ is the linear span of vectors of those fields at (m',ϑ').

Proposition 2.9 *The family* $\mathsf{H}^E_{(m,\vartheta)}$ *introduced in Section 2.1 is a connection on* $\pi^{-1}\mathfrak{U}_\alpha$.

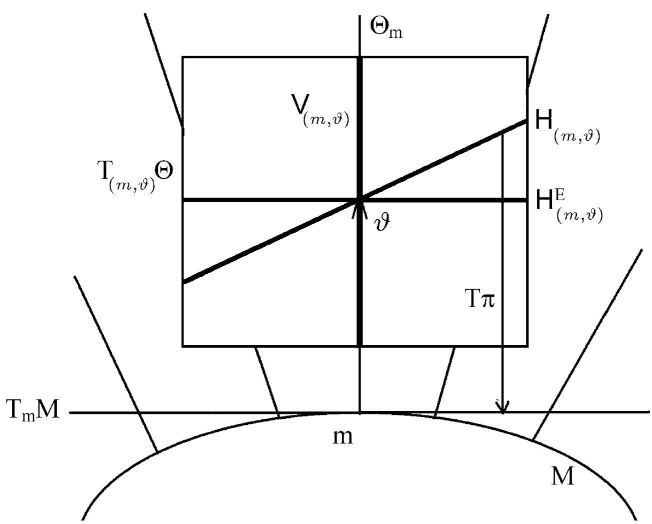
Proof. The presentation $T_{(m,\vartheta)}\Theta = \mathsf{H}^E_{(m,\vartheta)} \oplus \mathsf{V}_{(m,\vartheta)}$ was derived in Section 2.1 from the definition of $\mathsf{H}^E_{(m,\vartheta)}$. Also by definition $\mathsf{H}^E_{(m,\vartheta)}$ is the linear span of smooth linearly independent vectors $\frac{\partial}{\partial q^1}, \ldots, \frac{\partial}{\partial q^n}$.

At any point $(m,\vartheta) \in \Theta$ the space $\mathsf{H}^E_{(m,\vartheta)}$ is the set of all vectors whose quadruple presentation takes the form $(m,\vartheta,Y_1,0)$ (see Remark 2.6). By formula (2.8) we see that Ta is a one-to-one mapping sending $(m,\vartheta,Y_1,0)$ to the quadruple $(m,a\vartheta,Y_1,0)$. Thus $Ta\mathsf{H}_{(m,\vartheta)} = \mathsf{H}_{(m,a\vartheta)}$. □

Definition 2.10. The family of subspaces $\mathsf{H}^E_{(m,\vartheta)}$ is called the *Euclidean connection* of a given trivialization of $\pi^{-1}\mathfrak{U}_\alpha$.

Indeed, $\mathsf{H}^E_{(m,\vartheta)}$ depends on the trivialization (see Proposition 2.5).

There exist connections $\{\mathsf{H}_{(m,\vartheta)}\}$ that may not be presented as the Euclidean connection of a trivialization. Of course, at a given point (m,ϑ) for a subspace $\mathsf{H}_{(m,\vartheta)}$, complimentary to $\mathsf{V}_{(m,\vartheta)}$, one can find a trivialization such that $\mathsf{H}_{(m,\vartheta)}$ coincides with the Euclidean connection of a trivialization at (m,ϑ), but in general this cannot be achieved for subspaces at all points in a given neighborhood. In order not to exclude general connections, we have not limited ourselves to Euclidean connections.



Proposition 2.11 $T\pi : \mathsf{H}_{(m,\vartheta)} \to T_mM$ *is a linear isomorphism.*

Proof. Recall that $T\pi : T_{(m,\vartheta)}\Theta \to T_m M$ is surjective and (by Proposition 2.7) $\mathsf{V}_{(m,\vartheta)}$ is the kernel of $T\pi$. Thus, the Proposition follows from the general result of linear algebra that a surjective linear operator is one-to-one on a complement to the kernel. □

The above proof is also valid in the analogous case of $\mathsf{H}^E_{(m,\vartheta)}$ in Proposition 2.7. However, we used a coordinate proof there for simplicity.

Combining Propositions 2.7 and 2.11 we see that $T\pi$ is connected with the decomposition $T_{(m,\vartheta)}\Theta = \mathsf{H}_{(m,\vartheta)} \oplus \mathsf{V}_{(m,\vartheta)}$ as follows:

Lemma 2.12

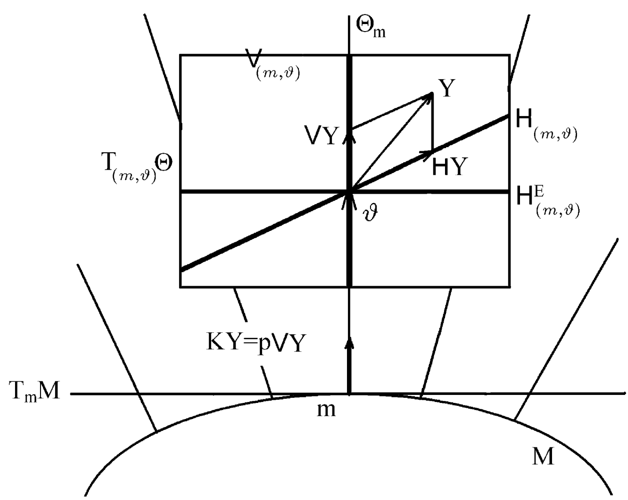
(i) *$T\pi : \mathsf{H}_{(m,\vartheta)} \to T_m M$ is an isomorphism.*
(ii) *$\mathsf{V}_{(m,\vartheta)} = \ker T\pi$.*

Our next step is to construct a map that is one-to-one on $\mathsf{V}_{(m,\vartheta)}$ and whose kernel is $\mathsf{H}_{(m,\vartheta)}$. Recall that the operator $\mathbf{p}$ that establishes a linear isomorphism between the vector space Θ_m and the tangent space $T_\vartheta\Theta_m$ to it acts by formula (2.11). Hence, for a vector from $\mathsf{V}_{(m,\vartheta)}$ with the quadruple $(m, \vartheta, 0, Y_2)$ (see Remark 2.6) where $Y_2 = \dot{Y}^i \frac{\partial}{\partial \dot{q}^i}$, we have

$$\mathbf{p}(m, \vartheta, 0, Y_2) = (m, \mathbf{p}Y_2) = \dot{Y}^i e_i. \tag{2.12}$$

The decomposition $T_{(m,\vartheta)}\Theta = \mathsf{H}_{(m,\vartheta)} \oplus \mathsf{V}_{(m,\vartheta)}$ yields the decomposition $Y_{(m,\vartheta)} = \mathsf{H}Y \oplus \mathsf{V}Y$ for every $Y_{(m,\vartheta)} \in T_{(m,\vartheta)}\Theta$, where $\mathsf{H}Y \in \mathsf{H}_{(m,\vartheta)}$ and $\mathsf{V}Y \in \mathsf{V}_{(m,\vartheta)}$. The symbols H and V may be considered as projections $\mathsf{H} : T_{(m,\vartheta)}\Theta \to \mathsf{H}_{(m,\vartheta)}$ and $\mathsf{V} : T_{(m,\vartheta)}\Theta \to \mathsf{V}_{(m,\vartheta)}$ in the above decomposition.

Definition 2.13. The map $K = \mathbf{p}\mathsf{V} : T_{(m,\vartheta)}\Theta \to \Theta$ is called the *connector* of the connection H.



Thus $K(Y_{(m,\vartheta)}) = \mathbf{p}(\mathsf{V}Y_{(m,\vartheta)})$. Evidently K on $\mathsf{V}_{(m,\vartheta)}$ coincides with $\mathbf{p}$ and so K maps $\mathsf{V}_{(m,\vartheta)}$ onto Θ_m isomorphically. On the other hand, $\mathsf{V}(\mathsf{H}_{(m,\vartheta)}) = 0 \in \mathsf{V}_{(m,\vartheta)}$ and so $K(\mathsf{H}_{(m,\vartheta)}) = 0 \in \Theta_m$. We summarize these properties in the following lemma:

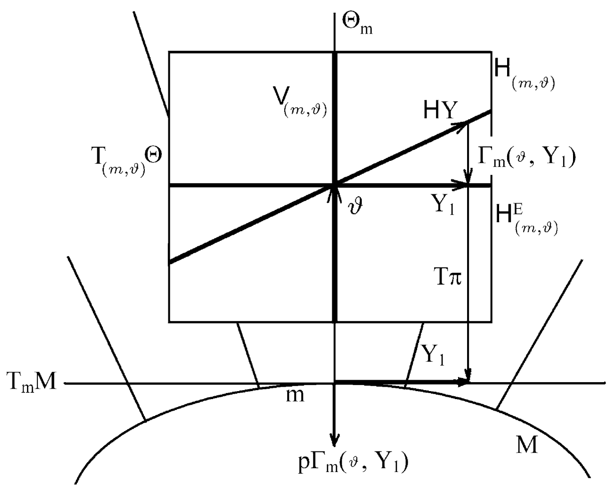
Lemma 2.14

(i) $K : \mathsf{V}_{(m,\vartheta)} \to \Theta_m$ *is a linear isomorphism.*
(ii) $\mathsf{H}_{(m,\vartheta)} = \ker K$.

Compare Lemmas 2.12 and 2.14. Notice the difference: we know that $\mathsf{V}_{(m,\vartheta)}$ and $T\pi$ exist on each vector bundle Θ while $\mathsf{H}_{(m,\vartheta)}$ and K must be given "by hand".

In order to work with $\mathsf{H}_{(m,\vartheta)}$ and K we need to describe them by means of coordinates. The best way to do that is to compare $\mathsf{H}_{(m,\vartheta)}$ with $\mathsf{H}^E_{(m,\vartheta)}$ of a certain trivialization over a chart $\mathfrak{U}_\alpha$ since the coordinate presentation of $\mathsf{H}^E_{(m,\vartheta)}$ is known.

Consider a vector $Y_1 \in T_mM$. Since $T\pi$ sends both $\mathsf{H}_{(m,\vartheta)}$ and $\mathsf{H}^E_{(m,\vartheta)}$ onto T_mM one-to-one, each of the spaces contains a unique vector whose image under $T\pi$ is Y_1. The vector in $\mathsf{H}^E_{(m,\vartheta)}$ is usually denoted by the same symbol Y_1. Denote the vector in $\mathsf{H}_{(m,\vartheta)}$ by $\mathsf{H}Y$. Consider the difference $\Gamma_m(\vartheta, Y_1) = Y_1 - \mathsf{H}Y \in T_{(m,\vartheta)}\Theta$. By construction we have $T\pi\Gamma_m(\vartheta, Y_1) = T\pi(Y_1) - T\pi(\mathsf{H}Y) = Y_1 - Y_1 = 0 \in T_mM$. Hence, $\Gamma_m(\vartheta, Y_1) \in \mathsf{V}_{(m,\vartheta)}$ since $\mathsf{V}_{(m,\vartheta)}$ is the kernel of $T\pi$. Thus we can apply $\mathbf{p}$ to $\Gamma_m(\vartheta, Y_1)$ and obtain the vector $\mathbf{p}\Gamma_m(\vartheta, Y_1) \in \Theta_m$. We have constructed the operator $\mathbf{p}\Gamma_m(\cdot,\cdot) : \Theta_m \times T_mM \to \Theta_m$.



Definition 2.15. The operator $\mathbf{p}\Gamma_m(\cdot,\cdot)$ is called the *local connector* (or *local connection coefficient*) of the connection H.

The word "local" means that the operator is constructed and calculated in a certain chart $\mathfrak{U}_\alpha$ on M with respect to a certain trivialization of $\pi^{-1}\mathfrak{U}_\alpha$.

Theorem 2.16 *The operator* $\mathbf{p}\Gamma_m(\cdot,\cdot)$ *is linear in the second argument.*

Indeed, $\mathbf{p}\Gamma_m(\vartheta, Y_1) = T\pi^{-1}(Y_1)_{|\mathsf{H}^E_{(m,\vartheta)}} - T\pi^{-1}(Y_1)_{|\mathsf{H}_{(m,\vartheta)}}$. Since the operation $T\pi^{-1}$ and the operation of taking the difference are both linear, $\mathbf{p}\Gamma_m(\vartheta, Y_1)$ is linear in Y_1.

Theorem 2.17 *The operator* $\mathbf{p}\Gamma_m(\cdot,\cdot)$ *is linear in the first argument.*

To prove Theorem 2.17 we need the following:

Lemma 2.18 *Let* $B : E \to E$ *be a map in the vector space* E*, smooth and homogeneous with degree* 1*. Then* B *is a linear operator.*

Proof. (of Lemma 2.18) Recall that B is homogeneous with degree k if for any vector $X \in E$ and any $\lambda \in \mathbb{R}$ we have $B(\lambda X) = \lambda^k B(X)$. From the homogeneity it follows that $B(0) = 0$.

Since B is smooth, we can expand it by the Taylor formula in a neighborhood of $0 \in E$ up to a certain degree greater than 1. Thus, since $B(0) = 0$, $B(X) = B'(X) + \frac{1}{2}B''(X,X) + \ldots$ where B' is the first derivative of B at the origin (recall that B' is a linear operator), B'' is the second derivative of B at the origin (recall that B'' is a bilinear operator), etc. On the right-hand side only B' is homogeneous with degree 1; $B''(X,X)$ is homogeneous with degree 2 and the other summands have greater degrees of homogeneity. Thus the left-hand side is homogeneous with degree 1 only if all summands on the right hand side except B' are equal to zero. Hence $B = B'$ and so it is a linear operator. □

Proof. (of Theorem 2.17) Since by Definition 2.8(i) both $\mathsf{H}_{(m,\vartheta)}$ and $\mathsf{H}^E_{(m,\vartheta)}$ are smooth in ϑ, so too is $\mathbf{p}\Gamma_m(\vartheta, Y_1)$. We shall show that $\mathbf{p}\Gamma_m(\vartheta, Y_1)$ is homogeneous with degree 1 in ϑ so that the statement of Theorem 2.17 will follow from Lemma 2.18.

The vector $\mathsf{H}Y = T\pi^{-1}(Y_1) \in \mathsf{H}_{(m,\vartheta)}$ is presented as a quadruple in the form $(m, \vartheta, Y_1, \Gamma_m(\vartheta, Y_1))$. By Definition 2.8(ii) and by formula (2.8) (describing Ta) the vector $Ta(\mathsf{H}Y) = (m, a\vartheta, Y_1, a\Gamma_m(\vartheta, Y_1))$ belongs to $\mathsf{H}_{(m,a\vartheta)}$. Using formula (2.3) we get $T\pi((m, a\vartheta, Y_1, a\Gamma_m(\vartheta, Y_1)) = (m, Y_1)$. Since $T\pi$ is one-to-one on $\mathsf{H}_{(m,a\vartheta)}$, it is the unique vector in $\mathsf{H}_{(m,a\vartheta)}$ whose image under $T\pi$ is (m, Y_1). But the vector $\mathsf{H}Y = T\pi^{-1}(Y_1) \in \mathsf{H}_{(m,a\vartheta)}$, whose quadruple takes the form $(m, a\vartheta, Y_1, \Gamma_m(a\vartheta, Y_1))$, also has this property: $T\pi(m, a\vartheta, Y_1, \Gamma_m(a\vartheta, Y_1)) = (m, Y_1)$. Hence,

$$(m, a\vartheta, Y_1, a\Gamma_m(\vartheta, Y_1)) = (m, a\vartheta, Y_1, \Gamma_m(a\vartheta, Y_1))$$

and so, since $\mathbf{p}$ is linear, $\mathbf{p}\Gamma_m(a\vartheta, Y_1) = a\mathbf{p}\Gamma_m(\vartheta, Y_1)$. □

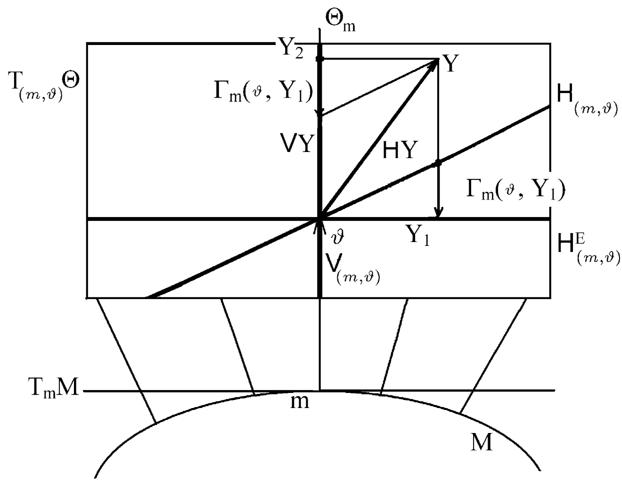
So, $\mathbf{p}\Gamma_m(\cdot,\cdot)$ is linear in both arguments. It is useful to find its values on basis vectors. Consider $\mathbf{p}\Gamma_m(e_i, \frac{\partial}{\partial q^j})$. It is a vector from Θ_m and so it can be expanded in coordinates Γ_{ij}^k with respect to the basis $e_1, \dots, e_d$: $\mathbf{p}\Gamma_m(e_i, \frac{\partial}{\partial q^j}) = \Gamma_{ij}^k e_k$. The coordinates Γ_{ij}^k depend on $m \in \mathfrak{U}_\alpha$ (as well as on a trivialization), this means that they are real-valued functions of $m \in \mathfrak{U}_\alpha$.

Definition 2.19. The functions Γ_{ij}^k are called *Christoffel symbols of the second kind* for the connection H.

Knowing Γ_{ij}^k, we can calculate the values $\mathbf{p}\Gamma_m(X,Y)$ for any $X \in \Theta_m$, $Y \in T_m M$. Indeed, let $X = X^i e_i$ and $Y = Y^j \frac{\partial}{\partial q^j}$, then by linearity we get

$$\mathbf{p}\Gamma(X,Y) = X^i Y^j \Gamma_{ij}^k e_k. \tag{2.13}$$

Now let us turn back to the connector K. Recall that $K(Y) = \mathbf{p}\mathsf{V}Y$ for $Y \in T_{(m,\vartheta)}\Theta$. Thus we need to describe $\mathbf{p}\mathsf{V}Y$. For Y we have two decompositions:



$Y = Y_1 + Y_2$ and $Y = \mathsf{H}Y + \mathsf{V}Y$. Hence $Y_1 + Y_2 = \mathsf{H}Y + \mathsf{V}Y$ and so $\mathsf{V}Y - Y_2 = Y_1 - \mathsf{H}Y = \Gamma_m(\vartheta, Y_1)$. Thus $\mathsf{V}Y = Y_2 + \Gamma_m(\vartheta, Y_1)$ and consequently $\mathsf{V}Y = \mathsf{V}(m, \vartheta, Y_1, Y_2) = (m, \vartheta, 0, Y_2 + \Gamma_m(\vartheta, Y_1))$. Finally we obtain the formula for K in the form:

$$K(m, \vartheta, Y_1, Y_2) = \mathbf{p}\mathsf{V}(m, \vartheta, Y_1, Y_2) = (m, \mathbf{p}Y_2 + \mathbf{p}\Gamma(\vartheta, Y_1)). \tag{2.14}$$

Compare (2.14) with (2.3) and (2.5).

Let $Y_2 = \dot{Y}^k \frac{\partial}{\partial \vartheta^k}$ and $\vartheta = \dot{q}^i e_i$. Using (2.12) we describe (2.14) in coordinates as follows

$$K(m, \vartheta, Y_1, Y_2) = (\dot{Y}^k + \dot{q}^i Y^j \Gamma_{ij}^k) e_k. \tag{2.15}$$

Remark 2.20. If we choose arbitrary functions $\Gamma_{ij}^k(m)$ on $\mathcal{U}_\alpha$ for all possible values of i, j and k, we shall be able to define a local connector $\mathbf{p}\Gamma_m(\cdot,\cdot)$ by formula (2.13) and consequently a connector K by formula (2.14) or (2.15) and then define the corresponding connection H on $\pi^{-1}\mathcal{U}_\alpha$ as kernels of K in all tangent spaces.

Remark 2.21. We say that $\mathbf{p}\Gamma_m(\vartheta, Y_1)$ is a vector in Θ_m. If we change the trivialization, this vector will no longer correspond to the local connector. Indeed, the Euclidean connection will be changed (see Proposition 2.5) and the old $\mathbf{p}\Gamma_m(\vartheta, Y_1)$ will not be the difference between $\mathsf{V}Y$ and the new Y_1. So, the change of $\mathbf{p}\Gamma_m(\vartheta, Y_1)$ under a change of coordinates on M and of a trivialization is described by a complicated "non-tensorial" formula that follows from (2.2). We shall derive it in explicit form for some special cases below (see formula (2.19)).

The covariant derivative and parallel translation

Here we present the general construction by which every connection defines its own method of differentiating a cross-section of Θ along a vector field on M. Notice that the use of a Euclidean connection of a natural trivialization of $\mathbb{R}^n \times \mathbb{R}^d$ gives the standard method of differentiating typically introduced in a classical course in mathematical analysis.

Let X be a smooth vector field on M and Y be a cross-section of a vector bundle Θ equipped with a connection H.

Definition 2.22. The *covariant derivative* $\nabla_X Y$ of a cross-section Y along a vector field X is the cross-section of Θ determined by the formula $\nabla_X Y = K \circ TY(X)$.

Let us discuss this definition. The cross-section Y can be considered as a smooth map $Y : M \to \Theta$. Its tangent map TY sends the vector $X \in T_m M$ to the tangent space $T_{(m,Y)}\Theta$. On applying K we again map into Θ_m.

Example 2.23. Consider the Euclidean connection of a natural trivialization of $\mathbb{R}^n \times \mathbb{R}^d$. The section Y can be presented as the map $m \mapsto (m, Y_m)$. We express the tangent map TY in coordinates and find the vector $TY(X)$. Here V coincides with the projection along H^E. One can easily see that the obtained covariant derivative coincides with the ordinary derivative of Y along the field X.

Theorem 2.24 *The covariant derivative has the following properties for any vector fields X, X_1 and X_2, smooth cross-sections Y, Y_1 and Y_2, smooth function $f : M \to \mathbb{R}$ and $\varkappa, \lambda \in \mathbb{R}$:*

(i) $\nabla_{(\varkappa X_1 + \lambda X_2)} Y = \varkappa \nabla_{X_1} Y + \lambda \nabla_{X_2} Y$;
(ii) $\nabla_{fX} Y = f \nabla_X Y$;
(iii) $\nabla_X (\varkappa Y_1 + \lambda Y_2) = \varkappa \nabla_X Y_1 + \lambda \nabla_X Y_2$;

(iv) $\nabla_X fY = (Xf)Y + f\nabla_X Y$,

where Xf is the derivative of f along X.

Proof. Properties (i) and (ii) follow immediately from the linearity of $TY : T_m M \to T_{(m,Y)}\Theta$, of the projection V and of $\mathbf{p}$ (see Definition 2.13 of K). In order to prove (iii) one should recall the action of TY derived in (2.6) and the representation of K via $\mathbf{p}\Gamma_m(\cdot,\cdot)$ in (2.14). Now (iii) follows from the fact that $\mathrm{d}_m Y$ is linear in Y and from the linearity of $\mathbf{p}\Gamma_m(\cdot,\cdot)$ in the first argument (Theorem 2.17). For the proof of (iv), we find a formula for $\mathrm{d}_m(fY)$ according to the usual rules of differentiation as follows:

$$\mathrm{d}_m(fY) = \begin{pmatrix} I \\ \left(\frac{\partial fY^i}{\partial q^j}\right) \end{pmatrix} = \begin{pmatrix} I \\ Y\mathrm{d}f + f\left(\frac{\partial Y^i}{\partial q^j}\right) \end{pmatrix}$$

where $\mathrm{d}f = \frac{\partial f}{\partial q^i}\mathrm{d}q^i$ is the differential of f (see (1.14)). Thus, taking into account that $Xf = \mathrm{d}f(X)$ (see formula (1.16)), we get

$$\begin{aligned} T(fY)X_m = T(fY)(m,X) &= \left(m, fY, X, (Xf)Y + \left(\frac{\partial Y^i}{\partial q^j}\right)X\right) \\ &= (m, fY, 0, (Xf)Y) + \left(m, fY, X, f\left(\frac{\partial Y^i}{\partial q^j}\right)X\right). \end{aligned}$$

By definition $K\left(\left(m, fY, X, f\left(\frac{\partial Y^i}{\partial q^j}\right)X\right)\right) = f\nabla_X Y$. Since $(m, fY, 0, (Xf)Y)$ is vertical (i.e., belongs to $\mathsf{V}_{(m,fY)}$), $K((m, fY, 0, (Xf)Y)) = (m, (Xf)Y)$.
□

Using the expression of K via Christoffel symbols (2.15), we find the expression for $\nabla_X Y$ in local coordinates in the form:

$$\nabla_X Y = \left(\frac{\partial Y^k}{\partial q^j}X^j + Y^i X^j \Gamma_{ij}^k\right)e_k. \tag{2.16}$$

Notice that $\frac{\partial Y^k}{\partial q^j}X^j e_k$ is the ordinary derivative of Y along X as in a trivial bundle. Under a change of trivialization this term transforms incorrectly. Only after adding $Y^i X^j \Gamma_{ij}^k e_k$ does (2.16) retain its form under a change of coordinates and trivialization. In the language used by physicists, this means that formula (2.16) is *covariant*. This is why we call the operation $\nabla_X Y$ the covariant derivative.

For further applications we also need a covariant construction for differentiating a cross-section in the "time" variable t along a certain curve $m(t)$ in M.

Let $m(t)$ be a smooth curve on M and $Y(t)$ be a cross-section of Θ over $m(\cdot)$. This means that at any point $m(t)$ there is associated a vector $Y(t) \in \Theta_{m(t)}$, and $Y(t)$ is smooth in t. The vector $\frac{\mathrm{d}}{\mathrm{d}t}Y(t)$ at any t belongs to the tangent space $T_{(m(t),Y(t))}\Theta$. Consider the vector $\frac{\mathrm{D}}{\mathrm{d}t}Y(t) = K \circ \frac{\mathrm{d}}{\mathrm{d}t}Y(t)$ in $\Theta_{m(t)}$.

Definition 2.25. The vector $\frac{\mathrm{D}}{\mathrm{d}t}Y(t) = K \circ \frac{\mathrm{d}}{\mathrm{d}t}Y(t)$ is called the *covariant derivative* of $Y(t)$ along $m(t)$ in t.

Let us discuss the relation between the operations ∇ and $\frac{\mathrm{D}}{\mathrm{d}t}$. We might hope that $\frac{\mathrm{D}}{\mathrm{d}t}Y(t)$ would be equal to $\nabla_{\dot{m}(t)}Y = K \circ TY(\dot{m}(t))$ if the latter expression were well-defined. Unfortunately this is not the case since the cross-section $Y(t)$ is given only at the points of the curve $m(t)$ while, when determining TY, it is necessary that Y is defined in a neighborhood of $m(t)$.

This is why we have to apply the following trick. On a subinterval of the domain, where the curve has neither self intersections nor periods where it is constant, define an auxiliary smooth vector field $\tilde{Y}$ in a neighborhood of $m(t)$ such that at the points of $m(t)$ it coincides with $Y(t)$: $\tilde{Y}_{m(t)} = Y(t)$. Various constructions of such fields are typically described in textbooks on differential geometry and topology. The expression $\nabla_{\dot{m}(t)}\tilde{Y} = K \circ T\tilde{Y}(\dot{m}(t))$ therefore makes sense.

Theorem 2.26 $\nabla_{\dot{m}(t)}\tilde{Y} = \frac{\mathrm{D}}{\mathrm{d}t}Y(t)$ *and so it does not depend on the choice of smooth vector field* $\tilde{Y}$*.*

Proof. Since the curve $m(t)$ and the map $\tilde{Y} : M \to TM$ are smooth, the curve $\tilde{Y}_{m(t)}$ in Θ is smooth and by the construction of the tangent map $T\tilde{Y}(\dot{m}(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\tilde{Y}_{m(t)}$. But $\tilde{Y}_{m(t)} = Y(t)$, hence $T\tilde{Y}(\dot{m}(t)) = \frac{\mathrm{d}}{\mathrm{d}t}Y(t)$ and so $\nabla_{\dot{m}(t)}\tilde{Y} = K \circ T\tilde{Y}(\dot{m}(t)) = K \circ \frac{\mathrm{d}}{\mathrm{d}t}Y(t) = \frac{\mathrm{D}}{\mathrm{d}t}Y(t)$. In particular $\nabla_{\dot{m}(t)}\tilde{Y}$ does not depend on the choice of $\tilde{Y}$. $\square$

Remark 2.27. Taking into account Theorem 2.26 we shall sometimes use the expression $\frac{\mathrm{D}}{\mathrm{d}t}Y(t) = \nabla_{\dot{m}}Y(t)$ where it is understood that in the right hand side $Y(t)$ represents some $\tilde{Y}$ such that $\tilde{Y}_{m(t)} = Y(t)$. This will simplify the formulae and arguments below.

Thus, in order to obtain a representation of $\frac{\mathrm{D}}{\mathrm{d}t}$ in terms of a local connector, analogous to (2.16), we should replace the vector field X by the velocity vector $\dot{m}(t) = \frac{\mathrm{d}m^j}{\mathrm{d}t}\frac{\partial}{\partial q^j}$ and $T\tilde{Y}(\dot{m}(t))$ by $\frac{\mathrm{d}}{\mathrm{d}t}Y(t)$. So, the analog of (2.16) takes the form

$$\frac{\mathrm{D}}{\mathrm{d}t}Y(t) = \left(\frac{\mathrm{d}Y^k}{\mathrm{d}t} + \Gamma^k_{ij}Y^i\frac{\mathrm{d}m^j}{\mathrm{d}t}\right)e_k. \tag{2.17}$$

If our vector bundle Θ is trivial and a trivialization is specified, the notion of a constant cross-section of Θ is well-defined. Indeed, since Θ is represented as a direct product $M \times \mathbb{R}^d$, the cross-section $M \times Y_0$ corresponding to the layer of a fixed $Y_0 \in \mathbb{R}^d$ can be considered where at any point $m \in M$ the same vector in Θ_m is applied. The visual image here is that all vectors of the cross-section are parallel to each other. The derivative of such a cross-section along any smooth curve in M is equal to zero.

In a general non-trivial bundle the idea of "applying the same vector" at each point of M cannot be realized. Nevertheless we still have a covariant derivative along a curve (rather than an ordinary derivative, which is not convenient, see above) and so we can consider cross-sections along curves with zero covariant derivatives and say that they consist of vectors parallel to each other. Let us give the exact definition.

Definition 2.28. A cross-section $Y(t)$ along a curve $m(t)$, $t \in [0, l]$, is called *parallel* if $\frac{\mathrm{D}}{\mathrm{d}t} Y(t) = 0$ for all $t \in [0, l]$.

It follows from (2.17) that a parallel cross-section is described by the system of first order linear differential equations

$$\frac{\mathrm{d}Y^k}{\mathrm{d}t} + \Gamma^k_{ij} Y^i \frac{\mathrm{d}m^j}{\mathrm{d}t} = 0. \tag{2.18}$$

Theorem 2.29 *For any initial vector $Y_0 \in \Theta_{m(0)}$ there exists a unique solution $Y(t)$ of the system* (2.18)*, well-defined for all $t \in [0, l]$.*

Indeed, this is a well-known existence and uniqueness theorem for linear first order differential equations. The only modification needed here is that one should prove the existence and uniqueness in a finite number of charts since (2.18) is given in terms of local coordinates.

Definition 2.30. The solution $Y(t)$ whose existence is asserted in Theorem 2.29 is called the *parallel translation of vector Y_0 along $m(\cdot)$*.

The idea of parallel translation can also be expressed in another language. Let a vector field X be given on M. At any point $m \in M$ consider the fiber Θ_m and the horizontal subspaces $\mathsf{H}_{(m,\vartheta)}$ at all points $(m, \vartheta) \in \Theta_m$. Recall that (see Proposition 2.7) $T\pi : \mathsf{H}_{(m,\vartheta)} \to T_m M$ is one-to-one and so at any (m, ϑ) we can define the vector $\tilde{X}_{(m,\vartheta)} = T\pi^{-1}(X_m)_{|\mathsf{H}_{(m,\vartheta)}}$.

Definition 2.31. The vector field $\tilde{X}$ on Θ is called the *horizontal lift* of the field X.

Now restrict the bundle Θ to the curve $m(\cdot)$ and consider on $\Theta_{m(\cdot)}$ the horizontal lift of the field $\dot{m}(t)$. This gives a smooth vector field on $\Theta_{m(\cdot)}$ and, taking the initial value $Y_0 \in \Theta_{m(0)}$, we can find the unique integral curve $Y(t)$ of this vector field. One can easily see that $Y(t)$ is the parallel translation of Y_0 according to Definition 2.30.

Let $m(t)$, $t \in [0, T]$, be a smooth curve on M and $\vartheta(t)$ be a cross-section of Θ along $m(\cdot)$ (i.e., $\vartheta(t)$ belongs to the fiber $\Theta_{m(t)}$ for all $t \in [0, T]$). Denote by $\Gamma_{s,t}$ the linear operator of parallel translation along $m(\cdot)$ from $\Theta_{m(t)}$ to $\Theta_{m(s)}$. Consider $\bar{\vartheta}(t) = \Gamma_{s,t}\vartheta(t)$, a curve in the fiber $\Theta_{m(s)}$. Its derivative $\frac{\mathrm{d}}{\mathrm{d}t}\bar{\vartheta}(t)_{|t=s}$ belongs to $T_{\bar{\vartheta}(s)}\Theta_{m(s)}$. Applying to it the operator $\mathbf{p}$, we obtain a vector in the fiber $\Theta_{m(s)}$. Everywhere below we regard $\frac{\mathrm{d}}{\mathrm{d}t}\bar{\vartheta}(t)_{|t=s}$ as a free vector lying in $\Theta_{m(s)}$ and so we do not distinguish in notation between $\mathbf{p}\frac{\mathrm{d}}{\mathrm{d}t}\vartheta(t)_{|t=s}$ and $\frac{\mathrm{d}}{\mathrm{d}t}\vartheta(t)_{|t=s}$.

Theorem 2.32 $\frac{\mathrm{D}}{\mathrm{d}t}\vartheta(t)_{|t=s} = \frac{\mathrm{d}}{\mathrm{d}t}(\Gamma_{s,t}\vartheta(t))_{|t=s}$.

Proof. Since the curve $\Gamma_{s,t}\vartheta(t)$ lies in the fiber $\Theta_{m(s)}$, its derivative is vertical. Clearly $\frac{\mathrm{d}}{\mathrm{d}t}\Gamma_{s,t}\vartheta(t) = T\Gamma_{s,t}\frac{\mathrm{d}}{\mathrm{d}t}\vartheta(t)$. Note that for any given t the vector $\Gamma_{s,t}\vartheta(t)$ is the value at s of the horizontal lift of $m(t)$ that at t goes through $q(t) \in \Theta_{m(t)}$. Then the vector tangent to the horizontal lift belongs to the kernel of the tangent mapping $T\Gamma_{s,t}$. But this vector is the horizontal component of $\frac{\mathrm{d}}{\mathrm{d}t}\vartheta(t)$. In particular, this means that $\frac{\mathrm{d}}{\mathrm{d}t}\Gamma_{s,t}\vartheta(t)_{|t=s}$ is the vertical component of $\frac{\mathrm{d}}{\mathrm{d}t}\vartheta(t)_{|t=s}$. Hence $\mathbf{p}\frac{\mathrm{d}}{\mathrm{d}t}\vartheta(t)_{|t=s} = \frac{\mathrm{D}}{\mathrm{d}t}\vartheta(t)_{|t=s}$. Since (see above) we do not distinguish between $\mathbf{p}\frac{\mathrm{d}}{\mathrm{d}t}\vartheta(t)_{|t=s}$ and $\frac{\mathrm{d}}{\mathrm{d}t}\vartheta(t)_{|t=s}$, the Theorem follows. □

2.3 Connections on Manifolds

Since the tangent bundle TM of a manifold M is a particular case of a vector bundle, all the constructions of Section 2.2 are also valid for tangent bundles.

Definition 2.33. A connection as in Section 2.2, given on the vector bundle TM, is called a *connection* on the manifold M.

Connections on manifolds have special features since here the fiber of the bundle is also a tangent space to the manifold (the base of the bundle). For this reason some constructions are simplified and some operators acquire new properties. In this Section we describe these special features. We use the notation and constructions from Section 2.1.

The vertical subspace $\mathsf{V}_{(m,X)} \subset T_{(m,X)}TM$ turns out to be the tangent space to the fiber of the tangent bundle, i.e. $\mathsf{V}_{(m,X)} = T_X T_m M$. This is why the operator $\mathbf{p}$, introduced by formula (1.2), is an isomorphism of $\mathsf{V}_{(m,X)}$ to $T_m M$.

When we specify a connection H on the tangent bundle, we introduce a subspace $\mathsf{H}_{(m,X)}$ in each $T_{(m,X)}TM$ that is complementary to $\mathsf{V}_{(m,X)}$ in such a way that the collection H satisfies Definition 2.8.

Recall that the tangent bundle of TM is called the second tangent bundle to M and is denoted by TTM or T^2M (see Definition 2.3). So, the connector K sends TTM onto TM and in particular it transforms each $T_{(m,X)}TM$ into $T_m M$. The subspaces $\mathsf{H}_{(m,X)}$ are kernels of K and the mapping K on $\mathsf{V}_{(m,X)}$ coincides with $\mathbf{p}$. As in the general case, $T\pi$ sends $\mathsf{H}_{(m,X)}$ isomorphically onto $T_m M$ and $\mathsf{V}_{(m,X)}$ is the kernel of $T\pi$. Thus for any vector $Y \in T_m M$ at any point $(m, X) \in TM$ there exists a unique vector $Y^l \in \mathsf{V}_{(m,X)}$ such that $\mathbf{p}Y^l = Y$, and a unique vector $Y^T \in \mathsf{H}_{(m,X)}$ such that $T\pi Y^T = Y$.

Definition 2.34. The vector Y^l is called the *vertical lift* of Y at the point (m, X), and the vector Y^T is called the *horizontal lift* of Y at the point (m, X).

Recall that a Euclidean connection and the local connector corresponding to it depend on a trivialization in $\pi^{-1}\mathfrak{U}_\alpha$. We retain the notation $\mathsf{H}^E_{(m,X)}$ for a trivialization by coordinate frames $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}$ (see Sections 1.1 and 2.1). For corresponding objects with respect to other trivializations we shall introduce the special notation below. For the sake of simplicity we denote by $\boldsymbol{\Gamma}_m(\cdot,\cdot)$ the local connector with respect to this trivialization, i.e., $\boldsymbol{\Gamma}_m(\cdot,\cdot) = \mathbf{p}\Gamma_m(\cdot,\cdot)$.

The local connector $\boldsymbol{\Gamma}_m(\cdot,\cdot)$ is a bilinear operator $\boldsymbol{\Gamma}_m : T_mM \times T_mM \to T_mM$. In particular, in this case the condition that $\boldsymbol{\Gamma}_m$ is symmetric is reasonable. The Christoffel symbols of the second kind Γ^k_{ij} are well-defined for indices $i, j, k = 1, \dots, n$. We emphasize that in the natural coordinate systems $\Gamma_m\left(\frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j}\right) = \Gamma^k_{ij}\frac{\partial}{\partial \dot{q}^k}$ while $\boldsymbol{\Gamma}_m\left(\frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j}\right) = \Gamma^k_{ij}\frac{\partial}{\partial q^k}$ where Γ^k_{ij} are Christoffel symbols of the second kind.

Since the operator $g_{\beta\alpha}$ in TM equals $\varphi'_{\beta\alpha}$ and $\Gamma_m(X, Y_1)$ as a quadruple is presented in the form $(m, X, Y_1, \Gamma_m(X, Y_1))$, from formula (2.10) it follows that under a change of coordinates $\varphi_{\beta\alpha}$ the local connector of a connection on a manifold transforms in the following manner

$$\boldsymbol{\Gamma}_m(X, Y_1)^\beta = -\varphi''_{\beta\alpha}(m^\alpha)(X^\alpha, Y_1^\alpha) + \varphi'_{\beta\alpha}(\boldsymbol{\Gamma}_m(X, Y_1)^\alpha). \tag{2.19}$$

The geometric interpretation of formula (2.19) is the same as that given in Remark 2.21.

Proposition 2.35 *The difference* $\boldsymbol{\Gamma}(\cdot,\cdot) - \bar{\boldsymbol{\Gamma}}(\cdot,\cdot)$ *of local connectors* $\boldsymbol{\Gamma}(\cdot,\cdot)$ *and* $\bar{\boldsymbol{\Gamma}}(\cdot,\cdot)$ *of different connections is a* $(1,2)$*-tensor.*

Indeed, by formula (2.19) the difference transforms under coordinate changes by the rule

$$\boldsymbol{\Gamma}_m(\cdot,\cdot)^\beta - \bar{\boldsymbol{\Gamma}}_m(\cdot,\cdot)^\beta = \varphi'_{\beta\alpha}[\boldsymbol{\Gamma}_m(\cdot,\cdot)^\alpha - \bar{\boldsymbol{\Gamma}}_m(\cdot,\cdot)^\alpha].$$

Since the cross-sections of a tangent bundle are vector fields on M, the covariant derivative $\nabla_X Y$ differentiates the vector field Y in the direction of the vector field X and $\frac{\mathrm{D}}{\mathrm{d}t}X(t)$ differentiates the vector field $X(t)$ in the time parameter along the curve $m(t)$ (see Section 2.2).

Equations (2.16) and (2.17) take the forms

$$\nabla_X Y = \left(\frac{\partial Y^k}{\partial q^j}X^j + \Gamma^k_{ij}Y^i X^j\right)\frac{\partial}{\partial q^k}, \tag{2.20}$$

$$\frac{\mathrm{D}}{\mathrm{d}t}Y(t) = \left(\frac{\mathrm{d}Y^k}{\mathrm{d}t} + \Gamma^k_{ij}Y^i\frac{\mathrm{d}m^j}{\mathrm{d}t}\right)\frac{\partial}{\partial q^k}. \tag{2.21}$$

Since in each chart the basis vectors $\frac{\partial}{\partial q^i}$ have constant coordinates in the decomposition with respect to the same basis (the i-th coordinate is 1 and all others equal zero), from formula (2.20) it follows that

$$\nabla_{\frac{\partial}{\partial q^i}} \frac{\partial}{\partial q^j} = \Gamma^k_{ij} \frac{\partial}{\partial q^k}. \tag{2.22}$$

Theorem 2.36 *Let ∇ and $\bar{\nabla}$ be covariant derivatives of two different connections. Then there exists a unique $(1,2)$-tensor $S(\cdot,\cdot)$, determined by the connections, such that for any pair of smooth vector fields X and Y the equality $\nabla_X Y - \bar{\nabla}_X Y = S(X,Y)$ holds.*

Theorem 2.36 follows from formula (2.20) and Proposition 2.35.

If $\frac{\mathrm{D}}{\mathrm{d}t}X(t) = 0$, by analogy with Definition 2.28 we say that $X(t)$ is a *parallel vector field* along the curve $m(t)$. From (2.21) it follows that a parallel vector field satisfies the system of equations

$$\frac{\mathrm{d}Y^k}{\mathrm{d}t} + \Gamma^k_{ij} Y^i \frac{\mathrm{d}m^j}{\mathrm{d}t} = 0. \tag{2.23}$$

A parallel vector field along a curve is an analog of a constant vector field in a linear space. We refer the reader to Section 2.2 where the analogy between a "constant" cross-section of a trivial vector bundle and a parallel cross-section along a curve is described. Notice that the Euclidean connection H^E on a linear space has zero local connector and so the covariant derivative generated by it coincides with the ordinary derivative of a vector field along a vector field or in time along a curve. Thus, on a linear space, parallel vector fields become constant.

Applying Theorem 2.29 to equation (2.23) we obtain that for every smooth curve $m(t)$ and for a specified initial vector $X \in T_{m_0}M$, there exists a unique parallel vector field $X(t)$ with initial condition $X(0) = X$ that is well-defined for all t in the domain of the curve. This vector field is called the *parallel translation* of X along $m(t)$.

Remark 2.37. In the case of a Riemannian manifold M, there is another commonly used trivialization of $\pi^{-1}\mathfrak{U}_\alpha$, namely by a field of orthonormal frames. Let in each tangent space T_mM, $m \in \mathfrak{U}_\alpha$, an orthonormal frame be specified that consists of vectors $e_1, \ldots, e_n$ and let each vector field e_i, $i = 1, \ldots, n$, be smooth. This frame field generates the trivialization in which a point $(m, X^i e_i) \in \pi^{-1}\mathfrak{U}_\alpha$ transforms into the point $(m, (X^1, \ldots, X^n)) \in \mathfrak{U}_\alpha \times \mathbb{R}^n$. The corresponding local connector is called a *tetrad connector* and is denoted by $\mathbf{p}\overset{\circ}{\Gamma}_m(\cdot,\cdot)$ (the term "tetrad" derives from general relativity where $n = 4$). The tetrad Christoffel symbols are denoted by $\overset{\circ}{\Gamma}{}^k_{ij}$ and are defined by the equality $\nabla_{e_i} e_j = \overset{\circ}{\Gamma}{}^k_{ij} e_k$. Since $\mathbf{p}\overset{\circ}{\Gamma}_m(\cdot,\cdot)$ is bilinear, it is uniquely determined by the tetrad symbols. For more detail, see e.g. [57].

2.4 Geodesics

The notion of a parallel vector field along a curve leads to another important notion.

Definition 2.38. A curve $m(t)$ along which its velocity vector field $\dot{m}(t)$ is parallel is called a *geodesic*.

On a manifold with connection the geodesics are analogs of straight lines in a vector space. Indeed, since a parallel vector field along a curve is an analog of a constant vector field in linear space, the property of a curve possessing a parallel velocity vector field is analogous to the property of a curve in a vector space possessing constant velocity. In a vector space the straight lines with natural parametrization, and only these lines, have the latter property.

From Definition 2.38 and the definition of parallel translation it follows that a curve $m(t)$ is a geodesic if and only if at each of its points the equality

$$\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) = 0 \tag{2.24}$$

holds. Equation (2.24) describes an analog of the property that straight lines in linear spaces have zero second derivative.

We now derive the equation of geodesics in local coordinates. For this purpose, in equation (2.23) we replace the coordinates of the vector Y by the coordinates of the vector $\dot{m}(t)$, since in our case the latter is parallel along $m(t)$. Then we obtain

$$\frac{\mathrm{d}^2 m^k}{\mathrm{d}t^2} + \Gamma^k_{ij}\frac{\mathrm{d}m^i}{\mathrm{d}t}\frac{\mathrm{d}m^j}{\mathrm{d}t} = 0. \tag{2.25}$$

Unlike (2.18) and (2.23), (2.25) is a non-linear second order differential equation (recall that (2.18) and (2.23) are linear first order differential equations). This is why we can apply only the most general existence of solution theorem for second order differential equations with smooth right-hand sides, from which we obtain the following statement of local existence and uniqueness of geodesics with given initial data.

Theorem 2.39 *For every point $m \in M$ and every vector $X \in T_m M$ there exists a unique geodesic $m(t)$, with initial conditions $m(0) = m$ and $\dot{m}(0) = X$, that is defined for $t \in [0, \varepsilon)$ where $\varepsilon > 0$ is a sufficiently small positive number.*

Theorem 2.39 is much weaker than existence Theorem 2.29 but it mirrors the physical situation if no additional hypotheses are assumed. For example, on an open manifold (e.g., consisting of only one open chart) the geodesic exists for $t \in [0, \varepsilon)$ where ε is the instant of time when the geodesic reaches the boundary, but it does not exist at any later time.

Definition 2.40. If each geodesic of a connection H exists for $t \in (-\infty, \infty)$, the connection H on M is said to be *complete*.

Let $X \in T_m M$ be a tangent vector at a point m. Denote by $m_X(t)$ the geodesic with initial data $m(0) = m$ and $\dot{m}(0) = X$ (which we know exists for $t \in [0, \varepsilon)$ by Theorem 2.39). Specify a positive number $\lambda < 1$. One can easily see that $m(\lambda t)$ is a geodesic with initial vector λX that exists for $t \in [0, \frac{1}{\lambda}\varepsilon)$. Thus, if X is close enough to the origin, the geodesic $m_X(t)$ exists at $t = 1$.

Definition 2.41. The mapping $\exp : \mathcal{O} \to M$, where $\mathcal{O}$ is a neighborhood of the origin in $T_m M$, is given by the formula $\exp(X) = m_X(1)$, and is called the *exponential mapping* of the connection H.

It is clear that if H is a complete connection, the exponential mapping is well-defined on $T_m M$. Sometimes, when dealing with exponential mappings from tangent spaces at various points of M, we shall use the notation $\exp_m : \mathcal{O}_m \to M$.

Theorem 2.42 *There exists a neighborhood $\mathcal{O}_m$ of the origin in $T_m M$ such that $\exp_m$ is a diffeomorphism of $\mathcal{O}_m$ onto $\exp_m \mathcal{O}_m$ and the exponential mapping is smooth on the neighborhood $\bigcup\limits_{m \in M} \mathcal{O}_m$ of the zero-section in TM.*

A proof of Theorem 2.42 can be found, for example, in [26] and [161].

Notice that the pair $(\mathcal{O}_m, \exp_m)$ satisfies the definition of chart. This pair is called the *normal chart* (or *normal neighborhood*) of the connection H at the point m. In this chart at m the connection space $\mathsf{H}_{(m,X)}$ at each $X \in T_m M$ coincides with the Euclidean connection space $\mathsf{H}^E_{(m,X)}$ and so $\boldsymbol{\Gamma}_m(\cdot, \cdot) = 0$. Hence in a normal chart at m all Christoffel symbols of the second kind $\Gamma^k_{ij}(m)$ for H at this point are equal to zero.

Suppose that the connection is complete and $\mathcal{O}_m$ is the maximal domain on which $\exp_m$ is one-to-one, i.e., such that the exponential map is one-to-one on $\mathcal{O}_m$ but not on the boundary $\partial\mathcal{O}_m$ in $T_m M$.

Definition 2.43. The set $\partial\mathcal{O}_m \subset T_m M$ is called the *cut locus* corresponding to the point m. The same term is also used to designate the image of $\partial\mathcal{O}_m$ under the mapping $\exp_m$.

All points of M besides the cut locus belong to the image of $\mathcal{O}_m$ under the diffeomorphism $\exp_m$. From this it follows that each manifold can be constructed from an open ball in a vector space by "gluing" the points of the boundary (according to a rule, determined by the manifold) so that the corresponding cut locus is obtained (see [140]).

Let the points m_0 and m_1 be connected by a geodesic $a(\cdot)$ of a connection H. This means that $m_1 = \exp_{m_0} X$ for some vector $X \in T_{m_0} M$.

Definition 2.44. If the differential $d_X \exp : T_X T_{m_0} M \to T_{m_1} M$ at X is degenerate, we say that $m_1 = \exp_{m_0} X$ is *conjugate* with m_0 along the geodesic $a(\cdot)$ joining them.

2.5 Curvature and Torsion Tensors

Let X and Y be smooth vector fields on a manifold M with connection. These vector fields determine a transformation of an arbitrary smooth vector field Z by the formula

$$\mathsf{R}_{XY}Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z. \tag{2.26}$$

Observe that the value $\mathsf{R}_{XY}Z$ at $m \in M$ depends only on the values of the vector fields X, Y, Z at m (it does not depend on their values in a neighborhood of m), i.e., $\mathsf{R}_{XY}Z$ is a tensor. (In particular this means that, in spite of the definition, (2.26) is well-defined for non-smooth vector fields X, Y and Z.)

Definition 2.45. $\mathsf{R}_{XY}Z$ is called the *curvature tensor*.

If $\mathsf{R}_{XY}Z = 0$ for all $X, Y, Z,$, the connection is called *flat*. An example of a flat connection is a Euclidean connection of any coordinate system.

The curvature is a $(1,3)$-tensor and its description as a polylinear form takes the form $\mathsf{R}(\alpha, X, Y, Z) = \alpha(\mathsf{R}_{XY}Z)$, where α is a covector field (1-form). We denote the components of the curvature tensor by R^i_{jkl}.

For two vector fields X and Y on M one can consider a third vector field

$$\mathsf{T}(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]. \tag{2.27}$$

Observe that the value $\mathsf{T}(X,Y)$ at $m \in M$ depends only on the values of X and Y at m (it does not depend on their values in a neighborhood of m), i.e., $\mathsf{T}(X,Y)$ is a tensor.

Definition 2.46. $\mathsf{T}(X,Y)$ is called the *torsion tensor*.

The curvature and torsion tensors together "measure" how the vector can be transformed under parallel translation along a closed infinitesimal loop (for details see, e.g., [26]).

Torsion is a $(1,2)$-tensor, i.e., its description as a polylinear form takes the form $\mathsf{T}(\alpha, X, Y) = \alpha(\mathsf{T}(X,Y))$ where α is a covector field (1-form). Denote the components of T by the symbols T^k_{ij}. To calculate these components we substitute into (2.27) the coordinate expressions of $\nabla_X Y$ and $\nabla_Y X$ from formula (2.20) as well as the coordinate expression for $[X,Y]$ from Proposition 1.7. We then obtain

$$\begin{aligned}\mathsf{T}(X,Y) &= \left\{ \left(\frac{\partial Y^k}{\partial q^j} X^j + Y^i X^j \Gamma^k_{ij} \right) - \left(\frac{\partial X^k}{\partial q^j} Y^j + X^i Y^j \Gamma^k_{ji} \right) \right. \\ &\qquad \left. - \left(\frac{\partial Y^k}{\partial q^j} X^j - \frac{\partial X^k}{\partial q^j} Y^j \right) \right\} \frac{\partial}{\partial q^k} \\ &= Y^i X^j \Gamma^k_{ij} - X^i Y^j \Gamma^k_{ji}.\end{aligned}$$

Hence,

$$T^k_{i,j} = \Gamma^k_{ij} - \Gamma^k_{ji}. \tag{2.28}$$

Formula (2.28) immediately yields:

Proposition 2.47 *The equality $\mathsf{T} = 0$ holds at all points $m \in M$ if and only if in all charts $\Gamma^k_{ij} = \Gamma^k_{ji}$, i.e., the local connector $\boldsymbol{\Gamma}_m(\cdot, \cdot)$ is a symmetric bilinear operator.*

2.6 Riemannian Connections. The Levi-Civitá Connection

From all the connections on a Riemannian manifold M we select one whose covariant derivative properties are the closest to those of the ordinary derivative in Euclidean space.

If on a manifold M a Riemannian metric and a connection are given independently, one should not expect, for the covariant derivative, to find an analog of the Leibnitz formula for differentiating the inner product. Nevertheless for every Riemannian manifold there exists a class of connections having this property.

Let a Riemannian or semi-Riemannian metric $\langle \cdot, \cdot \rangle$ be given on M. For two smooth vector fields Y and Z on M we consider the smooth function $\langle Y, Z \rangle$ that assigns the value of the Riemannian inner product $\langle Y_m, Z_m \rangle$ of the vectors of Y and Z at m to the point m. We find the derivative $X\langle Y, Z \rangle$ of the function $\langle Y, Z \rangle$ in the direction of a smooth vector field X.

Definition 2.48. A connection on M is said to be *Riemannian* if for all smooth vector fields X, Y and Z on M the following equality holds:

$$X\langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle. \tag{2.29}$$

Taking into account the interrelation between ∇ and $\frac{\mathrm{D}}{\mathrm{d}t}$ (see Remark 2.27) one can easily derive the following version of formula (2.29) for $\frac{\mathrm{D}}{\mathrm{d}t}$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle Y(t), Z(t) \rangle = \left\langle \frac{\mathrm{D}}{\mathrm{d}t} Y(t), Z(t) \right\rangle + \left\langle Y(t), \frac{\mathrm{D}}{\mathrm{d}t} Z(t) \right\rangle, \tag{2.30}$$

where $Y(t)$ and $Z(t)$ are smooth vector fields along a smooth curve $m(t)$.

An existence theorem for Riemannian connections will be proved below (see Remark 2.55).

Specify a Riemannian connection on a Riemannian manifold M.

Theorem 2.49 *Let $Y(t)$ and $Z(t)$ be parallel vector fields along a smooth curve $m(t)$. Then $\langle Y(t), Z(t) \rangle = \mathrm{const}$.*

Proof. By the definition of a parallel vector field, $\frac{\mathrm{D}}{\mathrm{d}t}Y(t) = 0$ and $\frac{\mathrm{D}}{\mathrm{d}t}Z(t) = 0$. Having substituted these expressions into (2.30) we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle Y(t), Z(t)\rangle = \left\langle \frac{\mathrm{D}}{\mathrm{d}t}Y(t), Z(t)\right\rangle + \left\langle Y(t), \frac{\mathrm{D}}{\mathrm{d}t}Z(t)\right\rangle = 0.$$

This means that the function $\langle Y(t), Z(t)\rangle$ is constant. □

Corollary 2.50 *If $Y(t)$ is a parallel vector field along a smooth curve $m(t)$, $\|Y(t)\| = \mathrm{const}$.*

Indeed, $\|Y(t)\| = \sqrt{\langle Y(t), Y(t)\rangle}$ and the assertion of Corollary 2.50 follows from Theorem 2.49.

Corollary 2.51 *If $Y(t)$ and $Z(t)$ are parallel vector fields along a smooth curve $m(t)$, the cosine of the angle between those vectors is constant.*

Since the cosine of the angle between $Y(t)$ and $Z(t)$ equals $\frac{\langle Y(t), Z(t)\rangle}{\|Y(t)\|\|Z(t)\|}$, the assertion of Corollary 2.51 follows from Theorem 2.49 and Corollary 2.50.

Definition 2.52. The functions $\Gamma_{ij,k} = \langle \nabla_{\partial q^i}\partial q^j, \partial q^k\rangle$ in a chart of a Riemannian manifold M are called *Christoffel symbols of the first kind.*

We now describe the interrelation between Christoffel symbols of the first and second kinds. By formula (2.22) $\nabla_{\frac{\partial}{\partial q^i}}\frac{\partial}{\partial q^j} = \Gamma^l_{ij}\frac{\partial}{\partial q^l}$. Thus

$$\Gamma_{ij,k} = \left\langle \Gamma^l_{ij}\frac{\partial}{\partial q^l}, \frac{\partial}{\partial q^k}\right\rangle = g_{lk}\Gamma^l_{ij}. \tag{2.31}$$

Applying the same arguments as in the derivation of formula (1.21), from (1.20) we obtain

$$\Gamma^k_{ij} = g^{lk}\Gamma_{ij,l}. \tag{2.32}$$

In particular, if the torsion tensor equals zero, i.e., $\Gamma^k_{ij} = \Gamma^k_{ji}$, then also $\Gamma_{ij,k} = \Gamma_{ji,k}$.

Note that here we use only the fact that the matrix (g_{ij}) is invertible, not that it is positive-definite. Thus formula (2.32) is well-defined both for Riemannian and semi-Riemannian metrics.

Lemma 2.53 (The principal lemma of Riemannian geometry) *On every manifold M with Riemannian or semi-Riemannian metric $\langle\cdot,\cdot\rangle$ there exists a unique Riemannian connection whose torsion tensor equals zero at all $m \in M$.*

Proof. Recall that by definition $g_{ij} = \left\langle \frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j}\right\rangle$. Since the connection that we are looking for is Riemannian, from formula (2.29) it follows that

$$\frac{\partial}{\partial q^l}g_{ij} = \frac{\partial}{\partial q^l}\left\langle \frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j}\right\rangle = \left\langle \nabla_{\frac{\partial}{\partial q^l}}\frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j}\right\rangle + \left\langle \frac{\partial}{\partial q^i}, \nabla_{\frac{\partial}{\partial q^l}}\frac{\partial}{\partial q^j}\right\rangle.$$

So, taking into account Definition 2.52, we obtain $\frac{\partial}{\partial q^l} g_{ij} = \Gamma_{li,j} + \Gamma_{lj,i}$. Considering all rearrangements of the given indices i, j, l we obtain a system of three equations of the same kind as above:

$$\begin{cases} \frac{\partial}{\partial q^l} g_{ij} = \Gamma_{li,j} + \Gamma_{lj,i} \\ \frac{\partial}{\partial q^i} g_{lj} = \Gamma_{il,j} + \Gamma_{ij,l} \\ \frac{\partial}{\partial q^j} g_{il} = \Gamma_{ji,l} + \Gamma_{jl,i}. \end{cases} \tag{2.33}$$

Recall that the torsion tensor equals zero, i.e., the Christoffel symbols of the first kind are symmetric in the first two indices. The system (2.33) of three linear algebraic equations has three unknowns. Adding the second equation to the third one and subtracting the first one from the sum, we obtain

$$\Gamma_{ij,l} = \frac{1}{2}\left(\frac{\partial}{\partial q^i} g_{lj} + \frac{\partial}{\partial q^j} g_{li} - \frac{\partial}{\partial q^l} g_{ij}\right). \tag{2.34}$$

Then by formula (2.32)

$$\Gamma_{ij}^k = \frac{1}{2}\left(\frac{\partial}{\partial q^i} g_{lj} + \frac{\partial}{\partial q^j} g_{li} - \frac{\partial}{\partial q^l} g_{ij}\right) g^{kl}. \tag{2.35}$$

Formula (2.35) uniquely determines the Christoffel symbols of the second kind. From this it follows that the connection we are looking for is unique. The existence is proved by an elementary verification that the connection with Christoffel symbols (2.35) has the properties described in the hypothesis. □

Definition 2.54. The connection whose existence is asserted in Lemma 2.53 is called the *Levi-Civitá connection* of the metric $\langle \cdot, \cdot \rangle$.

It is easy to see that in the Euclidean space $\mathbb{R}^n$ the Levi-Civitá connection of the standard inner product coincides with the Euclidean connection of the standard coordinate system.

Remark 2.55. If a connection is Riemannian but the torsion is not zero, system (2.33) consists of three equations but has six unknowns. This system has an infinite set of solutions, each of them determining a Riemannian connection.

Remark 2.56. The tetrad Christoffel symbols (see Remark 2.37) of the Levi-Civitá connection are determined by the formula

$$\overset{\circ}{\Gamma}{}_{ij}^k = \frac{1}{2}\left(c_{kj}^i + c_{ki}^j + c_{ij}^k\right), \tag{2.36}$$

where c_{pq}^l can be found from the equalities $[e_p, e_q] = c_{pq}^l e_l$, see [57].

The next property of the Levi-Civitá connection follows from the fact that its torsion tensor equals zero and so the property does not hold for other Riemannian connections.

Let $\gamma(t,s)$ be a smooth mapping from the rectangle $[a,b]\times(c,d)$ into M. Then one can consider the vector fields $\frac{\partial}{\partial t}$ and $\frac{\partial}{\partial s}$ on $\gamma([a,b]\times(c,d))$.

Lemma 2.57 (Lemma on the second covariant derivative)

$$\nabla_{\frac{\partial}{\partial t}}\frac{\partial}{\partial s}=\nabla_{\frac{\partial}{\partial s}}\frac{\partial}{\partial t}.$$

Proof. By construction, s and t are coordinates on $[a,b]\times(c,d)$. Hence on $\gamma([a,b]\times(c,d))$ the fields $\frac{\partial}{\partial s}$ and $\frac{\partial}{\partial t}$ commute, i.e., $\left[\frac{\partial}{\partial t},\frac{\partial}{\partial s}\right]=0$ (see Section 1.7). Then $\nabla_{\frac{\partial}{\partial t}}\frac{\partial}{\partial s}-\nabla_{\frac{\partial}{\partial s}}\frac{\partial}{\partial t}=(\Gamma^k_{ij}-\Gamma^k_{ji})\frac{\partial}{\partial q^k}=\mathsf{T}\left(\frac{\partial}{\partial t},\frac{\partial}{\partial s}\right)$ and the assertion of the Lemma follows from the fact that the torsion tensor T equals zero. □

Lemma 2.57 is an analog of the classical equality $\frac{\partial^2}{\partial x\partial y}=\frac{\partial^2}{\partial y\partial x}$.

For the curvature tensor of the Levi-Civitá connection of a Riemannian or semi-Riemannian metric we consider the following constructions. As above denote by R^i_{jkl} the components of the curvature tensor R. Contract R by the only contravariant and the second covariant indices (see Section 1.5). The result is a tensor Ric called the *Ricci curvature*. Its components take the form $R_{jl}=R^k_{jkl}$. If $\mathsf{R}=0$, it is evident that $\mathrm{Ric}=0$, but not vice versa. Ric is a symmetric $(0,2)$-tensor.

Let $\widehat{\mathrm{Ric}}$ be the $(1,1)$-tensor with components R^l_j that is physically equivalent to Ric. The contraction of $\widehat{\mathrm{Ric}}$, i.e., the scalar $S=R^j_j$, is called the *Gaussian* or *scalar curvature*. If $\mathrm{Ric}=0$, then $S=0$ but not vice versa.

Nevertheless, if $\dim M=2$ the Gaussian curvature determines both the Ricci curvature and the curvature tensor and if $\dim M=3$ the Ricci curvature determines the curvature tensor. If $\dim M\geq 4$ no such determinations are valid.

Definition 2.58. The operator $\nabla^2=\nabla\nabla^*$, where ∇^* is the operator conjugate to the operation of covariant derivation of the Levi-Civitá connection ∇, is called the *Laplace-Beltrami operator*.

In local coordinates of a chart the operator ∇^2 is described by the formula $\nabla^2=g^{ij}\nabla_i\nabla_j=-g^{ij}\Gamma^k_{ij}\frac{\partial}{\partial q^k}+g^{ij}\frac{\partial^2}{\partial q^i\partial q^j}$ where ∇_k is the covariant derivative in the direction of $\frac{\partial}{\partial q^k}$ and g^{ij} are the components of the metric tensor (g^{ij}). From this one can easily see that in a Euclidean space $\mathbb{R}^n$ with standard basis, ∇^2 coincides with the ordinary Laplacian. Note also that the above coordinate representation of ∇^2 defines its action on functions.

In general the Laplace-Beltrami operator does not coincide with the Laplace-de Rham operator $\varDelta=\mathrm{d}\delta+\delta\mathrm{d}$ (see Definition 1.73) in spite of the fact that in $\mathbb{R}^n$, modulo the sign, they both give the Laplacian. On functions both operators on all Riemannian manifolds take the same value. In the general case of differential forms (polyvectors) on manifolds the relation between the operators is described by the so-called *Weitzenbök formulae* (a special formula exists for each degree of the form), see [135]. For the material below we need the following Weitzenbök formula for 1-forms (and so for vector fields):

$$\Delta X = -\nabla^2 X + \widehat{\mathrm{Ric}} \circ X. \tag{2.37}$$

2.7 Connections on Principal Bundles

Let $\underline{G}$ be a principal bundle with fiber (structure group) G over base M. The fiber at $m \in M$ will be denoted G_m. Recall that G_m is homeomorphic to the group G. In the tangent space $T_g\underline{G}$ to $\underline{G}$ at $g \in \underline{G}$, as is the case for any bundle, we can consider the subspace that consists of the vectors tangent to the fiber $G_{\pi g}$. As usual we call this space the *vertical subspace* and denote it by $\mathsf{V}_g \subset T_g\underline{G}$. The vectors in V_g are said to be*vertical*.

The collection of vertical subspaces at all points forms the bundle $\mathsf{V} \to \underline{G}$ with fibers V_g.

Theorem 2.59 *The bundle* $\mathsf{V} \to \underline{G}$ *is trivial.*

Proof. Every point $g \in \underline{G}$ determines a diffeomophism of the group G onto the fiber $G_{\pi g}$ (which will be denoted by the same symbol g) by the formula $g \circ G$, where $\circ$ is the right action of G on $\underline{G}$ (see Section 1.3). It is clear that the diffeomorphism g sends the unit $e \in G$ to the point $g \in \underline{G}$. Since this is a diffeomorphism, the tangent map $Tg : TG \to TG_g$ is a linear isomorphism of $T_eG = \mathfrak{g}$ onto the tangent space to the fiber $G_{\pi g}$ at g, i.e., onto V_g. Thus every vertical subspace V_g is linearly isomorphic to the Lie algebra $\mathfrak{g}$ and the isomorphism smoothly depends on the point $g \in \underline{G}$. Hence, having specified a vector $X \neq 0 \in \mathfrak{g}$, we obtain the smooth vector field $\bar{X}_g = TgX \neq 0$ on $\underline{G}$. In particular, taking vectors of a basis in $\mathfrak{g}$, we obtain a basis at every V_g. So, we have represented $\mathsf{V} \to \underline{G}$ in the form of a direct product $\underline{G} \times \mathfrak{g}$. □

Definition 2.60. The vector field $\bar{X}$ on $\underline{G}$, constructed in the proof of Theorem 2.59 from the vector $X \in \mathfrak{g}$, is called a *fundamental vector field*.

Thus the fundamental vector fields trivialize the bundle $\mathsf{V} \to \underline{G}$ since they determine the frames in the fibers of V_g.

Recall that for a vector bundle Θ we also constructed vertical subspaces $\mathsf{V}_{(m,\vartheta)} \subset T_\vartheta \Theta_m$ that were sent onto the fibers $\Theta_{\pi q}$ by the linear isomorphism $\mathbf{p}$. In the case of a principal bundle an isomorphism like $\mathbf{p}$ does not exist since the fibers of the bundle are not vector spaces. However, Theorem 2.59 provides us with something that was not available in the case of vector bundles: all V_g are in a standard way isomorphic to a unique vector space $\mathfrak{g}$ while in the case of a vector bundle vertical subspaces at the points of Θ_m were sent onto "their own" fiber Θ_m.

Notation 2.61 *The isomorphism* $\mathsf{V}_g \to \mathfrak{g}$*, described above, is denoted by* p.

Definition 2.62. We say that a connection H is given on a principal bundle $\underline{G}$ if to every $T_g\underline{G}$ is associated a subspace H_g that is complementary to V_g, smoothly depends on the point $g \in \underline{G}$ and is such that the connection H is invariant with respect to the right action of the group G on $\underline{G}$ (i.e., for every element h of G and every $g \in \underline{G}$ the equality $TR_h\mathsf{H}_g = \mathsf{H}_{g\circ h}$ holds where $R_h(g) = g \circ h$ is the right action of h on $\underline{G}$). The subspaces H_g comprising a connection are called *horizontal*, as are the vectors belonging to them.

As in the case of a vector bundle, $Y \in T_g\underline{G}$ is uniquely represented as the sum $Y = \mathsf{H}Y + \mathsf{V}Y$ where $\mathsf{H}Y \in \mathsf{H}_g$ and $\mathsf{V}Y \in \mathsf{V}_g$.

Definition 2.63. The mapping $\varphi : T\underline{G} \to \mathfrak{g}$ whose value at $Y \in T\underline{G}$ is given by the formula $\varphi(Y) = \mathsf{pV}Y$ is called the *connection form* of H .

It is clear that the connection form is a direct analog of the connector (connection map) on vector bundles. It turns out that connection forms have a much richer collection of properties than connectors and their use allows one to obtain much deeper results. We refer the reader, e.g., to [26] and [161] for a more detailed exposition of the theory of general principal bundles and their connection forms. Here we only describe some objects and constructions that are used later.

For every k-form α on $\underline{G}$ with values in $\mathfrak{g}$ the so-called *covariant differential*

$$\mathrm{D}\alpha(\cdot, \ldots, \cdot) = \mathrm{d}\alpha(\mathsf{H}\cdot, \ldots, \mathsf{H}\cdot) \tag{2.38}$$

is introduced where, as above, the symbol H denotes the projection onto the connection subspace (i.e., H of a vector is the horizontal component of the vector).

Definition 2.64. The 2-form $\Phi = \mathrm{D}\varphi = \mathrm{d}\varphi(\mathsf{H}\cdot, \mathsf{H}\cdot)$ is called the *curvature form* of the connection H.

Since φ takes values in the Lie algebra $\mathfrak{g}$, the composition $[\varphi, \varphi]$ of the operators φ and bracket $[\cdot, \cdot]$ is well-defined.

The so-called *Bianchi identity*

$$\mathrm{D}\Phi = 0 \tag{2.39}$$

and the *structure equation*

$$\mathrm{d}\varphi = -\frac{1}{2}[\varphi, \varphi] + \Phi \tag{2.40}$$

hold (for the proofs see, e.g., [26]). Note that for a matrix group G

$$-\frac{1}{2}[\varphi, \varphi] = -\varphi^2 \tag{2.41}$$

(for details see, e.g., [26, 146]).

As in the case of vector bundles, by construction $T\pi : \mathsf{H}_g \to T_{\pi g}M$ is a linear isomorphism. Via this we also obtain the well-defined notion of a horizontal lift $\tilde{X}$ of a vector field X from the base M onto $\underline{G}$: $\tilde{X}_g = T\pi^{-1}_{|\mathsf{H}_g} X_{\pi g}$. Consider a smooth curve $m(t)$ on the base. The pull-back of the bundle $\underline{G}$ over this curve is a manifold on which the vector field $\widetilde{\dot{m}}(t)$, the horizontal lift of the velocity vector field $\dot{m}(t)$ of $m(t)$, is given. Take a point $g_0 \in G_{m(0)}$ and consider the integral curve $g(t)$ of $\widetilde{\dot{m}}(t)$ with initial data $g(0) = g_0$.

Definition 2.65. The curve $g(t)$ is called the *parallel translation* of g_0 along the curve $m(t)$.

Let Θ be a bundle with fiber F associated with a principal bundle $\underline{G}$. As on any other bundle, we can consider vertical subspaces in the tangent spaces to Θ, i.e., the subspaces tangent to fibers. The mapping $\lambda : \underline{G} \times F \to \Theta$ (see Notation 1.36) sends horizontal subspaces on $\underline{G}$ into subspaces of tangent spaces to Θ complementary to vertical subspaces. The collection of subspaces that we obtain in this way is called the *connection on* Θ. The parallel translation in the associated bundle Θ is defined by analogy with Definition 2.65.

If G is $GL(k, \mathbb{R})$, or one of its subgroups, with the standard action on $\mathbb{R}^k$, the associated bundle is a vector bundle.

Proposition 2.66 *Every connection on a vector bundle by means of Section 2.2 is an image of some connection on the corresponding principal bundle under the mapping* λ*.*

Now let us consider what is for us the most important case of a principal bundle, the frame bundle BM (see Definition 1.37). Recall that the tangent bundle TM is associated with BM, i.e., by the last statement every connection on the manifold M is obtained from some connection on BM as explained above.

We introduce a connection H on BM by means of Definition 2.62. Consider the bundle $\mathsf{H} \to BM$ whose fiber at every point $b \in BM$ is H_b.

Theorem 2.67 *The bundle* $\mathsf{H} \to BM$ *is trivial.*

Proof. Specify a vector $X \in \mathbb{R}^n$, i.e., a column with coordinates $X^1, \ldots, X^n$. Every $b \in BM$, i.e., a frame $b = e_1, \ldots, e_n$ in $T_{\pi b}M$, can be considered as a linear mapping $b : \mathbb{R}^n \to T_{\pi b}M$ defined by the formula $bX = X^i e_i$ (see Section 1.3). Denote by $\mathsf{E}_b(X)$ the vector in H_b of the form $\mathsf{E}_b(X) = T\pi^{-1}_{|\mathsf{H}_b} bX$. One can easily see that the mapping $\mathsf{E}_b : \mathbb{R}^n \to \mathsf{H}_b$ is a linear isomorphism and smoothly depends on $b \in BM$. In particular a basis in $\mathbb{R}^n$ determines a corresponding basis in every H_b so that, using coordinate decomposition of vectors of H_b with respect to this basis, we can represent $\mathsf{H} \to BM$ in the form $BM \times \mathbb{R}^n$. $\square$

Definition 2.68. The vector field $\mathsf{E}(X)$ on BM that is equal to $\mathsf{E}_b(X)$ at $b \in BM$ is called the *basic vector field.*

It is clear that basic vector fields are smooth. The basic vector fields trivialize the bundle $\mathsf{H} \to BM$ just as the fundamental vector fields trivialize $\mathsf{V} \to BM$.

Theorem 2.69 *The tangent bundle TBM is trivial.*

This statement is a corollary to Theorem 2.59 and Theorem 2.67. Indeed, by construction, for every $b \in BM$ we have $T_bBM = \mathsf{H}_b \oplus \mathsf{V}_b$, but by the theorems mentioned above the bundles $\mathsf{H} \to BM$ and $\mathsf{V} \to BM$ are trivial.

We introduce a mapping from TBM to $\mathbb{R}^n$ as follows. For $b \in BM$, where $b = (e_1, \dots, e_n)$ is a basis in $T_{\pi b}M$, consider a vector $X \in T_bBM$. Then $T\pi X \in T_{\pi b}M$ has the coordinate decomposition $T\pi X = \omega^i e_i$. The mapping $X \mapsto (\omega^1, \dots, \omega^n) \in \mathbb{R}^n$ is considered as a 1-form ω with values in $\mathbb{R}^n$ and is called the *displacement form*. Note that the displacement form ω exists without having to introduce a connection on BM. But if a connection H on BM is specified, we can consider the covariant differential $\mathrm{D}\omega = \mathrm{d}\omega(\mathsf{H}\cdot, \mathsf{H}\cdot)$ (see formula (2.38)) of ω with respect to this connection (cf. Definition 2.64).

Definition 2.70. The 2-form $\Omega = \mathrm{D}\omega$ is called the *torsion form* of H.

The curvature form Φ and the torsion form Ω determine the curvature and torsion tensors, respectively (see details, e.g., in [26]). In addition to (2.40) there is another structure equation for H on BM in terms of ω and Ω:

$$\mathrm{d}\omega = -\varphi\omega + \Omega \tag{2.42}$$

where φ is the connection form. The composition $\varphi\omega$ makes sense since φ is a transformation of $\mathbb{R}^n$ (a matrix from $\mathfrak{gl}(n, \mathbb{R})$) and ω takes values in $\mathbb{R}^n$ (see [26, 146] for details).

At the moment we have two constructions of a parallel vector field along a curve on a manifold: by general Definition 2.28 applied to connections on manifolds (see Section 2.3) and by analogy with Definition 2.65 for the case of associated bundles. Here we describe a third construction.

Let $m(t)$ be a smooth curve on M and $b(t)$ be the parallel translation of a basis $b_0 = b(0)$ in the tangent space $T_{m(0)}M$ along $m(t)$ by means of Definition 2.65. As said above, every basis $b(t)$ is a linear isomorphism $b(t) : \mathbb{R}^n \to T_{m(t)}M$. Let $X_0 \in T_{m(0)}M$ and consider the vector field $X(t) = b(t)(b_0^{-1}X_0) \in T_{m(t)}M$ along $m(t)$. Notice that $X(0) = b_0(b_0^{-1}X_0) = X_0$.

Proposition 2.71 *The vector field $X(t)$ along $m(t)$, introduced above, does not depend on the initial basis b_0 of the parallel translation $b(t)$.*

Proof. Specify another basis $\bar{b}_0$ in $T_{m(0)}M$ and let $\bar{b}(t)$ be the parallel translation of this basis along $m(t)$. It is clear that there exists an $h \in GL(n, \mathbb{R})$ such that $\bar{b}_0 = b_0 \circ h$ where $\circ$ denotes the right action of h on BM. Since by definition a connection H on BM is invariant with respect to the right action of $GL(n, \mathbb{R})$ (see Definition 2.62), one can easily see that $\bar{b}(t) = b(t) \circ h$. Then $\bar{b}(t)(\bar{b}_0^{-1}X_0) = b(t) \circ h((b_0 \circ h)^{-1}X_0) = b(t) \circ h((h^{-1} \circ b_0^{-1})X_0) = X(t)$. $\square$

Thus the formula $X(t) = b(t)(b_0^{-1}X_0)$ uniquely determines the translation of X_0 along $m(t)$. The following statement holds:

Theorem 2.72 *Let a connection on M be obtained from a connection on BM as described above. Then the parallel translation by means of Definition 2.28 applied to connections on manifolds, the parallel translation introduced analogously to Definition 2.65 for the case of associated bundles, and the translation by formula $b(t)(b_0^{-1}X_0)$ coincide.*

Remark 2.73. Let a connection on M be obtained from a connection on BM as described above. It is clear that the geodesics of this connection, and only these geodesics, are projections onto M of integral curves of basic vector fields on BM (see Definition 2.68).

Consider the bundle of orthonormal frames OM on a Riemannian manifold M. This is a principal bundle with a structure group $O(n)$ of orthogonal matrices. If a connection is given on BM, one can consider the spaces H_b of this connection at the points $b \in OM$. However, this collection of subspaces becomes a connection on OM only if it is invariant with respect to the right action of $O(n)$ on OM. In addition, a connection on a Riemannian manifold M is Riemannian if and only if it is obtained from some connection on OM as the image of the mapping λ.

Among the connections on OM there is unique connection with zero torsion form. This connection corresponds to the Levi-Civitá connection M. A detailed description of this material can be found in [26] and [161].

2.8 A Connection on the Total Space of a Vector Bundle

In this section we describe a construction that allows one to create a connection on the total space of a vector bundle (as on a manifold) from a connection of the bundle and a connection on the base (again as on a manifold). A more detailed presentation of this material (at least for the case of a tangent bundle) can be found in [23].

Denote by $\pi : \Theta \to M$ the vector bundle and by Θ_m its fiber at $m \in M$. Let a connection H^π be given on Θ by means of Section 2.2. Denote the connector of this connection by $K^\pi : T\Theta \to \Theta$.

In order to avoid confusion, in this section we denote the projection of a tangent bundle TM on M by $\tau : TM \to M$. Let a connection be given on the manifold M; for this connection we introduce the notation H^τ and denote its connector by $K^\tau : T^2M \to TM$ (recall that according to Section 2.3 a connection on a manifold M is a connection on its tangent bundle TM).

Using connections H^τ and H^π, we construct a connection H^Θ on the total space of Θ (i.e., on the manifold Θ) as follows. We define the connector $K : T^2\Theta \to T\Theta$ of this connection by the formula $K = K^H \oplus K^V$ with $K^H : T^2\Theta \to \mathsf{H}^\pi$ and $K^V : T^2\Theta \to \mathsf{V}$ where V is the vertical subspace at the corresponding point (recall that the fibers of the bundle V over Θ are the subspaces in $T\Theta$ that are tangent to the fibers of Θ, see above). These connectors we define in the form $K^H = \Gamma^\pi \circ K^\tau \circ T^2\pi$ where $\Gamma^\pi = T\pi^{-1}$ is a linear isomorphism of tangent spaces to M onto H^π (see Lemma 2.12) while $K^V = \mathbf{p}^{-1} \circ K^\pi \circ TK^\pi$ where $\mathbf{p} : \mathsf{V}_q \to \Theta_{\pi q}$ is the natural isomorphism of the tangent space V_q to the vector space $\Theta_{\pi q}$ onto the space $\Theta_{\pi q}$ introduced in (1.2) (see also (2.11)).

The covariant derivative on a manifold Θ corresponding to K will be denoted by $\frac{\mathrm{D}}{\mathrm{d}t} = K \circ \frac{\mathrm{d}}{\mathrm{d}t}$. By construction, $\frac{\mathrm{D}}{\mathrm{d}t} = \frac{\mathrm{D}}{\mathrm{d}t}^H + \frac{\mathrm{D}}{\mathrm{d}t}^V$ where $\frac{\mathrm{D}}{\mathrm{d}t}^H = K^H \circ \frac{\mathrm{d}}{\mathrm{d}t}$ and $\frac{\mathrm{D}}{\mathrm{d}t}^V = K^V \circ \frac{\mathrm{d}}{\mathrm{d}t}$. Notice the following important feature: $T\pi\frac{\mathrm{D}}{\mathrm{d}t} = T\pi\frac{\mathrm{D}}{\mathrm{d}t}^H$ and it is equal to the covariant derivative of the connection H^τ on M. From this it follows that for a parallel translation $X(t)$ along a curve $q(t)$ in Θ with respect to the connection H^Θ, the vector field $T\pi X(t)$ along the curve $m(t) = \pi q(t)$ in M is the parallel translation with respect to the connection H^τ. In particular the geodesics of the connection H^Θ on Θ are projected by π onto the geodesics of the connection H^τ on M.

2.9 Second Order Tangent Vectors and Connections

Definition 2.74. A *second order tangent vector* to a manifold M at a point $m \in M$ is a second order differential operator on M at m with zero constant term and a symmetric matrix of coefficients at second order derivatives in local coordinates. The linear space of second order tangent vectors at a point $m \in M$ is called the *second order tangent space* and is denoted by $\tau_m M$.

Usually the fact that the constant term of a second order differential operator $\mathcal{A}$ equals zero is expressed by the condition $\mathcal{A}1 = 0$ where 1 is the function identically equal to unity.

Recall that a vector (i.e., a first order vector) may be considered as a first order differential operator without constant term (the derivative in the direction of a vector, see Section 1.1). By analogy, second order differential operators without constant terms are called *second order tangent vectors*.

The set of all second order tangent vectors has the structure of a fiber bundle with fiber $\tau_m M$ and is called the *second order tangent bundle* τM.

In local coordinates every second order tangent vector $\mathcal{A} \in \tau_m M$ is uniquely represented in the form: $\mathcal{A}x = b^i \frac{\partial}{\partial q^i} + \beta^{ij} \frac{\partial^2}{\partial q^j \partial q^j}$ where the matrix (β^{ij}) is symmetric since $\frac{\partial^2 f}{\partial q^j \partial q^j} = \frac{\partial^2 f}{\partial q^j \partial q^i}$ for a smooth real-valued f. Thus $\frac{\partial}{\partial x^i}$ and $\frac{\partial^2}{\partial x^i \partial x^j}$, $i, j = 1, 2, \ldots, n$ form a basis in $\tau_m M$. The transformation of the components of a second order vector under coordinate changes is described by the formulae (see, e.g., [148])

$$\beta^{i'j'} = \frac{\partial q^{i'}}{\partial q^i}\frac{\partial q^{j'}}{\partial q^j}\beta^{ij},$$
$$b^{k'} = \frac{\partial q^{k'}}{\partial q^k}b^k + \frac{\partial^2 q^{k'}}{\partial q^i \partial q^j}\beta^{ij}. \tag{2.43}$$

From (2.43) it follows that at every $m \in M$ the first order tangent space T_mM is a subspace in $\tau_m M$ consisting of vectors with zero matrix (β^{ij}). However, if this matrix is not zero, the column (b^i) is not a first order tangent vector since it has another transformation rule. On the other hand, by (2.43) the field of matrices (β^{ij}) is a symmetric $(2,0)$-tensor field and it is symmetric in every coordinate system.

There is an analogous construction of second order differential forms.

The theory of second order vectors and differential forms is presented in detail, for example, in [69, 179, 180, 204, 205]. In these works one can also find an interesting approach to stochastic differential equations on manifolds.

At every $m \in M$ there is a canonical isomorphisms between the space $T_mM \odot T_mM$ (where $\odot$ denotes the symmetric tensor product, see Section 1.5) and the quotient space $\tau_m M/T_mM$, and hence between $TM \odot TM$ and $\tau M/TM$ (see [205]). Taking into account this factorization, we construct the morphism $\mathcal{Q} : \tau M \to TM \odot TM$, i.e., the field of linear projectors $\mathcal{Q}_m : \tau_m M \to T_mM \odot T_mM$ such that

$$\mathcal{Q}B(t,m) = \mathcal{Q}\left(b^i \frac{\partial}{\partial q^i} + \beta^{ij}\frac{\partial^2}{\partial q^i \partial q^j}\right) = \beta^{ij}\frac{\partial}{\partial q^i} \otimes \frac{\partial}{\partial q^j}. \tag{2.44}$$

Every connection H on M determines a linear operator from $\tau_m M$ to T_mM at any point $m \in M$ as follows:

$$\mathcal{H}\left(b^k \frac{\partial}{\partial q^k} + \beta^{ij}\frac{\partial^2}{\partial q^j \partial q^j}\right) = \left(b^k + \Gamma^k_{ij}\beta^{ij}\right)\frac{\partial}{\partial q^k}, \tag{2.45}$$

where Γ^k_{ij} are the Christoffel symbols of the connection H. Thus connections, and only connections, are smooth cross-sections of the bundle $\mathrm{Hom}(\tau M, TM)$ of fiber-wise linear operators from τM to TM.

Let $m(t) = (q^1(t), \ldots, q^n(t))$ be a smooth curve in a chart $\mathcal{U}$. The second order vector $\mathrm{D}^2 m(t) = \ddot{q}^k \frac{\partial}{\partial q^k} + \dot{q}^i \dot{q}^j \frac{\partial^2}{\partial q^i \partial q^j}$ is called the *acceleration* of $m(t)$.

Proposition 2.75 *For any smooth curve the equality* $\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) = \mathcal{H}\mathrm{D}^2 m(t)$ *holds where* $\frac{\mathrm{D}}{\mathrm{d}t}$ *is the covariant derivative of a connection* H.

Indeed, by formula (2.45) we obtain that $\mathcal{H}\mathrm{D}^2 m(t) = (\ddot{q}^k + \Gamma^k_{ij}\dot{q}^i\dot{q}^j)\frac{\partial}{\partial q^k}$.

Corollary 2.76 *A curve* $m(t)$ *is a geodesic of a connection* H *if and only if* $\mathcal{H}D^2 m(t) = 0$.

Proof. By Proposition 2.75 the equality $\mathcal{H}\mathrm{D}^2 m(t) = 0$ means that for $m(t)$ the geodesic equation (2.25) holds. $\square$

Chapter 3
Ordinary Differential Equations

3.1 Global in Time Existence of Solutions of Ordinary Differential Equations

Many criteria for the extendability to $(-\infty, \infty)$ of the solutions of differential equations in vector spaces are known (see, e.g., the bibliography in [144]). The main aim of this section is to modify some conditions of this sort in such a way that they become necessary and sufficient. The trick here is the transition to extended phase spaces and an analysis of the so-called proper functions or complete Riemannian metrics on manifolds.

3.1.1 A necessary and sufficient condition for completeness of a vector field of one-sided type

Here we use the notation and notions from Section 1.1 concerning vector fields.

Let M be a smooth manifold with dimension $n < \infty$ and a smooth vector field X be given on M.

Definition 3.1. A vector field X and its flow are called *complete* if all its integral curves are well-defined for $t \in (-\infty, +\infty)$.

Denote by $m(s) : M \to M$, $s \in R$, the flow of X. For any point $m \in M$ and time t the *orbit* $m(s)(t, m) = m_{t,m}(s)$ of the flow is the solution of the equation

$$\dot{m}_{t,m}(s) = X(s, m_{t,m}(s)) \tag{3.1}$$

with the initial condition

$$m_{t,m}(t) = m. \tag{3.2}$$

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_3,

If the right-hand side of an equation is a complete vector field, we say that the flow of this equation is complete, i.e., the completeness of vector fields and their flows are equivalent.

Definition 3.2. A mapping $F : \mathbf{X} \to \mathbf{Y}$ of topological spaces is called *proper* if the pre-image of every relatively compact set in $\mathbf{Y}$ is relatively compact in $\mathbf{X}$. In particular, a function $f : \mathbf{X} \to \mathbb{R}$ on a topological space $\mathbf{X}$ is called proper if the pre-image of any relatively compact set in $\mathbb{R}$ is a relatively compact set in $\mathbf{X}$.

Recall that in any finite-dimensional space (in particular, in $\mathbb{R}$) a set is relatively compact if and only if it is bounded.

Examples of proper functions are the norm in a Euclidean space and the distance function on a complete Riemannian manifold.

In what follows we shall mainly deal with proper functions on smooth manifolds.

In questions connected with the completeness of flows in $\mathbb{R}^n$ the functions $f : \mathbb{R}^n \to \mathbb{R}$ such that $f(x) \to \infty$ as $x \to \infty$ are often considered. Clearly such functions are proper. The classical completeness theorem with so-called one-sided estimates says that if there exists a function f as above such that $\mathrm{grad} f \cdot X < C$ (where $\cdot$ denotes the inner product in $\mathbb{R}^n$) for a certain constant C, the flow of the vector field X on $\mathbb{R}^n$ exists up to $+\infty$. Note that $\mathrm{grad} f \cdot X = Xf$, the derivative of f in direction of X. In this subsection we show how to modify this completeness theorem in order to get a necessary and sufficient condition for completeness.

Consider the extended phase space $M^+ = R \times M$ with the natural projection $\pi^+ : M^+ \to M$, $\pi^+(t, m) = m$. Introduce the vector field X^+ on M^+ given at the point $(t, m) \in M^+$ as $X^+_{(t,m)} = (1, X(t, m))$. It is clear that its coordinate representation is given in the form $X^+ = \frac{\partial}{\partial t} + X^1 \frac{\partial}{\partial q^1} + ... + X^n \frac{\partial}{\partial q^n}$. Hence the corresponding differential operator on the space of C^1-smooth functions on M^+ takes the form $\frac{\partial}{\partial t} + X$.

Theorem 3.3 *A smooth vector field X on a finite-dimensional manifold M is complete if and only if there exists a smooth proper function $\varphi : M^+ \to \mathbb{R}$ such that the absolute value of the derivative $|X^+\varphi|$ of φ in the direction of X^+ is uniformly bounded, i.e., $|X^+\varphi| = |(\frac{\partial}{\partial t} + X)\varphi| \le C$ at any $(t, m) \in M^+$ for some constant $C > 0$ that does not depend on (t, m).*

Proof. Sufficiency.

Consider the flow $m^+(s) : M^+ \to M^+$, $s \in \mathbb{R}$, with orbits $m^+(s)(t, m) = m^+_{(t,m)}(s)$ being the solutions of the equation

$$\dot{m}^+_{(t,m)}(s) = X^+\left(m^+_{(t,m)}(s)\right)$$

with initial conditions

$$m^+_{(t,m)}(t) = (t, m).$$

Consider the derivative $X^+\varphi$ of φ along X^+. At $(t,m) \in M^+$, by definition of the derivative in the direction of a vector field, we get the equality

$$X^+\varphi(t,m) = \frac{\mathrm{d}}{\mathrm{d}s}\,\varphi\left(m^+_{(t,m)}(s)\right)\Big|_{s=t},$$

and under the hypothesis of our Theorem

$$\left|\frac{\mathrm{d}}{\mathrm{d}s}\varphi\left(m^+_{(t,m)}(s)\right)\Big|_{s=t}\right| \le C. \tag{3.3}$$

Represent the values of φ along the orbit $m^+_{(t,m)}(s)$ as follows:

$$\varphi\left(m^+_{(t,x)}(s)\right) - \varphi((t,m)) = \int_0^s \frac{\mathrm{d}}{\mathrm{d}\tau}\varphi\left(m^+_{(t,m)}(\tau)\right)\mathrm{d}\tau.$$

From the last equality and from inequality (3.3) we obtain that if s belongs to a finite interval, the values $\varphi(m^+_{(t,x)}(s))$ are bounded in $\mathbb{R}$. Then since φ is proper, this means that the set $m^+_{(t,m)}(s)$ is relatively compact in M^+.

Recall that by the classical solution existence theorem the domain of any solution of an ODE is an open interval in $\mathbb{R}$. In particular for $s > t$ the solution of the above Cauchy problem is well-defined for $s \in [t, t+\varepsilon)$. If $\varepsilon > 0$ is finite, then from the above arguments it follows that the corresponding values of the solution belong to a relatively compact set in M and so the solution is well-defined for $s \in [t, \varepsilon]$. The same arguments are valid also for $s < t$. Thus the domain is both open and closed and so it coincides with the entire real line $(-\infty, \infty)$.

Taking into account the construction of the vector field X^+, we can represent the integral curves $m^+_{(t,m)}(s)$ in the form $m^+_{(t,m)}(s) = (s, m_{t,m}(s))$. Hence from the global existence of integral curves of X^+ we easily obtain the global existence of integral curves of X. So, the vector field X is complete.

Necessity.

Let the vector field X be complete. Thus all orbits $m_{t,m}(s)$ of the flow $m(s)$ are well-defined on the entire real line. Let $V = \{V_i\}_{i\in N}$ be a countable locally-finite cover of M where all V_i are open and relatively compact. Such a cover exists since every finite-dimensional manifold is locally compact and by definition satisfies the second countability axiom. Introduce the functions $\psi_i : M \to R$ by the formula

$$\psi_i(m) = \begin{cases} i \ \text{ if } \ m \in V_i \\ 0 \ \text{ if } \ m \notin V_i. \end{cases}$$

Denote by $\{\varphi_i\}_{i=1}^{\infty}$ the smooth partition of unity subordinate to the above cover and define the function ψ on M by:

$$\psi(m) = \sum_{i=1}^{\infty} \psi_i(m)\varphi_i(m). \tag{3.4}$$

It is clear that $\psi(m)$ is smooth and proper by construction. Our construction of the function $\psi(m)$ is taken from [134].

Introduce the function $\Phi : M^+ \to R$ as follows. For any point $(t, m) \in M^+$ set $\Phi(t, m) = \psi(m_{t,m}(0))$. By construction the function Φ is constant along any orbit of the flow $m^+(s)$. Indeed, for $m^+(s)(t, m) = (s, m_{t,m}(s))$ the equality $m_{s,m_{t,m}(s)}(0) = m_{t,m}(0)$ holds.

Consider the function $\varphi : M^+ \to R$, $\varphi(t, m) = \Phi(t, m) + t$. Clearly φ is smooth and proper. Consider $X^+\varphi$. By the construction of the vector field X^+ and of the function φ we get

$$X^+\varphi = X^+(\Phi(t, m)) + X^+t = 0 + 1 = 1.$$

Thus we have proven the necessary part of our Theorem for $C = 1$. This completes the proof. □

3.1.2 A generalization to the infinite-dimensional case

A direct infinite-dimensional generalization of the results of the previous subsection cannot be obtained at least because of the absence of proper functions on infinite dimensional manifolds. To avoid this difficulty, here we replace functions which are proper with respect to the strong topology by functions which are proper with respect to a weaker topology.

Another difference is that infinite-dimensional ordinary differential equations with smooth right-hand sides very rarely arise in applications. Generally the right-hand sides are locally Lipschitz continuous and/or completely continuous (recall that equations with simply continuous right-hand sides in infinite-dimensional cases may have no solutions at all) or even are given only on an everywhere dense subset. Examples of the latter are, say, parabolic equations, considered as ordinary differential equations on some function spaces, or equations, close to hydrodynamical ones, on the Sobolev vector fields on a compact manifold (see, e.g., [61]). In both cases the differential operator on the right-hand side makes functions (vector fields) less smooth so that the vectors of the right-hand side that are tangent to the manifold (phase space) are well-defined only on an everywhere dense subset of "more smooth" functions (vector fields).

In this subsection we obtain infinite-dimensional necessary and sufficient conditions for the global existence of solutions both for the right-hand sides given on the entire manifold and for those given on everywhere dense subsets, under some additional conditions that seem to be natural.

3.1.2.1 Basic results for linear Banach spaces

Let B be a Banach space and $X(t,m)$ be a continuous non-autonomous vector field on B, $t \in \mathbb{R}$. For autonomous vector fields no simplification of the construction and results occurs and so everywhere below we consider the general non-autonomous case. For such a vector field we consider the following Cauchy problem

$$\frac{\mathrm{d}}{\mathrm{d}s} m_{(t,m)}(s) = X(s, m_{(t,m)}(s)) \tag{3.5}$$

$$m_{(t,m)}(t) = m. \tag{3.6}$$

Since B is infinite dimensional, without additional conditions (3.5)–(3.6) may not have solutions. Nevertheless solutions do exist if, say, $X(t,m)$ is completely continuous, i.e., it is continuous jointly in $t \in \mathbb{R}$ and $m \in B$ and for each bounded set $A \subset B$ and each interval $[a,b] \subset \mathbb{R}$ the mapping X sends the set $[a,b] \times A$ into a compact set.

Definition 3.4. We say that the Cauchy problem for equation (3.5) is *locally well-posed* if for each $m \in B$ and $t \in \mathbb{R}$ there exists a unique local solution of problem (3.5)–(3.6), well-defined on an interval $(t - \varepsilon'_{(t,m)}, t + \varepsilon_{(t,m)})$, that continuously depends on initial data, i.e., from $m_n \to m$, $t_n \to t$ it follows that $m_{(t_n,m_n)}(s) \to m_{(t,m)}(s)$ for $s \in \bigcap_n (t - \varepsilon'_{(t_n,m_n)}, t + \varepsilon_{(t_n,m_n)})$, where $(t - \varepsilon'_{(t_n,m_n)}, t + \varepsilon_{(t_n,m_n)})$ is the domain of $m_{(t_n,m_n)}(\cdot)$.

Remark 3.5. On $(t - \varepsilon'_{(t,m)}, t]$ the curve $m_{(t,m)}(s)$ may be presented as a solution of the equation $\frac{\mathrm{d}}{\mathrm{d}s} m_{(t,m)}(s) = -X(s, m_{(t,m)}(s))$, inverse to (3.5). This is why instead of saying that the Cauchy problem (3.5)–(3.6) is locally well-posed one sometimes says both the direct and inverse Cauchy problems for (3.5) are locally well-posed.

We retain Definition 3.1 of complete vector fields for infinite-dimensional systems.

Definition 3.6. A function $f : E \to \mathbb{R}$, where E is a Banach space, is called *weakly proper* if for any relatively compact set in $\mathbb{R}$ its pre-image in E is bounded.

Remark 3.7. Recall Definition 3.2 of a proper mapping of topological spaces. Taking into account the features of weakly compact sets in Banach spaces (see, e.g., [156, 223, 235]), one can easily see that if B is a reflexive Banach space, Definition 3.6 means that a weakly proper function is indeed proper with respect to the weak topology on B.

Consider the extended phase space $B^+ = \mathbb{R} \times B$ and the vector field $X^+_{(t,m)} = (1, X(t,m))$ on it. From the hypothesis one can easily derive that $X^+_{(t,m)}$ is completely continuous. It is clear that the curves

$$m^+_{(t,m)}(s) = (s, m_{(t,m)}(s)) \tag{3.7}$$

satisfy the following Cauchy problem on B^+

$$\frac{\mathrm{d}}{\mathrm{d}s} m^+_{(t,m)}(s) = X^+ \left(m^+_{(t,m)}(s) \right) \tag{3.8}$$

$$m^+_{(t,m)}(t) = (t, m). \tag{3.9}$$

Theorem 3.8 *Let $X(t,m)$ be a completely continuous vector field on a Banach space B. Let there exist a continuous weakly proper function $f : B^+ \to \mathbb{R}$ such that for any curve $m^+_{(t,m)}(s)$, as defined by (3.7), there exists a real constant $C > 0$ for which the relation*

$$\left| f(m^+_{(t,m)}(s_1)) - f(m^+_{(t,m)}(s_2)) \right| < C|s_1 - s_2| \tag{3.10}$$

holds for every pair s_1 and s_2 in the domain of the curve $m_{(t,m)}(s)$. Then $X(t,m)$ is complete.

Proof. Let f be as in the hypothesis and let $m_{(t,m)}(s)$ be an arbitrary solution of (3.5)–(3.6) that is well-defined on some interval $s \in (t - \varepsilon', t + \varepsilon)$. The corresponding curve $m^+_{(t,m)}(s)$ is well-defined on the same interval.

If both ε and ε' are infinite, the Theorem is proved. Let ε be finite. From the hypothesis it follows that for $s \in [t, t+\varepsilon)$ the inequality $|f(m^+_{(t,m)}(t)) - f(m^+_{(t,m)}(s))| < C\varepsilon$ holds. Thus on this interval the values $|f(m^+_{t,m}(s))|$ are bounded by $|f(m^+_{t,m}(t))| + C\varepsilon$. Then from Definition 3.6 it follows that the points of this curve for $s \in [t, t+\varepsilon)$ belong to a bounded set $\Theta = f^{-1}([t, t+\varepsilon))$. Since Θ is bounded, the image of Θ under the completely continuous mapping X^+ is compact. Thus the norm of X^+ as a continuous function is bounded on Θ and so $\|X^+(m^+_{(t,m)}(s)\|$ is bounded on $[t, t+\varepsilon)$. Thus the length of $m^+_{t,m}(s)$ for $s \in [t, t+\varepsilon)$, which is equal to $\int_t^\varepsilon \|\frac{\mathrm{d}}{\mathrm{d}s} m^+_{(t,m)}(s)\| \mathrm{d}s = \int_t^\varepsilon \|X^+(m^+_{(t,m)}(s))\| \mathrm{d}s$, is bounded. Since B is complete and since $m^+_{(t,m)}(s)$ is continuous in s, one can easily derive from this fact that the limit $\lim_{s \to t+\varepsilon} m^+_{(t,m)}(s)$ exists, i.e., $m^+_{(t,m)}(s)$ exists on the closed interval $[t, t+\varepsilon]$.

In complete analogy with these arguments we show that if ε' is finite, $m^+_{(t,m)}(s)$ exists on the closed interval $[t - \varepsilon', t]$. This means that the domain of $m^+_{(t,m)}(s)$ is both open and closed, hence it is equal to $\mathbb{R}$. Obviously the same is true for the domain of $m_{(t,m)}(s)$. Thus, since the solution $m_{(t,m)}(s)$ was arbitrary, X is complete. □

Corollary 3.9 *Let $X(t,m)$ be a completely continuous vector field on a Banach space B. Suppose there exists a continuous weakly proper function $f : B \to \mathbb{R}$ such that for any solution $m_{(t,m)}(s)$ of (3.5)–(3.6) there exists a* *real constant $C > 0$ for which the relation (3.10) holds for every pair s_1 and s_2 in the domain of $m_{(t,m)}(s)$. Then $X(t, m)$ is complete.*

Corollary 3.9 is an analog of a well-known sufficient condition for the completeness of vector fields in finite-dimensional spaces (see the previous subsection). The proof of Corollary 3.9 is the same as that of Theorem 3.8 with a certain simplification since here we do not use the extended phase space. Notice that the conditions of Theorem 3.8 after some modification become necessary and sufficient for the completeness of vector fields (see Theorem 3.13 below) while Corollary 3.9 gives only sufficient conditions for completeness.

The next statement describes a particular case where (3.10) is fulfilled. The relation obtained here easier to verify in practice.

Theorem 3.10 *Let a continuous weakly proper function $f : B^+ \to \mathbb{R}$ be such that its derivative $X^+ f$ in the direction of a vector field X^+ (in the ordinary sense) is well-defined and satisfies the estimate $|X^+ f| < C$ at any point $(t, m) \in B^+$ for some constant $C > 0$ that is independent of (t, m). Then along any curve $m^+_{(t,m)}(s)$, as introduced by (3.7), relation (3.10) is satisfied with this C and so by Theorem 3.8 $X(t, m)$ is complete.*

Indeed,

$$\left| f\left(m^+_{(t,m)}(s_1)\right) - f\left(m^+_{(t,m)}(s_2)\right)\right| = \left| \int_{s_1}^{s_2} (X^+ f)\left(m^+_{(t,m)}(s)\right) \mathrm{d}s \right|$$
$$\leq \left| \int_{s_1}^{s_2} \left|(X^+ f)\left(m^+_{(t,m)}(s)\right)\right| \mathrm{d}s \right| < \left| \int_{s_1}^{s_2} C \mathrm{d}s \right| = C|s_1 - s_2|.$$

The next statement gives a necessary condition for the completeness of $X(t, m)$ of the same sort as the sufficient condition of Theorem 3.8.

Theorem 3.11 *Let $X(t, m)$ be a continuous vector field on a Banach space B such that the Cauchy problem (3.5)–(3.6) is locally well-posed (see Definition 3.4). If $X(t, m)$ is complete, there exists a continuous weakly proper function $f : B^+ \to \mathbb{R}$ such that for any curve $m^+_{(t,m)}(s)$, as introduced by (3.7), there exists a real constant $C > 0$ for which relation (3.10) holds for every pair s_1 and s_2.*

Proof. Let the solutions $m_{(t,m)}(s)$ of the problem (3.5)–(3.6) exist on the entire line. Introduce the weakly proper continuous function $r : B \to \mathbb{R}$, $r(m) = \|m\|$.

Denote by $g_s : B \to B$ the flow of the vector field X, i.e., $g_s(m) = m_{(0,m)}(s)$. Since $X(t, m)$ is complete, under the hypothesis of our Theorem g_s is well-defined and forms a continuous family of homeomorphisms of B. For the same reason, for every point $(t, m) \in B^+$ the value $m_{(t,m)}(0)$ of the curve $m_{(t,m)}(s)$ is well-defined and continuously depends on (t, m). Now construct the continuous function $\Phi : B^+ \to \mathbb{R}$ by assigning the value $\Phi(t, m) = r^2(m_{t,m}(0))$ to the point $(t, m) \in B^+$.

Consider the continuous function $f : B^+ \to \mathbb{R}$, $f(t, m) = \Phi(t, m) + t$. One can easily see that f is weakly proper.

By construction, Φ takes constant values along the curves $m^+_{(t,m)}(s)$ on B^+. Indeed, for $m^+_{(t,m)}(s) = (s, m_{(t,m)}(s))$ the equality $m_{(s,m_{(t,m)}(s))}(0) = m_{(t,m)}(0)$ holds for all s. Hence,

$$\begin{aligned}&\left|f\left(m^+_{(t,m)}(s_1)\right) - f\left(m^+_{(t,m)}(s_2)\right)\right|\\ &= \left|r^2\left(m_{(t,m)}(0)\right) + s_1 - r^2\left(m_{(t,m)}(0)\right) - s_2\right| = |s_1 - s_2|.\end{aligned}$$

Thus, the constant C in the conclusion of the theorem may take any value greater than or equal to 1. □

For a smooth vector field $X(t, m)$, under some additional hypotheses we obtain a necessary condition for completeness of the same sort as the sufficient condition of Theorem 3.10.

Theorem 3.12 *Let a Banach space B have a smooth norm and let $X(t, m)$ be a smooth and complete vector field on B. Then there exist a smooth weakly proper function $f : B^+ \to \mathbb{R}$ and a real constant $C > 0$ such that for the derivative $X^+ f$ of f in the direction of X^+ the inequality $|X^+ f| < C$ holds on B^+.*

Proof. Since $X(t, m)$ is smooth, it is locally Lipschitz continuous and so the Cauchy problem (3.5)–(3.6) is locally well-posed. From the fact that both the vector field $X(t, m)$ and the function $r^2(m)$ are smooth it follows that the function $f : B^+ \to \mathbb{R}$, constructed in the proof of Theorem 3.11, is smooth. Since Φ takes constant values along the curves $m^+_{(t,m)}(s)$ on B^+, a direct calculation shows that $|X^+ f| = 1$. □

Now let a vector field $X(t, m)$ on a Banach space B be completely continuous and be such that the Cauchy problem (3.5)–(3.6) is locally well-posed. Then as a corollary to Theorems 3.8 and 3.11 we obtain the following:

Theorem 3.13 *Let B be a Banach space. A completely continuous vector field $X(t, m)$ on B for which the Cauchy problem (3.5)–(3.6) is locally well-posed is complete if and only if there exists a continuous weakly proper function $f : B^+ \to \mathbb{R}$ such that for any curve $m^+_{(t,m)}(s)$, as introduced by (3.7), there exists a real constant $C > 0$ for which relation (3.10) holds for every pair s_1 and s_2 in the domain of the curve $m_{(t,m)}(s)$.*

Corollary 3.14 *Let B be a Banach space. Both completely continuous and locally Lipschitz continuous vector fields $X(t, m)$ on B are complete if and only if there exists a continuous weakly proper function $f : B^+ \to \mathbb{R}$ such that for any curve $m^+_{(t,m)}(s)$, as introduced by (3.7), there exists a real constant $C > 0$ for which relation (3.10) holds for every pair s_1 and s_2 in the domain of the curve $m_{(t,m)}(s)$.*

Indeed, for a locally Lipschitz continuous vector field the Cauchy problem (3.5)–(3.6) is locally well-posed.

As a corollary to Theorems 3.10 and 3.12 we obtain the following:

Theorem 3.15 *Let a Banach space B have smooth norm and let $X(t,m)$ be a smooth and completely continuous vector field on B. $X(t,m)$ is complete if and only if there exist a smooth weakly proper function $f : B^+ \to \mathbb{R}$ and a real constant $C > 0$ such that for the derivative X^+f of f in the direction of X^+ the inequality $|X^+f| < C$ holds on B^+.*

3.1.2.2 The case when the right-hand side of the equation is defined on an everywhere dense subset

Let D be a Banach space that is embedded into B by a continuous map so that the image of D is everywhere dense in B. For the sake of simplicity we do not distinguish between D and its image, so we regard D as a subset of B.

Let for every $m \in D$ and $t \in \mathbb{R}$ a vector $X(t,m) \in B$ be given. Consider the Cauchy problem

$$\frac{\mathrm{d}}{\mathrm{d}s} m_{(t,m)}(s) = X\left(s, m_{(t,m)}(s)\right) \tag{3.11}$$

$$m_{(t,m)}(t) = m \in D. \tag{3.12}$$

In analogy with the notions from the theory of partial differential equations we give the following:

Definition 3.16. We say that the Cauchy problem (3.11)–(3.12) is *locally well-posed* if for each $m \in D$ and $t \in \mathbb{R}$ there exists a unique local solution, given on a certain interval $(t - \varepsilon', t + \varepsilon)$, belonging to D and continuously depending on the initial data, i.e., from $m_n \to m$ it follows that $m_{(t,m_n)}(s) \to m_{(t,m)}(s)$ for $s \in \bigcap_n (t - \varepsilon'_n, t + \varepsilon_n)$, where $\bigcap_n (t - \varepsilon'_n, t + \varepsilon_n)$ is the domain of $m_{(t,m_n)}(\cdot)$. Here the convergence is understood to be in the topology of D.

We refer the reader, say, to [163, 164] for examples of conditions that guarantee for an equation the well-posedness of the Cauchy problem (3.11)–(3.12).

Definition 3.17. The Cauchy problem (3.11)–(3.12) is called *regular* if its local solutions exist and, for a solution $m_{(t,m)}(s)$ with $m \in D$, from the fact that at the time s^* the point $m_{(t,m)}(s^*)$ belongs to B it follows that $m_{(t,m)}(s^*) \in D$.

Examples of regular Cauchy problems can be found, e.g., in [15, 34, 61].

In analogy with the above notation, we introduce $D^+ = \mathbb{R} \times D$ and $m^+_{(t,m)}(s) = (s, m_{(t,m)}(s))$.

Theorem 3.18 *Let the embedding $i : D \to M$ be completely continuous, let the Cauchy problem* (3.11)–(3.12) *be regular and let $f : D^+ \to \mathbb{R}$ be a continuous weakly proper function such that for any curve $m^+_{(t,m)}(s)$ there exists a real constant $C > 0$ for which the relation*

$$\left| f\left(m^+_{(t,m)}(s_1)\right) - f\left(m^+_{(t,m)}(s_2)\right)\right| < C|s_1 - s_2|$$

holds for every pair s_1 and s_2 in the domain of the curve $m_{(t,m)}(s)$. Then all solutions of (3.11)–(3.12) *are well-defined for $s \in (-\infty, \infty)$.*

Proof. Let $m_{(t,m)}(s)$ be an arbitrary solution of the Cauchy problem (3.11)–(3.12) that exists for $s \in (t - \varepsilon', t + \varepsilon)$. If both ε and ε' are infinite, the Theorem is proved. Suppose they are finite.

As in the proof of Theorem 3.8, from the inequality

$$\left| f\left(m^+_{(t,m)}(s_1)\right) - f\left(m^+_{(t,m)}(s_2)\right)\right| < C|s_1 - s_2|$$

along the curve $m^+_{(t,m)}(s)$ it follows that the curve $m_{(t,m)}(s)$ on $(t - \varepsilon', t + \varepsilon)$ belongs to a bounded set in D. Since the embedding of D into B is completely continuous, this set is relatively compact in B. So, the curve $m_{(t,m)}(s)$, continuous and relatively compact on $(t-\varepsilon', t+\varepsilon)$, can be extended to $[t-\varepsilon', t+\varepsilon]$ in B. But since the solutions of (3.11)–(3.12) are regular, the extension belongs to D. Thus the domain of $m_{(t,m)}(s)$ is both open and closed and so it coincides with $\mathbb{R}$. □

Theorem 3.19 *Let the Cauchy problem* (3.11)–(3.12) *be locally well-posed and all solutions of* (3.11)–(3.12) *exist for $s \in (-\infty, \infty)$. Then there exists a continuous weakly proper function $f : D^+ \to \mathbb{R}$ such that for any curve $m^+_{(t,m)}(s)$ there exists a real constant $C > 0$ for which the relation*

$$\left| f\left(m^+_{(t,m)}(s_1)\right) - f\left(m^+_{(t,m)}(s_2)\right)\right| < C|s_1 - s_2|$$

holds for every pair s_1 and s_2.

Proof. This proof is a simple modification of that for Theorem 3.11. The space B^+ is replaced by D^+. Existence, uniqueness and continuity of solutions in D^+ now follow from the local well-posedness of the Cauchy problem according to Definition 3.16. The remaining arguments are the same as in the proof of Theorem 3.11. □

Theorem 3.20 *Let the embedding $i : D \to B$ be completely continuous and let the Cauchy problem* (3.11)–(3.12) *be locally well-posed and regular. Then all solutions of* (3.11)–(3.12) *are well-defined for all $s \in (-\infty, \infty)$ if and only if there exists a continuous weakly proper function $f : D^+ \to \mathbb{R}$ such that for any curve $m^+_{(t,m)}(s)$ there exists a real constant $C > 0$ for which the relation*

$$\left| f\left(m^+_{(t,m)}(s_1)\right) - f\left(m^+_{(t,m)}(s_2)\right)\right| < C|s_1 - s_2|$$

holds for every pair s_1 *and* s_2 *in the domain of the curve* $m_{(t,m)}(s)$*.*

Theorem 3.20 is a simple corollary to Theorems 3.18 and 3.19.

3.1.2.3 The case of manifolds

Let M be a smooth Banach manifold with model Banach space B. For the sake of convenience we consider charts on M as triples (U, V, φ), where V is an open ball in the model space B, U is an open set in M and $\varphi : V \to U$ is a homeomorphism. Recall that the structure of a smooth manifold is given by the maximal atlas for whose charts $(\mathcal{U}_\alpha, V_\alpha, \varphi_\alpha)$ and $(\mathcal{U}_\beta, V_\beta, \varphi_\beta)$, such that $\mathcal{U}_\alpha \cap \mathcal{U}_\beta = \mathcal{U}_{\alpha\beta} \neq \emptyset$, the changes of coordinates $\varphi_\beta^{-1} \circ \varphi_\alpha : \varphi_\alpha^{-1}(\mathcal{U}_{\alpha\beta}) \to \varphi_\beta^{-1}(\mathcal{U}_{\alpha\beta})$ are smooth.

In addition to the maximal atlas, which defines the smooth structure, we shall also consider atlases (called *continuous*) that define the structure of a topological manifold on M with the same model space B. This means that the changes of coordinates between charts of such an atlas are only continuous. The charts of a continuous atlas will be denoted by $(\tilde{\mathcal{U}}_\alpha, \tilde{V}_\alpha, \tilde{\varphi}_\alpha)$.

In this section we also assume an additional property of the manifold:

Condition 3.21 *There exists a continuous atlas such that the images of all bounded sets in* B *under its coordinate changes are bounded.*

Sometimes we shall also assume that M satisfies the following:

Condition 3.22 M *has a continuous (as in Condition 3.21) countable locally finite atlas such that for any its charts* $(\tilde{\mathcal{U}}_\alpha, \tilde{V}_\alpha, \tilde{\varphi}_\alpha)$ *the set* $\tilde{V}_\alpha \subset B$ *is bounded with respect to the norm of the model space* B*.*

Definition 3.23. A set $\Theta \subset M$ is called *relatively weakly compact* if there exists a finite collection of charts $(\tilde{\mathcal{U}}_i, \tilde{V}_i, \tilde{\varphi}_i)$ of a continuous atlas on M such that $\Theta \subset \bigcup_i \tilde{\mathcal{U}}_i$ and for every i the set $\tilde{\varphi}_i^{-1}(\Theta \bigcap \tilde{\mathcal{U}}_i) \subset \tilde{V}_i$ is bounded with respect to the norm of the model space B that contains $\tilde{V}_i$.

Remark 3.24. Unlike the case of linear Banach spaces, the weak topology is (generally speaking) ill-defined on Banach manifolds while the slightly stronger topology of weak convergence is well-defined (for details, see [195]). If the model space B of M is a reflexive Banach space, then under some natural conditions the relatively weakly compact set described in Definition 3.23 is relatively compact with respect to the topology of weak convergence on M (see [195]). If M itself is a reflexive Banach space, then any relatively weakly compact set (described in Definition 3.23) is relatively weakly compact (i.e., with respect to weak convergence) by the well-known properties of reflexive Banach spaces (cf. Remark 3.7).

Definition 3.25. A function $f : N \to \mathbb{R}$ on a Banach manifold N is called *weakly proper* if for any relatively compact set in $\mathbb{R}$ its pre-image is relatively weakly compact in N as in Definition 3.23.

Definition 3.26. A vector field $X(t,m)$ on M, $t \in \mathbb{R}$, is called *completely continuous* if its restriction to any chart (U, V, φ), as a mapping from $\mathbb{R} \times V$ into the model space B, is completely continuous in the ordinary sense.

For $X(t,m)$ we shall consider the Cauchy problem (3.5)–(3.6) on M. We retain the Definition 3.1 of a complete vector field as well as the Definition 3.4 of local well-posedness of solutions of the Cauchy problem.

As above, we consider the extended phase space $M^+ = \mathbb{R} \times M$ and the vector field $X^+_{(t,m)} = (1, X(t,m))$ on it. By formula (3.7) we also introduce the curves $m^+_{(t,m)}(s)$ that satisfy the Cauchy problem (3.8)–(3.9) on M^+.

Theorem 3.27 *Let $X(t,m)$ be a completely continuous vector field on a Banach manifold M. Let there exist a continuous weakly proper function $f : M^+ \to \mathbb{R}$ on M^+ such that for any curve $m^+_{(t,m)}(s)$, as introduced by* (3.7)*, there exists a real constant $C > 0$ for which the relation* (3.10) *holds for every pair s_1 and s_2 in the domain of the curve $m_{(t,m)}(s)$. Then $X(t,m)$ is complete.*

The proof of Theorem 3.27 is a simple modification of that for Theorem 3.8 and is left to the reader.

We cannot prove a necessary condition of completeness in the same way as in Theorem 3.11 since a weakly proper function such as $r^2(m)$ may not exist on an arbitrary infinite-dimensional manifold and so it is a problem to determine whether at least one weakly proper function on a Banach manifold exists. We say that a Banach manifold M *admits a continuous weakly proper function* if at least one such function is well-defined on M. The following statements describe sufficient conditions for a Banach manifold M to admit a weakly proper function of a sort different from $r(m)$ on a Banach space.

Theorem 3.28 *If M satisfies Condition 3.22, it admits a continuous weakly proper function.*

Proof. Let $\{(\tilde{\mathfrak{U}}_i, \tilde{V}_i, \tilde{\varphi}_i)\}$ be an atlas as in the hypothesis. In particular, the covering $\{\tilde{\mathfrak{U}}_i\}$ is locally finite and so there exists a continuous partition of unity $\theta_i(m)$, $i = 1, \ldots, \infty$ corresponding to $\{\tilde{\mathfrak{U}}_i\}$ (see, e.g., [172]).

Define the functions $\psi_i : M \to \mathbb{R}$ by the formulae

$$\psi_i(m) = \begin{cases} i \text{ if } m \in \tilde{\mathfrak{U}}_i \\ 0 \text{ if } m \notin \tilde{\mathfrak{U}}_i \end{cases}$$

and construct the function $\psi(m)$ on M by the formula:

$$\psi(m) = \sum_{i=1}^{\infty} \theta_i(m)\psi_i(m).$$

By construction $\psi(m)$ is continuous and weakly proper. □

The construction of ψ above is a modification of the construction in [134] (cf. the proof of Theorem 3.3).

Corollary 3.29

(i) *If M is a paracompact Lindelöf topological space (see, e.g., [168]), it admits a continuous weakly proper function.*
(ii) *If M is paracompact and separable, it admits a continuous weakly proper function.*

Proof. (i) For each $m \in M$ choose a chart $(\mathcal{U}_\alpha, V_\alpha, \varphi_\alpha)$ such that $m \in \mathcal{U}_\alpha$. Then take an open bounded ball $V_m \subset V_\alpha$ centered at $\varphi_\alpha^{-1}m$ and define $\mathcal{U}_m = \varphi_\alpha V_m$. Since M is paracompact, there exists a locally finite refinement of the covering $\{\mathcal{U}_m\}$ (see [172]). Since M is a Lindelöf space, we can select a countable subcovering $\{\mathcal{U}_i\}$ of that refinement. By construction each $\mathcal{U}_i$ is a subset of some $\mathcal{U}_\alpha$ and so there exists an open ball $V_i = \varphi_\alpha^{-1}\mathcal{U}_i \subset V_\alpha$. Denote by φ_i the restriction of φ_α to V_i. Then the atlas $\{(\mathcal{U}_i, V_i, \varphi_i)\}$ satisfies the conditions of Theorem 3.28.

(ii) Let Ξ be a countable everywhere dense subset of M. For each $m \in \Xi$ construct a chart $(\mathcal{U}_m, V_m, \varphi_\alpha)$ and a locally finite refinement of the covering $\{\mathcal{U}_m\}$ in the same manner as in the proof of (i). Since every point of Ξ is contained only in a finite number of sets from that refinement, the collection of sets in the refinement is countable. The rest of the proof is the same as in (i). □

Theorem 3.30 *Let M admit a continuous weakly proper function and let a vector field $X(t,m)$ on M be such that the Cauchy problem* (3.5)–(3.6) *is locally well-posed. If $X(t,m)$ is complete, there exists a continuous weakly proper function $f : M^+ \to \mathbb{R}$ on M^+ such that for any curve $m^+_{(t,m)}(s)$, as defined by* (3.7)*, there exists a real constant $C > 0$ for which the relation* (3.10) *holds for every pair s_1 and s_2.*

Proof. Let $X(t,m)$ be complete, i.e., the solutions $m_{(t,m)}(s)$ of the problem (3.5)–(3.6) exist on the entire line. Let ψ be a weakly proper function on M that exists by the hypothesis.

Denote by $g_s : M \to M$ the flow of the vector field X, i.e., $g_s(m) = m_{(0,m)}(s)$. Since $X(t,m)$ is complete and the Cauchy problem is locally well-posed for it, g_s is well-defined and forms a continuous family of homeomorphisms of M. From the completeness of $X(t,m)$ it also follows that for every pair $(t,m) \in M^+$ the value $m_{(t,m)}(0)$ of the solution $m_{(t,m)}(s)$ is well-defined and continuously depends on (t,m).

Now construct a continuous atlas on M^+ as follows: for a chart $(\mathcal{U}_\alpha, V_\alpha, \varphi_\alpha)$ from the smooth atlas on M and for an interval $(s_1, s_2) \subset \mathbb{R}$ define

$$\tilde{\mathcal{U}}_{\alpha,(s_1,s_2)} = \bigcup_{s\in(s_1,s_2)} (s, g_s(\mathcal{U}_\alpha)), \qquad \tilde{V}_{\alpha,(s_1,s_2)} = (s_1,s_2) \times V_\alpha$$

and

$$\tilde{\varphi}_{\alpha,(s_1,s_2)}(s,x) = (s, g_s(\varphi_\alpha(x))$$

for $x \in V_\alpha$. Construct also the continuous function $\Phi : M^+ \to \mathbb{R}$ by assigning the value $\Phi(t,m) = \psi(m_{t,m}(0))$ to the point $(t,m) \in M^+$.

Consider the continuous function $f : M^+ \to \mathbb{R}$, $f(t,m) = \Phi(t,m) + t$. Taking into account the construction of the continuous atlas

$$\left\{\tilde{\mathfrak{U}}_{\alpha,(s_1,s_2)}, \tilde{V}_{\alpha,(s_1,s_2)}, \tilde{\varphi}_{\alpha,(s_1,s_2)}\right\},$$

one can easily see that f is weakly proper.

By its construction, Φ takes constant values along the curves $m^+_{(t,m)}(s)$ on M^+. Indeed, for $m^+_{(t,m)}(s) = (s, m_{(t,m)}(s))$ the equality $m_{(s,m_{(t,m)}(s))}(0) = m_{(t,m)}(0)$ holds for all s. Hence,

$$\begin{aligned} &\left|f\left(m^+_{(t,m)}(s_1)\right) - f\left(m^+_{(t,m)}(s_2)\right)\right| \\ &= \left|\psi\left(m_{(t,m)}(0)\right) + s_1 - \psi\left(m_{(t,m)}(0)\right) - s_2\right| = |s_1 - s_2|. \end{aligned}$$

Thus, the constant C in the conclusion of the theorem may take any value greater than or equal to 1. □

Theorem 3.31 *Let M be a Banach manifold that admits a continuous weakly proper function. A completely continuous vector field $X(t,m)$ on M such that the Cauchy problem (3.5)–(3.6) is locally well-posed is complete if and only if there exists a continuous weakly proper function $f : M^+ \to \mathbb{R}$ such that for any curve $m^+_{(t,m)}(s)$, as defined by (3.7), there exists a real constant $C > 0$ for which the relation*

$$\left|f\left(m^+_{(t,m)}(s_1)\right) - f\left(m^+_{(t,m)}(s_2)\right)\right| < C|s_1 - s_2|$$

holds for every pair s_1 and s_2 from the domain of the curve $m_{(t,m)}(s)$.

Theorem 3.31 follows from Theorems 3.27 and 3.30.

As in the case of linear Banach spaces the following statement holds since for locally Lipschitz continuous vector fields the Cauchy problems are locally well-posed.

Corollary 3.32 *Let M be a Banach manifold that admits a continuous weakly proper function. Both completely continuous and locally Lipschitz continuous vector fields $X(t,m)$ are complete if and only if there exists a continuous weakly proper function $f : M^+ \to \mathbb{R}$ on M^+ such that for any curve $m^+_{(t,m)}(s)$, as defined by (3.7), there exists a real constant $C > 0$ for which the relation*

$$\left|f\left(m^+_{(t,m)}(s_1)\right) - f\left(m^+_{(t,m)}(s_2)\right)\right| < C|s_1 - s_2|$$

holds for every pair s_1 and s_2 in the domain of the curve $m_{(t,m)}(s)$.

An analog of Theorem 3.15 takes the following form:

Theorem 3.33 *Let a Banach manifold M have a smooth countable locally finite atlas such that for any its charts $(\mathfrak{U}_\alpha, V_\alpha, \varphi_\alpha)$ the set $V_\alpha \subset B$ is bounded with respect to the norm of the model space B. Let M also admit a smooth partition of unity. A smooth completely continuous vector field $X(t,m)$ on M is complete if and only if there exists a smooth weakly proper function $f : M^+ \to \mathbb{R}$ and a real constant $C > 0$ such that for the derivative $X^+ f$ of f in the direction of X^+ the inequality $|X^+ f| < C$ holds on M^+.*

A new point in the proof of Theorem 3.33 is that having taken a smooth partition of unity in the construction of the function ψ in the proof of Theorem 3.28, one obtains a smooth ψ. Since the flow of a smooth X^+ is smooth, the function f constructed from ψ in the proof of Theorem 3.30 is also smooth. A direct calculation shows that $|X^+ f| = 1$. The rest of proof is analogous to the previous ones.

Remark 3.34. All finite dimensional manifolds are paracompact and separable and they all admit smooth partitions of unity. This means that they all admit smooth proper functions (all finite dimensional weakly proper functions are proper with respect to the strong topology). Notice also that all continuous finite-dimensional vector fields are completely continuous. Thus, it follows from Theorem 3.33 that any smooth vector field $X(t,m)$ on any finite-dimensional manifold M is complete if and only if there exists a smooth proper function $f : M^+ \to R$ such that $|X^+ f| < C$ on M^+ for some constant $C > 0$. This is the assertion of Theorem 3.3.

Let D be a Banach manifold that is embedded into M by a continuous map so that the image of D is everywhere dense in M. Let for every $m \in D$ and $t \in \mathbb{R}$ a vector $X(t,m) \in T_m M$ be given. Consider the Cauchy problem

$$\frac{\mathrm{d}}{\mathrm{d}s} m_{(t,m)}(s) = X\left(s, m_{(t,m)}(s)\right) \tag{3.13}$$

$$m_{(t,m)}(t) = m \in D. \tag{3.14}$$

We use the same definitions of local well-posedness and regularity as in the case of linear spaces (see Definitions 3.16 and 3.17). As above, we introduce the manifold $D^+ = \mathbb{R} \times D$, the vector field $X^+(t,m) = (1, X(t,m))$ and the curves $m^+_{(t,m)}(s) = (s, m_{(t,m)}(s))$.

By a combination of arguments used above in the proofs for cases of manifolds and for right-hand sides defined on an everywhere dense subset in a Banach space, we obtain the following:

Theorem 3.35 *Let D be a Banach manifold that admits a weakly proper function and let the embedding $i : D \to M$ be completely continuous. Let the Cauchy problem (3.13)–(3.14) be locally well-posed and regular. Then all* *solutions of* (3.13)–(3.14) *are well-defined for all* $s \in (-\infty, \infty)$ *if and only if there exists a continuous weakly proper function* $f : D^+ \to \mathbb{R}$ *such that for any curve* $m^+_{(t,m)}(s)$ *there exists a real constant* $C > 0$ *for which the relation*

$$\left| f\left(m^+_{(t,m)}(s_1)\right) - f\left(m^+_{(t,m)}(s_2)\right) \right| < C|s_1 - s_2|$$

holds for every pair s_1 *and* s_2 *in the domain of the curve* $m_{(t,m)}(s)$.

3.1.3 A necessary and sufficient condition for completeness of a vector field of two-sided type

An alternative type of sufficiency condition for the completeness of flows may be formulated in terms of the so-called two-sided estimates, i.e., estimates of the norm of the right-hand side under some additional condition. In $\mathbb{R}^n$ we have, for example, the condition of sub-linear growth $\|X(t,x)\| < C(1 + \|x\|)$ and the famous Wintner theorem [144] (formulated below). Under the hypotheses of some of these theorems, one can define a new Riemannian metric on the phase space in such a way that the right-hand side of the equation is uniformly bounded by a constant with respect to this metric. Thus, in these cases, the extendability of solutions (the completeness of a vector field) follows from the fact that a solution has bounded length on every finite interval with respect to a complete Riemannian metric and, therefore, is relatively compact.

It turns out that the requirement that the vector field should be bounded with respect to a complete Riemannian metric can be modified in such a way that it becomes necessary and sufficient.

Let M be a finite-dimensional smooth manifold and $X(t,m)$ be a vector field which is jointly smooth in t and m. Denote the extended phase space $\mathbb{R}\times M$ by M^+. Clearly, $T_{(t,m)}M^+ = \mathbb{R}\times T_m M$. As in the previous subsections, define a vector field X^+ on M^+ setting $X^+_{(m,t)} = \big(1, X(m,t)\big)$.

Theorem 3.36 *A field* X *on* M *is* complete *if and only if there exists a complete Riemannian metric on* M^+ *with respect to which* X^+ *is uniformly bounded.*

Proof. Clearly, the completeness of X is equivalent to the completeness of the vector field X^+.

Assume that there exists a complete Riemannian metric on M^+ with respect to which the field X^+ is bounded. Then every integral curve of X^+ has finite length on every finite interval. Since the metric is complete, the last assertion implies the relative compactness of the integral curve on every finite interval. As above, we deduce that the domain of every integral curve is both open and closed in $\mathbb{R}$. This yields the completeness of the field.

Let us prove the "only if" assertion. Let X be complete, then so is X^+. Since, by hypothesis, the field X is smooth, the field X^+ is also smooth. Consider an arbitrary smooth proper real-valued function ψ on the manifold M (see Definition 3.2). The function g, satisfying the aforesaid conditions, can be constructed in the same way as ψ in (3.4).

Pick an inner product depending smoothly on (m,t) on each tangent space $T_{(m,t)}(\{t\}\times M)$ to the submanifold $\{t\}\times M$ of the manifold M^+. For example, one can take a Riemannian metric on M and extend it in a natural way. Now we can construct a Riemannian metric $\langle\cdot,\cdot\rangle_1$ on M^+ by regarding the vectors of the field X^+ as being of unit length and orthogonal to the subspaces $T_{(m,t)}(M\times\{t\})$.

Denote by Φ_t the diffeomorphism of the manifold $M\times\{0\}$ to the manifold $M\times\{t\}$ along the trajectories of the field X^+. The function ψ can be regarded as given on $M\times\{0\}$. Since the integral curves of the field X^+ are globally extendable, the function $f\colon M^+\to\mathbb{R}$ given by the formula

$$f(m,t)=\psi\left(\Phi_t^{-1}(m,t)\right)+t$$

is, obviously, smooth and proper. Clearly, $X^+f=1$, where X^+f is the derivative of the function f in the direction of the field X^+.

Let us now choose an arbitrary smooth function $\varphi\colon M^+\to\mathbb{R}$ such that

$$\varphi(m,t)>\max\ \exp(Yf)^2,$$

where $Y\in T_{(m,t)}\bigl(M\times\{t\}\bigr)$ and $\|Y\|_1=1$. Such a function can be defined as follows. For a relatively compact neighborhood of each point $(m',t')\in M^+$, there exists a constant greater than $\sup\max\exp(Yf)^2$, where, as above, $Y\in T_{(m,t)}\bigl(M\times\{t\}\bigr)$ and $\|Y\|_1=1$, and the supremum is taken over all points (m,t) from the neighborhood. Then, using the paracompactness of M^+ and, as a consequence, the existence of a smooth partition of unity, we glue together the function φ so that it is defined on the whole of M^+.

At every point $(m,t)\in M^+$, define the inner product on $T_{(m,t)}M^+$ by the formula

$$\langle Y,Z\rangle_2=\varphi^2(m,t)\langle p_mY,p_mZ\rangle_1+p_XY\cdot p_XZ,$$

where $Y,Z\in T_{(m,t)}M^+$ and p_m, p_X are (in the metric $\langle\,,\rangle_1$) orthogonal projections of $T_{(m,t)}M^+$ onto $T_{(m,t)}\bigl(M\times\{t\}\bigr)$ and X^+, respectively. Clearly, $\|X^+\|_2=1$.

Lemma 3.37 *The Riemannian metric $\langle\cdot,\cdot\rangle_2$ is complete on M^+.*

Proof. [of the lemma] By the Hopf-Rinow Theorem (Theorem 3.68) it is enough to prove that every geodesic is extendable to the whole real axis. It suffices to consider the geodesics with unit velocity vector norm. The other geodesics can be obtained from these by linear changes of time.

Let $c(s)$ be a geodesic with unit velocity vector norm, i.e., $\|\dot c(s)\|_2=1$ for all s. One can easily see that $\frac{\mathrm{d}}{\mathrm{d}s}f(c(s))=\dot c(s)f=(p_m\dot c(s))f+(p_X\dot c(s))f$.

Since $\|\dot{c}(s)\|_2 = 1$ and since $p_m \dot{c}(s)$ and $p_X \dot{c}(s)$ are orthogonal to each other with respect to the metric $\langle\cdot,\cdot\rangle_2$, $\|p_m \dot{c}(s)\|_2 \leq 1$ and $\|p_X \dot{c}(s)\|_2 \leq 1$. Hence,

$$\begin{aligned}\left|\frac{\mathrm{d}}{\mathrm{d}s} f(c(s))\right| &\leq \left|\frac{p_m \dot{c}(s)}{\|p_m \dot{c}(s)\|_2} f\right| + \left|\frac{p_X \dot{c}(s)}{\|p_X \dot{c}(s)\|_2} f\right| \\ &= \left|\frac{1}{\varphi(c(s))} \frac{p_m \dot{c}(s)}{\|p_m \dot{c}(s)\|_1} f\right| + |X^+ f| < 2\end{aligned}$$

by construction of the functions φ and f.

Thus, the function $f\big(c(s)\big)$ is bounded on any finite interval (a,b) and, since f is proper, the set of points $c(s)$ for $s \in (a,b)$ is relatively compact. This proves the unlimited extendability of the geodesic. □

As mentioned above, $\|X^+\|_2 = 1$, which completes the proof of Theorem 3.36. □

Remark 3.38. If the field $X(t,m)$ is C^k-smooth on M^+, then the above construction gives a C^k-smooth complete Riemannian metric on M^+, with respect to which X^+ is bounded.

3.1.4 Some sufficient conditions

As already mentioned above, sufficient conditions for the completeness of a vector field are known and, when they hold, it is easy to find a Riemannian metric such that the vector field is bounded. In this subsection, we construct a complete Riemannian metric using a Wintner type hypothesis. Later, this metric will be used to study complicated differential equations (stochastic, with delay, etc).

Let us recall the classical Wintner theorem (see, e.g., [144]). Consider the following differential equation on the Euclidean space $\mathbb{R}^n$:

$$\dot{x} = f\big(t, x(t)\big) \tag{3.15}$$

where $f(x,t)$ is continuous in (t,x).

Theorem 3.39 (Wintner). *Suppose that*

$$\big\|f(x,t)\big\| \leq \varphi(t) \cdot L\big(\|x\|\big),$$

where the function $\varphi(t)$ is positive and integrable on any finite interval $[0,l]$, and $L\colon [0,\infty) \to (0,\infty)$ is continuous and satisfies the condition

$$\int_0^\infty \frac{\mathrm{d}u}{L(u)} = \infty. \tag{3.16}$$

Then all solutions of (3.36) are defined on $(-\infty,\infty)$.

The Wintner theorem will be derived from the following result.

Theorem 3.40 *Let M be a complete Riemannian manifold with Riemannian metric $\langle\,,\rangle$ and $L\colon [0,\infty) \to (0,\infty)$ be a smooth function satisfying* (3.16)*. Choose a point $m_0 \in M$ and define a Riemannian metric $\langle\cdot,\cdot\rangle^*$ by*

$$\langle\cdot,\cdot\rangle^*_m = \frac{1}{L^2\big(\rho(m_0,m)\big)}\langle\cdot,\cdot\rangle_m \tag{3.17}$$

at a point $m \in M$, where ρ is the Riemannian distance on M in the metric $\langle\cdot,\cdot\rangle$. If $\langle\cdot,\cdot\rangle$ is complete, then so is $\langle\cdot,\cdot\rangle^*$.

Lemma 3.41 *The minimal geodesics of $\langle\cdot,\cdot\rangle$ and $\langle\cdot,\cdot\rangle^*$ starting at m_0 coincide to within a parametrization.*

Proof. The lemma can be proved with a straightforward calculation: it can be shown that the minimal geodesics (beginning at m_0) of the first metric satisfy the geodesic equation of the second metric and the cut loci of both metrics coincide.

Let us continue the proof of the theorem. It suffices to prove that the metric ball $\mathcal{U}_T$, centered at m_0, of any fixed radius T in the metric $\langle\cdot,\cdot\rangle^*$ is compact. (By the Hopf-Rinow theorem, this implies completeness.) Suppose the contrary, i.e., assume for some l that the ball $\mathcal{U}_l$ is not compact. Since $\langle\cdot,\cdot\rangle$ is a complete metric, it follows by Lemma 3.41 that $\mathcal{U}_l$ contains a minimal geodesic $u(s)$ of infinite length in the metric $\langle\cdot,\cdot\rangle$, beginning at m_0. Here s is the natural parameter (the length) in the metric $\langle\cdot,\cdot\rangle$. Denote the reparametrization of $u(s)$ with the length in $\langle\cdot,\cdot\rangle^*$ by $u(t)$. By Lemma 3.41, $u(t)$ is a minimal geodesic of the metric $\langle\cdot,\cdot\rangle^*$.

The definition of the metric $\langle\cdot,\cdot\rangle^*$ yields

$$\frac{\mathrm{d}s}{\mathrm{d}t} = L\left(u(s)\right) = L(s)$$

and since $u(s)$ lies in $\mathcal{U}_l$,

$$\int_0^\infty \frac{\mathrm{d}s}{L(s)} \le \int_0^l \mathrm{d}t = l.$$

This contradicts condition (3.16). □

Corollary 3.42 *Let M be a complete Riemannian manifold, $X(t,m)$ a vector field continuous in (t,m), and $L\colon [0,\infty) \to (0,\infty)$ a continuous function satisfying* (3.16)*. Suppose there exists a point $m_0 \in M$ such that at every point $m \in M$ the inequality $\big\|X(t,m)\big\| < \varphi(t)\cdot L\big(\rho(m_0,m)\big)$ holds, where ρ is the Riemannian distance on M and φ is a positive function integrable on every finite interval. Then the field $X(t,m)$ is complete.*

Proof. It is obvious that for any constant $C > 0$, the function $L + C$ satisfies (3.16). There exists a smooth function $\Psi(u)$ such that $L < \Psi < L + C$ for

$u \in [0, \infty)$. Clearly, Ψ satisfies (3.16). Let $\langle \cdot, \cdot \rangle$ be a Riemannian metric on M. Consider $\langle \cdot, \cdot \rangle_m^* = \Psi^{-2}(m)\big(\rho(m_q, m)\big)\langle \cdot, \cdot \rangle_m$. By Theorem 3.40, $\langle \cdot, \cdot \rangle^*$ is complete. Now it suffices to observe that the length of every integral curve of the field $X(t, m)$ in the metric $\langle \cdot, \cdot \rangle^*$ is bounded above on any interval $[a, b)$ by $\int_a^b \varphi(t)\,\mathrm{d}t$. □

Wintner's theorem is a particular case of the corollary, where M is the Euclidean space $\mathbb{R}^n$.

3.2 Integral Operators with Parallel Translation

Ordinary differential equations on a vector space can be turned into equivalent Volterra type integral equations. In fact, this method is very often used in the investigation of ordinary differential equations. For example, one can turn the Cauchy problem $\dot{x} = f\big(t, x(t)\big)$, $x(0) = x_0$ in $\mathbb{R}^n$ into the integral equation $x(t) = x_0 + \int_0^t f\big(\tau, x(\tau)\big)\,\mathrm{d}\tau$. To do so, we use the fact that for any continuous curve $y(t)$ there exists a unique curve $z(t) = \int_0^t f\big(\tau, y(\tau)\big)\,\mathrm{d}\tau$ such that the derivative of $z(t)$ at any point t is equal to $f(t, y(t))$. The existence of the curve $z(t)$ is possible only because of global parallelism on the tangent bundle to the vector space. Indeed, the vectors $\dot{z}(t)$ and $f(t, y(t))$ belong to the tangent spaces at different points, and the equation $\dot{z}(t) = f(t, y(t))$ makes sense only by virtue of the existence of global parallelism, i.e., a canonical isomorphism between all tangent spaces and the vector space itself.

Global parallelism does not exist on an arbitrary manifold. Therefore classical integrals can be used only locally (in charts) and, moreover, the integrals themselves depend on the choice of the coordinate system. In this section, following [86, 88, 94], we describe a construction of an analog of the integral operator, in which global parallelism is replaced with parallel translation (with respect to a connection) along a chosen curve. For simplicity we use the Levi-Civitá connection on a complete finite-dimensional manifold.

Similar notions of absolute and covariant integrals were introduced in a different way by Vujičić. (See, e.g., [227, 228, 229] and the bibliography in [229].) There, the integral is defined in a local coordinate system with the connection coefficients used in such a way that the integral becomes covariant with respect to changes of coordinates.

3.2.1 The operator $\mathcal{S}$

Let M be a complete Riemannian manifold, let $m_0 \in M$, $I = [0, l]$ and let $v\colon I \to T_{m_0}M$ be a continuous curve.

Theorem 3.43 *There exists a unique C^1-curve $\gamma\colon I \to M$ such that $\gamma(0) = m_0$ and the tangent vector $\dot\gamma(t)$ is parallel to the vector $v(t) \in T_{m_0}M$ for every $t \in I$.*

Proof. Let $b_0 = (e_1^0, \dots, e_n^0)$ be a basis in the tangent space $T_{m_0}M$; b_0 gives rise to an isomorphism between $\mathbb{R}^n$ and $T_{m_0}M$ by the formula $b_0(x^1, \dots, x^n) = x^1 e_1^0 + \ldots + x^n e_n^0$, where $\mathbb{R}^n$ is the model space of M.

Consider the time-dependent basic vector field $E(b_0^{-1}v(t))$ on the frame bundle $B(M)$. Clearly, this field is locally Lipschitz in $b \in B(M)$. Hence, for every point $b \in B(M)$, there exists a unique integral curve $b(t)$ passing through b, $b(0) = b$. The curve $\gamma(t) = \pi b_0(t)$ is the one we are looking for. (Here π is the natural projection of $B(M)$ to M.) Indeed, for any point t^* in the domain of $\gamma(\cdot)$, the vectors $\dot\gamma(t^*)$ and $v(t^*)$ are connected along $\gamma(\cdot)$ by the parallel vector field $b_0(t)\big(b_0^{-1}v(t^*)\big)$.

It remains to prove that $\gamma(\cdot)$ is defined on the whole interval $[0, l]$. Since the metric on M is complete, the metric ball of radius $\int_0^l \|v(s)\|\, \mathrm{d}s$ centered at m_0 is compact. Now assume that $\gamma(t)$ is defined on $[\,0, t^*)$, where $t^* \in [0, l]$. The length of $\gamma(\cdot)$ on $[\,0, t^*)$ equals

$$\int_0^{t^*} \|v(s)\|\ \mathrm{d}s \le \int_0^l \|v(s)\|\ \mathrm{d}s,$$

i.e., $\gamma(t)$ belongs to a compact set and, therefore, can be extended to $[0, t^*]$. It is clear that one can extend $\gamma(\cdot)$ to a neighborhood of t^*. Thus, the domain of $\gamma(\cdot)$ is open and closed in $[0, l]$, i.e., it coincides with $[0, l]$. The theorem is proved. $\square$

In what follows, we denote by $\mathcal{S}v(\cdot)$ the curve γ constructed as above beginning with v.

Remark 3.44. It is important to emphasize that $\mathcal{S}(v(t))$, for a continuous curve $v \in C^0(I, T_{m_0}M))$, is independent of the basis b_0 used in the proof of Theorem 3.43. To see this, let us go back to the construction of $\gamma(\cdot)$ and replace the basis b_0 by b_1 in $T_{m_0}M$. Since there exists an invertible $n \times n$ matrix $\mathbf{b}$ such that $b_1 = b_0 \circ \mathbf{b}$ and since the connection is invariant with respect to the right action of $GL(n, \mathbb{R})$ on BM, it follows from the definition of a basic vector field that $\pi b_1(t) = \gamma(t)$.

Remark 3.45. Let $m(t)$ be a C^1-smooth curve in M, $t \in I$, $m(0) = m_0$. Denote by $\varGamma$ the operator of parallel translation along $m(\cdot)$ at $T_{m_0}M$. The curve $C(m)(t) = \int_0^t \varGamma \dot m(s)\mathrm{d}s$ is known as *Cartan's development* of $m(t)$ at $T_{m_0}M$. The curve $\mathcal{S}v(\cdot)$ introduced above is expressed via Cartan's development as $\mathcal{S}v(t) = C^{-1}(\int_0^t v(s)\mathrm{d}s)$.

Consider the Banach space $C^0(I, T_{m_0}M)$ of continuous maps from I to $T_{m_0}M$ and the Banach manifold $C^1(I, M)$ of C^1-smooth maps from I to M.

As follows from Theorem 3.43, the operator $\mathcal{S}\colon C^0(I, T_{m_0}M) \to C^1(I, M)$ is well-defined. If M is a Euclidean space, $\mathcal{S}v$ is a primitive of v.

It is easy to see that $\mathcal{S}$ is a homeomorphism between $C^0(I, T_{m_0}M)$ and its image $C^1_{m_0}(I, M)$ in $C^1(I, M)$, where the manifold $C^1_{m_0}(I, M)$ consists of all C^1-curves γ with $\gamma(0) = m_0$.

Theorem 3.46 *Let $\mathcal{U}_K$ be the ball of radius K centered at the origin of $C^0(I, T_{m_0}M)$. Then, at every point $t \in I$, the inequality $\|\dot{\gamma}(t)\| \leq K$ holds for all curves $\gamma(\cdot)$ in the set $\mathcal{S}\mathcal{U}_K$.*

This is obvious since parallel translation preserves the norm of a vector.

Theorem 3.47 *Assume that the point $m_1 \in M$ is not conjugate to m_0 along some geodesic of the Levi-Civitá connection on M. Then for any geodesic $\alpha(\cdot)$, $\alpha(0) = m_0$, $\alpha(1) = m_1$, along which m_0 and m_1 are not conjugate, and for any $K > 0$, there exists a constant $\bar{L}(m_0, m_1, K, \alpha) > 0$ with the following property: for any t_1, $0 < t_1 < \bar{L}(m_0, m_1, K, \alpha)$, and for any curve $u(\cdot) \in \mathcal{U}_K \subset C^0\big([0, t_1], T_{m_0}M\big)$, there exists a unique vector $C_u \in T_{m_0}M$, such that $\mathcal{S}(u + C_u)(t_1) = m_1$, which belongs to a bounded neighborhood of $t_1^{-1} \cdot \dot{\alpha}(0) \in T_{m_0}M$ and depends continuously on u.*

Proof. We divide the proof into two lemmas. Without loss of generality we assume that the parameter t on $g(t)$ is chosen so that $g(0) = m_0$ and $g(1) = m_1$.

Lemma 3.48 *There exists a ball $\mathfrak{U}_\varepsilon \subset C^0([0, 1], T_{m_0}M)$ with radius $\varepsilon > 0$ such that for any curve $\hat{u}(t) \in \mathfrak{U}_\varepsilon \subset C^0([0, 1], T_{m_0}M)$ there exists a unique vector $\mathbf{C}_{\hat{u}}$, belonging to a bounded neighborhood V of the vector $\dot{\alpha}(0)$ in $T_{m_0}M$, that is continuous in $\hat{u}$ and such that $\mathcal{S}(\hat{u} + \mathbf{C}_{\hat{u}})(1) = m_1$.*

Proof. [of Lemma 3.48] Consider the mapping from $C^0\big([0, 1], T_{m_0}M\big) \times T_{m_0}M$ to M sending a pair (u, C), where $u \in C^0\big([0, 1], T_{m_0}M\big)$ and $C \in T_{m_0}M$, to the point $\mathcal{S}(u + C)(1)$. Note that the vector field $E(b_0^{-1}(v(t))$ on $B(M)$ is smooth in $v(\cdot)$. (See the proof of Theorem 3.43.) Using this fact, the definition of $\mathcal{S}$, and the classical theorem on the smooth dependence of solutions of differential equations on parameters, one can easily show that this map is jointly smooth in u and C. Clearly, we have $\mathcal{S}(C)(1) = \exp(C)$ when $u = 0$. Thus, by the hypotheses of the theorem, $\mathcal{S}\big(\dot{\alpha}(0)\big)(1) = m_1$ and $\mathcal{S}(C)(1)$ is a diffeomorphism from a neighborhood of $\dot{\alpha}(0)$ in $T_{m_0}M$ onto a neighborhood of m_1 in M.

Let us now think of $\mathcal{S}(u + C)(1)$ as a perturbation of $\mathcal{S}(C)(1) = \exp(C)$. Thus, there exists a $\varepsilon > 0$ such that for any fixed $\hat{u} \in \mathcal{U}_\varepsilon \subset C^0\big([0, 1], T_{m_0}M\big)$ the operator $\mathcal{S}(\hat{u} + C)(1)$ is a local diffeomorphism. Therefore, there is a ball $D \subset T_{m_0}M$ centered at $\dot{\alpha}(0)$ such that, for any $\hat{u} \in \mathcal{U}_\varepsilon$, there exists a vector $C_{\hat{u}} \in D$ solving the equation $\mathcal{S}(\hat{u} + C_{\hat{u}})(1) = m_1$. Using the implicit function theorem one may show that, when D is sufficiently small, the vector $C_{\hat{u}} \in D$ is unique and $C_{\hat{u}}$ depends continuously on $\hat{u}$. □

We introduce the notation $\sup_{\mathbf{C}\in V} \|\mathbf{C}\| = C$ where V is as defined in Lemma 3.48.

Remark 3.49. One can easily show that $\varepsilon < C$.

Lemma 3.50 *In the conditions and notation of Lemma 3.48 let $K > 0$ and $t_1 > 0$ be such that $t_1^{-1}\varepsilon > K$. Then for any curve $u(t) \in \mathfrak{U}_K \subset C^0([0,t_1], T_{m_0}M)$ there exists a unique vector C_u in a neighborhood $t_1^{-1}V$ of the vector $t_1^{-1}\dot{\gamma}(0)$ in $T_{m_0}M$, continuously depending on u and such that $S(u + C_u)(t_1) = m_1$.*

Proof. [of Lemma 3.50] For $u(t) \in \mathfrak{U}_K \subset C^0([0,t_1], T_{m_0}M)$ let $\hat{u}(t) = t_1 u(t_1 \cdot t) \in \mathfrak{U}_\varepsilon \subset C^0([0,1], T_{m_0}M)$ and $C_u = t_1^{-1}\mathbf{C}_{\hat{u}}$. From Lemma 3.48 we get $\mathcal{S}(\hat{u} + \mathbf{C}_{\hat{u}})(1) = m_1$ and $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(\hat{u} + \mathbf{C}_{\hat{u}})(t)$ is parallel to $\hat{u}(t) + \mathbf{C}_{\hat{u}}$. For the curve $\gamma(t) = \mathcal{S}(\hat{u} + \mathbf{C}_{\hat{u}})(t \cdot t_1)$ we have $\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t) = t_1^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(\hat{u} + \mathbf{C}_{\hat{u}})(t \cdot t_1)$ and this vector is parallel along the same curve to the vector $t_1^{-1}(\hat{u}(t) + \mathbf{C}_{\hat{u}}) = u(t) + C_u$. Thus $\gamma(t) = \mathcal{S}(u + C_u)(t) = \mathcal{S}(\hat{u} + \mathbf{C}_{\hat{u}})(t \cdot t_1^{-1})$ for $t \in [0,t_1]$. Hence $\mathcal{S}(u + C_u)(t_1) = \mathcal{S}(\hat{u} + \mathbf{C}_{\hat{u}})(1) = m_1$. ☐

This completes the proof of Theorem 3.47. ☐

Lemma 3.51 *For specified $t_1 > 0$ and $K > 0$ all curves $S(v(t) + C_v)(t)$ with $v \in \mathfrak{U}_K \subset C^0([0,t_1]T_{m_0}M)$ lie in a compact set $\Xi \subset M$, where Ξ depends on ε and C is as defined above.*

Indeed, since parallel translation preserves the norm of a vector, for any $v(t)$ as above, the length of $S(v(t) + C_v)(t)$ is no greater than $\int_0^{t_1}(K + \|C_v\|)\mathrm{d}t \le \int_0^{t_1} t_1^{-1}(\varepsilon + C)\mathrm{d}t = \int_0^1(\varepsilon + C)\mathrm{d}t = \varepsilon + C$. Since M is complete, by the Hopf-Rinow theorem any closed metric ball of finite radius $\varepsilon + C$ is compact.

Remark 3.52. Note that if M is a Euclidean space, one can take any constant as ε in the proof of Theorem 3.3., i.e., the theorem holds for every t_1, $0 < t_1 < \infty$.

3.2.2 The operator Γ

Let $\gamma(t)$, $t \in I$, be a C^1-curve in M and $X(\gamma(t))$ a continuous vector field along $\gamma(\cdot)$. Consider the curve $\Gamma X\big(\gamma(t)\big)$ in $T_{\gamma(0)}M$ where Γ is the operator of parallel translation along $\gamma(\cdot)$ at $\gamma(0)$.

Lemma 3.53 (Compactness lemma). *Let $\Xi \subset C^0(I, TM)$ be such that $\pi\Xi \subset C^1(I, M)$, where $\pi\colon TM \to M$ is the natural projection, and the norms of the derivatives of the curves $\pi\Xi$ are uniformly bounded by a certain constant $K > 0$. If Ξ is relatively compact in $C^0(I, TM)$, then so is $\Gamma\Xi$.*

Proof. Since the closure $\bar{\Xi}$ of the set Ξ is compact in $C^0(I, TM)$, the vectors of $\{\xi(t) \mid \xi(\cdot) \in \Xi\}$ are bounded and the set $\pi\bar{\Xi}$ is compact in $C^0(I, M)$. This implies, in particular, that all curves in $\pi\Xi$ lie in a compact subset of‘M.

Let $\xi^*(\cdot)$ be a limit curve of the set Ξ. It is clear that the inequality

$$\rho\left(\pi\xi^*(t), \pi\xi^*(t')\right) \leq K\left|t - t'\right|,$$

where $t, t' \in I$ and ρ is the Riemannian distance, holds for the curve $\pi\xi^*(\cdot)$. Note that this curve may not be smooth. Parallel translation along such curves was defined in [29] as the limit of parallel translations along their piecewise geodesic approximations. Moreover, it was shown that under the hypothesis above, the procedure of parallel translation converges uniformly on any bounded set of vectors. If the curve is smooth, the new definition of parallel translation is equivalent to the classical one. It was also shown that the parallel translation along a limit curve is just the limit of parallel translations along curves converging to it. Therefore, Γ sends convergent sequences to convergent ones. □

If $X(\gamma(t)) = X(t, \gamma(t))$ is the restriction to $\gamma(\cdot)$ of a continuous vector field $X(t, m)$, $t \in I$ and $m \in M$, then we use the notation $\Gamma \circ \gamma$ for $\Gamma X(t, \gamma(t))$. Thus, for a specified vector field $X(t, m)$, we may consider the operator $\Gamma\colon C^1(I, M) \to C^0(I, TM)$, which is clearly continuous.

Let Ω_K be the set of curves from $C^1(I, M)$ satisfying the inequality $\|\dot{\gamma}(t)\| \leq K$, where $K > 0$, at every point $t \in I$ and such that the set $\{\gamma(0) \mid \gamma(\cdot) \in \Omega_K\}$ is bounded in M.

Theorem 3.54 *The set of curves $\Gamma(\Omega_K)$ is relatively compact in $C^0(I, TM)$.*

Proof. Because M is complete, it is clear that Ω_K is relatively compact in $C^0(I, M)$. Since the field $X(t, m)$ is continuous, the set of curves $\{X(t, \gamma(t)) \mid \gamma(\cdot) \in \Omega_K\}$ is relatively compact in $C^0(I, TM)$. The theorem follows from Lemma 3.53. □

Corollary 3.55 *The operator Γ is locally compact.*

Proof. For every $\gamma \in C^1(I, M)$, the continuous function $\|\dot{\gamma}(t)\|$ assumes its supremum K_γ on I. By the definition of the C^1-topology, the inequality $\|\dot{\gamma}_1\| < K_\gamma + \varepsilon$ holds for every $\gamma_1(\cdot)$ in a small neighborhood of $\gamma(\cdot)$. □

3.2.3 Integral operators

Consider the continuous composition operator

$$\mathcal{S} \circ \Gamma\colon C^1_{m_0}(I, M) \to C^1_{m_0}(I, M).$$

Theorem 3.56 *The fixed points of $\mathcal{S} \circ \Gamma$ are precisely the integral curves of the field $X(t, m)$ with the initial condition $\gamma(0) = m_0$.*

Proof. Let $\gamma(t)$ be an integral curve of the field $X(t, m)$, i.e., $\dot{\gamma}(t) = X\big(t, \gamma(t)\big)$. Then the operator Γ on $\gamma(\cdot)$ is equal to $\mathcal{S}^{-1}$, and so γ is a fixed point of $\mathcal{S} \circ \Gamma$. Conversely, let $\gamma(\cdot)$ be a fixed point of the operator $\mathcal{S} \circ \Gamma$. Using the parallel translation along $\gamma(\cdot)$, we transport the vector $X\big(t, \gamma(t)\big)$ to $\gamma(0) = m_0$ and then back to $\gamma(t)$. The resulting vector coincides with $\dot{\gamma}(t)$ by the definitions of $\mathcal{S}$ and Γ. Therefore, $\dot{\gamma}(t) = X\big(t, \gamma(t)\big)$. □

Thus, $\mathcal{S} \circ \Gamma$ is a direct analog of the standard Urysohn-Volterra integral operator from the theory of ordinary differential equations on vector spaces.

Theorem 3.57 *The operator $\mathcal{S} \circ \Gamma$ is locally compact.*

The assertion of Theorem 3.57 follows from the local compactness of Γ and the continuity of $\mathcal{S}$.

Let Θ be a closed bounded set in M and $C^1_{m_0}(I, \bar{\Theta})$ the subset in $C^1_{m_0}(I, M)$ formed by curves lying in the closure of Θ. Consider the second iteration $(\mathcal{S} \circ \Gamma)^2$ of the operator $\mathcal{S} \circ \Gamma$.

Theorem 3.58 *The set $(\mathcal{S} \circ \Gamma)^2 C^1_{m_0}(I, \bar{\Theta})$ is compact in $C^1_{m_0}(I, M)$.*

Proof. Since Θ is bounded and M is complete, Θ is compact. Therefore, $\|X(t, m)\|$ is bounded on $I \times \Theta$ by some constant K. Since parallel translation preserves the norm, the curves $\Gamma C^1_{m_0}(I, \bar{\Theta})$ lie in $\bigcup_{m \in \Theta} \mathcal{U}_K(m)$ and, by Theorem 3.46, the set $\mathcal{S} \circ \Gamma C^1_{m_0}(I, \bar{\Theta})$ is formed by curves which satisfy the inequality $\|\dot{\gamma}(t)\| \le K$ at every point $t \in I$. Now the theorem follows from Theorem 3.54 and the continuity of the operator $\mathcal{S}$. □

Composition operators, such as $\mathcal{S} \circ \Gamma$, are employed to solve certain problems in the theory of differential equations on manifolds (for example, to find periodic solutions for some special classes of differential equations). The construction of such operators and their applications are described, e.g., in the survey [33]. The theory of topological characteristics is also developed in [28, 33, 109] for a large class of maps of infinite-dimensional manifolds. This theory enables one to prove the existence of fixed points of these operators.

Let us discuss one more class of integral operators that can be used to reduce certain problems on manifolds to problems on vector spaces. Consider the composition $\Gamma \circ \mathcal{S}$. This operator is continuous and acts on the Banach space $C^0(I, T_{m_0}M)$. If $v = \Gamma \circ \mathcal{S}v$, then $\mathcal{S}v = \mathcal{S} \circ \Gamma \circ \mathcal{S}v = (\mathcal{S} \circ \Gamma)\mathcal{S}v$ is a fixed point of $\mathcal{S} \circ \Gamma$ and so, an integral curve of the field $X(t, m)$. Conversely, $\mathcal{S}v = \mathcal{S} \circ \Gamma(\mathcal{S}v)$ implies that $v = \Gamma \circ \mathcal{S}v$ because $\mathcal{S}$ is one-to-one.

Theorem 3.59 *The operator $\Gamma \circ \mathcal{S}$ is completely continuous.*

Proof. Let $\mathcal{U}_K$ be a ball of radius K in $C^0(I, T_{m_0}M)$. By Theorem 3.46, $\mathcal{S}\mathcal{U}_K \subset \Omega_K$ and, by Theorem 3.54, the set $\Gamma \circ \mathcal{S}\mathcal{U}_K$ is compact. □

Remark 3.60. As mentioned in the introduction to this section, we consider the Levi-Civitá connection of a complete Riemannian metric only to simplify the presentation of the material. Under certain hypotheses, the constructions of the integral operators may be generalized to other connections. Note, for example, that we have never used the fact that the connection has zero torsion, i.e., all of our constructions hold for any Riemannian connection of a complete Riemannian metric. In particular, the construction leads to the classical multiplicative integral for a special choice of the connection on a Lie group. (See, e.g., [78] for matrix groups.)

3.3 Second Order Differential Equations (Special Vector Fields)

Let M be a manifold with tangent bundle TM. On the manifold TM there is a class of vector fields that is naturally coordinated with the bundle structure of TM. This class describes the second order differential equations on M.

Recall the well-known trick of reducing a second order differential equation in $\mathbb{R}^n$ to a first order differential equation in $\mathbb{R}^{2n}$: the differential equation $\ddot{x} = f(t, x, \dot{x})$ in $\mathbb{R}^n$ is equivalent to the system of first order differential equations

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= f(t, x, y) \end{aligned} \tag{3.18}$$

in $\mathbb{R}^{2n}$. We emphasize that a general system in $\mathbb{R}^{2n}$ has the form

$$\begin{aligned} \dot{x} &= f_1(t, x, y) \\ \dot{y} &= f_2(t, x, y), \end{aligned}$$

i.e., (3.18) is a system of a special form.

It is clear that in every chart of a manifold a second order differential equation must be reduced to a special system of first order equations on the tangent bundle. It turns out that such systems are vector fields of a special type connected with the bundle structure. Let us introduce the exact definition.

Definition 3.61. A vector field $Y(t, (m, X))$ on the tangent bundle TM is called a *special vector field* on TM or a *second order differential equation* on M if at every point $(m, X) \in TM$ the equality

$$T\pi Y(t, (m, X)) = X_m \tag{3.19}$$

holds where $\pi : TM \to M$ is the natural projection of TM onto M.

Recall that by Convention 1.3 X_m and (m, X) are two equivalent designations of the same object. Below, if it is not necessary to emphasize that the vector $Y(t,(m,X))$ is given at $(m,X) \in TM$, we shall also denote it by $Y(t,m,X)$. Represent a vector Y at a point (m,X) on TM as a quadruple

$$Y(t,(m,X)) = (m, X, Y_1(t,(m,X)), Y_2(t,(m,X))),$$

as described in Section 2.1. By applying formula (2.3) we find $T\pi Y(t,(m,X))$ and substituting into the above, we obtain:

$$T\pi(m, X, Y_1(t,(m,X)), Y_2(t,(m,X))) = (m, Y_1(t,(m,X))) = (m,X).$$

Thus, if Y is a second order differential equation $Y_1(t,(m,X)) = X$, and so the presentation as a quadruple takes the form

$$Y_{(m,X)} = (m, X, X, Y_2(t,(m,X))). \tag{3.20}$$

We show that second order differential equations (special vector fields), as defined in Definition 3.61, do indeed yield ordinary second order differential equations in the charts of M. Let $(m(t), X(t))$ be an integral curve of a special vector field Y on TM. This means that

$$\frac{\mathrm{d}}{\mathrm{d}t}(m(t), X(t)) = Y_{(m(t),X(t))} = (m(t), X(t), X(t), Y_2(t,(m(t),X(t)))),$$

i.e. $\frac{\mathrm{d}}{\mathrm{d}t}m(t) = X(t)$ and $\frac{\mathrm{d}}{\mathrm{d}t}X(t) = Y_2(t,(m(t),X(t)))$. Thus the curve $m(t) = \pi(m(t), X(t))$ on M (the projection of the integral curve of Y) satisfies the equation $\ddot{m}(t) = Y_2(t, m(t), X(t))$ in the charts. The curve $m(t)$ in M is called a solution of the second order differential equation Y. It follows from the above arguments that the integral curve of Y on TM is represented via $m(t)$ by the formula $(m(t), \dot{m}(t))$ where $\dot{m}(t) = \frac{\mathrm{d}}{\mathrm{d}t}m(t)$.

Let a connection H be given on a manifold M. Recall (see Lemma 2.12) that at any point $(m,X) \in TM$ the mapping $T\pi$ is a linear isomorphism of $\mathsf{H}_{(m,X)}$ onto T_mM. Hence in $\mathsf{H}_{(m,X)}$ there is a unique vector $\mathcal{Z}_{(m,X)}$ such that

$$T\pi\mathcal{Z}_{(m,X)} = X_m. \tag{3.21}$$

On constructing such a vector at each point $(m,X) \in TM$, we obtain the vector field $\mathcal{Z}$ on TM. From formula (3.21) it immediately follows that $\mathcal{Z}$ is a second order differential equation (special vector field, see Definition 3.61).

Definition 3.62. The second order differential equation $\mathcal{Z}$ constructed above is called the *geodesic spray* of the connection H.

Remark 3.63. A *spray* is a second order differential equation Y that has the following property: for any real number $a \in \mathbb{R}$ and every point $(m,X) \in TM$ the equality $Y_{(m,aX)} = Ta(aY_{(m,X)})$ holds (see formula (2.7) where the action of the real line on a vector bundle is defined). One can easily verify that $\mathcal{Z}$ satisfies the definition of a spray.

For a given connection H the geodesic spray $\mathcal{Z}$ is a universal object for the description of all second order differential equations. Let Y be such an equation. Since every tangent space to TM is presented as direct sum $T_{(m,X)}TM = \mathsf{H}_{(m,X)} \oplus \mathsf{V}_{(m,X)}$, there is a unique decomposition $Y_{(m,X)} = \mathsf{H}Y_{(m,X)} + \mathsf{V}Y_{(m,X)}$ where $\mathsf{H}Y_{(m,X)} \in \mathsf{H}_{(m,X)}$ is called the *horizontal component* of $Y_{(m,X)}$ and $\mathsf{V}Y_{(m,X)} \in \mathsf{V}_{(m,X)}$ is the *vertical component* (see Section 2.2).

Proposition 3.64 *For any second order differential equation Y its horizontal component is $\mathcal{Z}$.*

Proof. By the definition of second order differential equations, $T\pi Y_{(m,X)} = X_m$. Since $\mathsf{V}_{(m,X)}$ is the kernel of $T\pi$ (see Lemma 2.12), we get that $T\pi \mathsf{V}Y_{(m,X)} = 0$ and so $T\pi\mathsf{H}Y_{(m,X)} = T\pi Y_{(m,X)} = X_m$. But in $\mathsf{H}_{(m,X)}$, as mentioned above, there is unique vector $\mathcal{Z}_{(m,X)}$ having this property. Thus $\mathsf{H}Y_{(m,X)} = \mathcal{Z}_{(m,X)}$. □

Let a curve $m(t)$ in M be a solution of the second order differential equation Y, i.e., $(m(t), \dot{m}(t))$ is an integral curve of the vector field Y on TM, $\frac{\mathrm{d}}{\mathrm{d}t}(m(t), \dot{m}(t)) = Y_{(m(t),\dot{m}(t))}$ (see above). Apply the connector K of the connection H to both sides of this equality. Since $K \circ \frac{\mathrm{d}}{\mathrm{d}t} = \frac{\mathrm{D}}{\mathrm{d}t}$ and $KY = \mathbf{p}\mathsf{V}Y$ (see Section 2.2), we obtain

$$\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) = \mathbf{p}\mathsf{V}Y. \tag{3.22}$$

On the other hand, for the curve $m(t)$ the derivative of the corresponding curve $(m(t), \dot{m}(t))$ in TM is always a vector of some second order differential equation for each t (i.e., property (3.19) is fulfilled). By Proposition 3.64, once a connection on M given, any second order differential equation is uniquely defined by its vertical component. If in addition $m(t)$ satisfies (3.22), this equation coincides with Y.Thus, we have proved the following:

Proposition 3.65 *The solutions $m(t)$ of the second order differential equation Y, and only these solutions, satisfy equation (3.22).*

Corollary 3.66 *The solutions of the geodesic spray $\mathcal{Z}$, and only these solutions, are geodesics of the connection H.*

Proof. Since $\mathcal{Z}_{(m,X)} \in \mathsf{H}_{(m,X)}$ and $\mathsf{H}_{(m,X)}$ is the kernel of K (see Lemma 2.14(ii)), one obtains that $K\mathcal{Z} = 0$ and equation (3.22) takes the form $\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) = 0$, i.e., $m(t)$ satisfies (2.24). □

Corollary 3.66 together with Remark 3.63 clarify the name “geodesic spray” for $\mathcal{Z}$.

Consider the equation

$$\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) = Y(t, m(t), \dot{m}(t)), \tag{3.23}$$

where $Y(t, m, X)$ is a non-autonomous vector field on M which depends, at each point $m \in M$, on the vector parameter $X \in T_m M$. Denote by Y^l the vector field on TM that at $(m, X) \in TM$ and $t \in \mathbb{R}$ is the vertical lift of $Y(t, m, X)$ at the point (m, X) (see Definition 2.34). Evidently we obtain

Proposition 3.67 *If a curve $m(t)$ satisfies (3.23), it is a solution of the second order differential equation $Y^l + \mathcal{Z}$.*

Without proof we present the following classical result of finite-dimensional Riemannian geometry (for a proof, see e.g., [26, 161]).

Theorem 3.68 (Hopf-Rinow theorem) *For a finite-dimensional Riemannian manifold M the following four statements are equivalent:*

(i) *M is a complete Riemannian manifold (see Definition 1.49), i.e., it is complete as a metric space with respect to the Riemannian distance ρ;*

(ii) *every set that is bounded with respect to the Riemannian distance ρ is relatively compact;*

(iii) *the Levi-Civitá connection is complete in the sense of Definition 2.40;*

(iv) *the geodesics of the Levi-Civitá connection, starting at some specified point, are well-defined for all $t \in (-\infty, \infty)$.*

From equivalent statements (i)–(iv) it follows that:

(v) *for every pair of points $m_0, m_1 \in M$ there exists a geodesic of the Levi-Civitá connection that joins them and whose length is equal to $\rho(m_0, m_1)$.*

Remark 3.69 (Hamiltonian systems). The additional vector bundle structure on the tangent bundle TM yielded above the special class of vector fields on TM, second order differential equations. Analogously, the additional structure on the cotangent bundle yields a single special object called the *canonical 1-form*, the differential form θ whose value on a vector $Y \in T_{(m,\alpha)}T^*M$ at a point $(m, \alpha) \in T^*M$ is given by the formula $\theta_{(m,\alpha)}(Y) = \alpha_m(T\pi Y)$, where $\pi : T^*M \to M$ is the natural projection.

By routine calculation one can easily show that at $\alpha_m = p_i \mathrm{d}q^i$ the canonical 1-form obtains the coordinate presentation $\theta_{(m,\alpha)} = p_i \mathrm{d}q^i$ that coincides with the coordinate expression of α at $m \in M$. For θ the coordinates correspond to the covectors from the first half of the basis in $T^*_{(m,\alpha)}T^*M$ while for α the entire basis in $T^*_m M$ is involved.

The canonical 2-form on T^*M is $\Omega = \mathrm{d}\theta$. Its coordinate presentation is $\Omega = \mathrm{d}p_i \wedge \mathrm{d}q^i$. We obtain directly from the definition that Ω is exact and so it is closed. In addition one can easily prove that Ω is not degenerate, i.e., for every 1-form β at every point $(m, \alpha) \in T^*M$ there exists a unique vector $X_\beta \in T_{(m,\alpha)}T^*M$ such that for every vector $Y \in T_{(m,\alpha)}T^*M$ the equality $\beta(Y) = \Omega(Y, X_\beta)$ holds.

Let $H : T^*M \to R$ be a smooth function and consider its differential $\mathrm{d}H$ (i.e. a covector field). The vector field X_H such that for every vector field Y the equality $\mathrm{d}H(Y) = \Omega(Y, X_H)$ holds is called the *screw gradient* of H or *Hamiltonian vector field* with *Hamiltonian* H. For the screw gradient $X_H = X^i \frac{\partial}{\partial q^i} + \tilde{X}_i \frac{\partial}{\partial p_i}$ the coordinates take the form $X^i = \frac{\partial H}{\partial p_i}$, $\tilde{X}_i = -\frac{\partial H}{\partial q^i}$, and so the integral curve of X_H with coordinates $(q^1(t), \dots, q^n(t), p_1(t), \dots, p_n(t))$ satisfies the system $\frac{\mathrm{d}}{\mathrm{d}t}q^i = \frac{\partial H}{\partial p_i}$, $\frac{\mathrm{d}}{\mathrm{d}t}p_i = -\frac{\partial H}{\partial q^i}$. The latter system is called the *Hamiltonian system* with *Hamiltonian H*.

Hamiltonian systems describe a broad class of processes in physics and mechanics. For conservative mechanical systems (see Section 11.3 below) the Hamiltonian H is the total energy expressed in terms of momenta.

The Hamiltonian H is constant along the integral curves of X_H. Indeed, $X_H H = \mathrm{d}H(X_H) = \Omega(X_H, X_H) = 0$ since Ω is skew symmetric. Taking into account that the usual interpretation of the Hamiltonian in mechanics is the total energy, this is a version of the conservation of energy law.

Consider two smooth functions f and g as Hamiltonians and (as usual) denote by X_f and X_g, respectively, their Hamiltonian vector fields. The function $\{f, g\} = \Omega(X_f, X_g)$ is called the *Poisson bracket* of f and g. Its coordinate representation is easily derived in the form $\{f, g\} = \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i}$.

The Poisson bracket is evidently skew-symmetric. It is easily shown that it satisfies the Jacobi identity (1.8). So, the Poisson bracket defines the structure of a Lie algebra on the space of smooth functions on T^*M.

The Poisson bracket is useful for applications. In physical applications it is more important to find the values of a function along a solution than the solution itself (for example, the solution can describe a process in a chemical reactor and the function of the solution, the temperature of the wall of the reactor). Let $\gamma(t)$ be a solution of a Hamiltonian system with Hamiltonian H and $f : T^*M \to \mathbb{R}$ be a smooth function. Then one can easily show that the real-valued function of the real variable $f(\gamma(t))$ satisfies the equation $\frac{\mathrm{d}}{\mathrm{d}t}f = \{H, f\}$.

The notion of a Hamiltonian system can be generalized to a class of manifolds much broader than cotangent bundles. A manifold M, on which a smooth closed non-degenerate 2-form Ω is given, is called a *symplectic manifold* and in this case Ω is called the *symplectic form*.

Every 2-form on an odd-dimensional manifold is degenerate. Hence all symplectic manifolds are even-dimensional. The construction and results on Hamiltonian systems can be translated to symplectic manifolds. Locally (i.e., in the charts) every symplectic manifold is organized as the cotangent bundle of a certain manifold. This follows from a famous theorem of Darboux.

There is also a broad generalization of Hamiltonian theory to so-called *Poisson manifolds*, i.e., manifolds in the space of smooth functions on which a bracket operation $\{\cdot, \cdot\}$ satisfying the Jacobi identity is given ($\{\cdot, \cdot\}$ is the Poisson bracket). In this case equations of the type $\frac{\mathrm{d}}{\mathrm{d}t}f = \{H, f\}$ are considered instead of ordinary Hamiltonian systems.

We refer the reader, say, to [212] where a detailed geometrical description of Hamiltonian theory is given.

Chapter 4
Elements of the Theory of Set-Valued Mappings

In this chapter we survey some notions in the theory of set-valued mappings which will be used below for the description of complicated mechanical systems such as systems with discontinuous forces, with control, etc.

More details can be found, for example, in [31, 155], where in particular the proofs of many of the results presented here are given.

4.1 Set-Valued Mappings and Differential Inclusions

A *set-valued mapping* F from a set X into a set Y is a correspondence that assigns a non-empty subset $F(x) \subset Y$ to each point $x \in X$. $F(x)$ is called the *value* of x.

In order to distinguish set-valued mappings from single-valued mappings we shall denote a set-valued mapping F sending X to Y by the symbol $F : X \multimap Y$ while for a single-valued mapping we shall retain the notation $f : X \to Y$.

If X and Y are metric spaces, for set-valued mappings there are several different analogs of continuity that in the case of single-valued mappings coincide with the usual definition of continuity (here we do not deal with the description of such a notion for set-valued mappings of topological spaces, see, e.g., [31]).

Definition 4.1. A set-valued mapping F is *upper semicontinuous* at the point $x \in X$ if for each $\varepsilon > 0$ there exists a neighborhood $U(x)$ of x such that from $x' \in U(x)$ it follows that $F(x')$ belongs to the ε-neighborhood of the set $F(x)$. F is *upper semicontinuous* on X if it is upper semicontinuous at every point of X.

Definition 4.2. A set-valued mapping F is *lower semicontinuous* at the point $x \in X$ if for each $\varepsilon > 0$ there exists a neighborhood $U(x)$ of x such that from $x' \in U(x)$ it follows that $F(x)$ belongs to the ε-neighborhood of $F(x')$.

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_4,

F is *lower semicontinuous* on X if it is lower semicontinuous at every point of X.

Definition 4.3. If F is both upper and lower semicontinuous, it is said to be *continuous* (sometimes it is also called *Hausdorff continuous*).

A continuous set-valued mapping F with the property that, for each x, its value $F(x)$ is a closed bounded set, is continuous with respect to the so-called *Hausdorff metric* on the space of all non-empty closed bounded subsets in Y. In order to describe the Hausdorff metric we first introduce the submetric $\bar{H}(A, B) = \sup\limits_{a \in A} \rho(a, B)$ where ρ is the metric in Y. Then the Hausdorff metric is defined by the formula

$$H(A, B) = \max(\bar{H}(A, B), \bar{H}(B, A)). \tag{4.1}$$

A set-valued mapping is said to be *closed* if its graph is a closed subset in $X \times Y$. If F is closed and for each point $x \in X$ there exists a neighborhood $U(x)$ such that $F(U(x))$ is relatively compact, F is upper semicontinuous.

Definition 4.4. We say that $F(t, x)$ satisfies the *upper Carathéodory condition* if:

1) for every $x \in X$ the map $F(\cdot, x) : I \multimap Y$ is measurable;
2) for almost all $t \in I$ the map $F(t, \cdot) : X \multimap Y$ is upper semicontinuous.

Definition 4.5. Let $I = [0, l] \subset R$. The set-valued mapping $F : I \times X \multimap Y$ is said to be *almost lower semicontinuous* if there exists a countable sequence of disjoint compact sets $\{I_n\}$, $I_n \subset I$ such that:

(i) the measure of $I \setminus \cup_n I_n$ is equal to zero;
(ii) the restriction of F on each $I_n \times X$ is lower semicontinuous.

An important technical role in the investigation of set-valued mappings is played by single-valued mappings that approximate the set-valued mappings in some sense. We describe two kinds of such single-valued mappings: selectors and ε-approximations.

Definition 4.6. Let $F : X \multimap Y$ be a set-valued mapping. A single-valued mapping $f : X \to Y$ such that for each $x \in X$ the inclusion $f(x) \in F(x)$ holds is called a *selector* of F.

Not every set-valued mapping has a continuous selector. However, for lower semicontinuous set-valued mappings with convex closed values, their existence is proved in the following classical Theorem.

Theorem 4.7 (Michael's Theorem) *If X is an arbitrary metric space and Y is a Banach space, then a lower semicontinuous mapping such that the value of every point of X is a convex closed set has a continuous selector.*

If the values of a lower semicontinuous set-valued mapping are not, in general, convex, it may not have continuous selectors. In this case the following construction is often very useful.

Definition 4.8. Let E be a separable Banach space. A non-empty set $\mathcal{M} \subset L^1([0,l];E)$ is called *decomposable* if $f \cdot \chi_{\mathfrak{M}} + g \cdot \chi_{[0,l]\setminus\mathfrak{M}} \in \mathcal{M}$ for all $f, g \in \mathcal{M}$ and for every measurable subset $\mathfrak{M}$ in $[0,l]$ where χ is the characteristic function of the corresponding set.

The reader can find more details about decomposable sets in [48] and [155].

Theorem 4.9 (Bressan-Colombo Theorem) *Let (Ω, d) be a separable metric space, X be a Banach space and $(J, \mathcal{A}, \mu)$ be a measurable space with a σ-algebra $\mathcal{A}$ and a non-atomic measure μ such that $\mu(J) = 1$. Consider the space $Y = L^1_X(J, \mathcal{A}, \mu)$ of integrable mappings from $(J, \mathcal{A}, \mu)$ into X. If a set-valued mapping $F : \Omega \multimap Y$ is lower semicontinuous and has closed decomposable values, F has a continuous selector.*

The assertion of Theorem 4.9 is proved, for example, as Lemma 9.2 in [48].

Upper semicontinuous mappings arise in applications more often than lower semicontinuous mappings. Generally speaking, they do not have continuous selectors (but they do have measurable selectors). The so-called ε-approximations are very useful for investigating upper semicontinuous mappings.

Definition 4.10. For a given $\varepsilon > 0$ a continuous single-valued mapping $f_\varepsilon : X \to Y$ is called an *ε-approximation* of a set-valued mapping $F : X \multimap Y$ if the graph of f, as a set in $X \times Y$, belongs to the ε-neighborhood of the graph of F.

We mention the following classes of upper semicontinuous set-valued mappings of finite-dimensional spaces, for which the existence of ε-approximations is proved for each $\varepsilon > 0$:

(i) the mappings with convex closed values;
(ii) the mappings with values that are aspheric in all dimensions from 1 to $n-1$ and weakly aspheric in the dimension n (see [32]). This class of set-valued mappings was first considered by A.D. Myshkis in 1954 [184]. In [32] and [87] topological characteristics of topological index and Lefschetz number types were constructed for such mappings. Later (in the 1980s) this class was rediscovered and described as "the mappings whose values at every point have the so-called uv^k-property for $k = 1, \ldots, n$" (for the exact definition see, e.g., [166]).

Let X be a Banach space and $F : X \multimap X$ be an upper semicontinuous set-valued mapping with convex closed values. For each bounded subset $\Omega \subset X$, let the image $F(\Omega)$ be relatively compact. Then if F sends a ball B of X

into itself, in B there exists a fixed point $x \in F(x)$ of F (this is an analog of Schauder's principle known as the *Glicksberg-Ky Fan Theorem*).

Let $F : \mathbb{R} \times \mathbb{R}^n \multimap \mathbb{R}^n$ be a set-valued mapping. A *differential inclusion*

$$\dot{x} \in F(t, x) \tag{4.2}$$

is an analog of a differential equation and transforms into the latter if F is single-valued.

A solution of (4.2) is an absolutely continuous curve $x(t)$ such that (4.2) is satisfied for $x(t)$ almost everywhere.

If F is upper semicontinuous and has convex closed bounded values, for each pair $x_0 \in \mathbb{R}^n$, $t_0 \in \mathbb{R}$ there exists a local in time solution of (4.2) with the initial condition $x(t_0) = x_0$. It is also known that for an upper semicontinuous F with closed bounded (not necessarily convex) values there exists a solution of the Cauchy problem for the differential inclusion

$$\dot{x} \in \overline{\mathrm{co}} F(t, x),$$

where $\overline{\mathrm{co}} F(t, x)$ is the convex closure of $F(t, x)$.

The existence of solutions of (4.2) for lower semicontinuous F is possible also for non-convex values. Often such existence can be proved by applying the Bressan-Colombo Theorem (Theorem 4.9).

4.2 Special Approximations

Here, following [10, 11], we prove the existence of special approximations for upper semicontinuous mappings in finite-dimensional spaces with convex closed values which point-wise converge to a Borel measurable selector of the set-valued mapping as $\varepsilon \to 0$. These results will often be used below.

Theorem 4.11 *Let $\Phi : \mathbb{R}^n \multimap \mathbb{R}^n$ be an upper semi-continuous set-valued mapping with convex closed bounded values. For a sequence $\varepsilon_i \to 0$ there exists a sequence of continuous ε_i-approximations for Φ that point-wise converges to a Borel measurable selector of Φ. If Φ takes values in a convex set Ξ in $\mathbb{R}^n$, these ε-approximations also take values in Ξ.*

Proof. It is shown in [79, Theorem 2] that, in the case under consideration, for any ε_i there exists a lower semi-continuous set-valued map $\Psi_i : \mathbb{R}^n \multimap \mathbb{R}^n$ with closed convex bounded values such that: (i) for any $x \in \mathbb{R}^n$ the inclusion $\Phi(x) \subset \Psi_i(x)$ holds and (ii) the graph of Ψ_i belongs to the ε_i-neighborhood of the graph of Φ. From the construction it follows that if Φ takes values in a convex set Ξ in $\mathbb{R}^n$, then the values of all $\Psi_i(x)$ belong to Ξ. Notice that for an upper semi-continuous mapping with compact values the sum of such mappings and the products with a continuous function are upper semi-continuous. Hence, from the proof of Theorem 2 of [79] it follows that, in the case under consideration, all Ψ_i are continuous set-valued mapping and, in particular in our case, they are continuous with respect to the Hausdorff metric.

Consider the minimal selector $\psi_i(\cdot)$ of $\Psi_i(\cdot)$, i.e., $\psi_i(x)$ is the closest point to the origin in $\Psi_i(x)$, $x \in \mathbb{R}^n$. We refer the reader to [5] for a complete description of minimal selectors. In particular, it is shown there that minimal selectors are continuous. Thus, ψ_i is an ε_i-approximation of Φ.

Let $x \in \mathbb{R}^n$. Since $\Phi(x) \subset \Psi_i(x)$ for each i, for the Hausdorff submetric $\bar{H}$ we have $\bar{H}(\Phi(x), \Psi_k(x)) = 0$. Hence for the Hausdorff metric H we obtain that $H(\Psi_i(x), \Phi(x)) = \bar{H}(\Psi_i(x), \Phi(x))$ for each i.

Now specify ε_k. By the definition of upper semi-continuity, for any $x \in \mathbb{R}^n$ there exists a $\delta_k > 0$ such that for any x' in the δ_k-neighborhood of x the value $\Phi(x')$ belongs to the ε_k-neighborhood of $\Phi(x)$. Since $\varepsilon_i \to 0$, $\varepsilon_{k+l} < \delta_k$ for some $l = l(k, x)$ and without loss of generality we may take $l(k, x) \geq 0$. Thus $\bar{H}(\Phi(x'), \Phi(x)) < \varepsilon_k$ for each x' in the ε_{k+l}-neighborhood of x.

Since the graph of Ψ_{k+l} belongs to the ε_{k+l}-neighborhood of the graph of Φ, there exists a point x'' in the ε_{k+l}-neighborhood of x such that $\Psi_{k+l}(x)$ belongs to the ε_{k+l}-neighborhood of $\Phi(x'')$, i.e., $\bar{H}(\Psi_{k+l}(x), \Phi(x'')) < \varepsilon_{k+l}$.

Thus

$$\begin{aligned} H(\Psi_{k+l}(x), \Phi(x)) &= \bar{H}(\Psi_{k+l}(x), \Phi(x)) \\ &\leq \bar{H}(\Psi_{k+l}(x), \Phi(x'')) + \bar{H}(\Phi(x''), \Phi(x)) \\ &< \varepsilon_{k+l} + \varepsilon_k < 2\varepsilon_k. \end{aligned}$$

Hence at each x the convex set $\Psi_i(x)$ tends to the convex set $\Phi(x)$ with respect to the Hausdorff metric as $i \to \infty$. It follows that $\psi_i(x)$ tends to the point $\varphi(x) \in \Phi(x)$ that is the closest to the origin. The well-known fact that the point-wise limit $\varphi(\cdot)$ of the sequence of continuous mappings $\psi_i(\cdot)$ is a Borel measurable mapping completes the proof. □

We introduce $\tilde{\Omega} = C^0([0, T], \mathbb{R}^n)$ – the Banach space of continuous curves in $\mathbb{R}^n$ given on $[0, T]$, with the usual uniform norm – and the σ-algebra $\tilde{\mathcal{F}}$ on $\tilde{\Omega}$ generated by cylinder sets. By $\mathcal{P}_t$ we denote the σ-subalgebra of $\mathcal{F}$ generated by cylinder sets with bases over $[0, t] \subset [0, T]$. Recall that $\tilde{\mathcal{F}}$ is the Borel σ-algebra on $\tilde{\Omega}$ (see [208]).

Let $B : [0, T] \times \tilde{\Omega} \to Z$ be a mapping to some metric space Z. Below we shall often suppose that such mappings, for various spaces Z, satisfy the following condition:

Condition 4.12 *For each $t \in [0, T]$, from the fact that the curves $x_1(\cdot)$ and $x_2(\cdot) \in \tilde{\Omega}$ coincide for $0 \leq s \leq t$, it follows that $B(t, x_1(\cdot)) = B(t, x_2(\cdot))$.*

Remark 4.13. The fact that a mapping B satisfies Condition 4.12 is equivalent to the fact that B is measurable at each t with respect to a Borel σ-algebra in Z and $\mathcal{P}_t$ in $\tilde{\Omega}$ (see [83], cf. Condition 6.19(ii) below).

Theorem 4.14 *Let (ε_k) be a sequence of positive numbers such that $\varepsilon_k \to 0$ as $k \to \infty$. Let $\boldsymbol{B}$ be an upper semi-continuous set-valued mapping with compact convex values sending $[0,T] \times \tilde{\Omega}$ to a finite-dimensional Euclidean space Y and satisfying Condition 4.12. Then there exists a sequence of continuous single-valued mappings $B_k : [0,T] \times \tilde{\Omega} \to Y$ with the following properties:*

(i) *each B_k satisfies Condition 4.12;*
(ii) *the sequence B_k point-wise converges to a selector of $\boldsymbol{B}$ that is measurable with respect to the Borel σ-algebra in Y and the product σ-algebra of the Borel σ-algebra on $[0,T]$ and $\tilde{\mathcal{F}}$ on $\tilde{\Omega}$;*
(iii) *at each $(t, x(\cdot)) \in [0,T] \times \tilde{\Omega}$ the inequality $\|B_k(t, x(\cdot))\| \le \|\boldsymbol{B}(t, x(\cdot))\|$ holds for all k;*
(iv) *if $\boldsymbol{B}$ takes values in a closed convex set $\Xi \subset Y$, the values of all B_k belong to Ξ.*

Proof. In this proof we combine and modify the ideas used in the proofs of [79, Theorem 2] by Gel'man and Theorem 4.11 above.

For $t \in [0,T]$ define the mapping $f_t : \tilde{\Omega} \to \tilde{\Omega}$ by the formula

$$f_t x(\cdot) = \begin{cases} x(s) & \text{if} \quad 0 \le s \le t \\ x(t) & \text{if} \quad t \le s \le T. \end{cases} \tag{4.3}$$

Clearly $f_t x(\cdot)$ is jointly continuous in $t \in [0,T]$ and $x(\cdot) \in \tilde{\Omega}$. Since $\boldsymbol{B}$ satisfies Condition 4.12, $\boldsymbol{B}(t, x(\cdot)) = \boldsymbol{B}(t, f_t x(\cdot))$ for each $x(\cdot) \in \tilde{\Omega}$ and $t \in [0,T]$.

Choose an element ε_k from the sequence. Since $\boldsymbol{B}$ is upper semi-continuous, for every $(t, x(\cdot)) \in [0,T] \times \tilde{\Omega}$ there exists a $\delta_k(t,x) > 0$ such that for every $(t^*, x^*(\cdot))$ from the $\delta_k(t,x)$-neighborhood of $(t, x(\cdot))$ the set $\boldsymbol{B}(t^*, x^*(\cdot))$ is contained in the $\frac{\varepsilon_k}{2}$-neighborhood of the set $\boldsymbol{B}(t, x(\cdot))$. Without loss of generality we can suppose $0 < \delta_k(t,x) < \varepsilon_k$ for every $(t, x(\cdot))$. Consider the $\frac{\delta_k(t,x)}{4}$-neighborhood of $(t, x(\cdot))$ in $[0,T] \times \tilde{\Omega}$ and construct the open covering of $[0,T] \times \tilde{\Omega}$ by such neighborhoods for all $(t, x(\cdot))$. Since $[0,T] \times \tilde{\Omega}$ is paracompact, there exists a locally finite refinement $\{V_j^k\}$ of this covering. Without loss of generality we can consider each V_j^k as an $\eta_k(t_j^k, x_j^k)$-neighborhood of some $(t_j^k, x_j^k(\cdot))$ where by construction the radius $\eta_k(t_j, x_j) \le \frac{\delta_k(t_j, x_j)}{4}$.

Consider a continuous partition of unity $\{\varphi_j^k\}$ adapted to $\{V_j^k\}$ and introduce the set-valued mapping $\Phi_k(t, x(\cdot)) = \sum\limits_j \varphi_j^k(t, x(\cdot)) \overline{\text{co}} \boldsymbol{B}(V_j^k)$ where $\overline{\text{co}}$ denotes the convex closure. Since $\boldsymbol{B}(t, x(\cdot))$ is upper semi-continuous and has compact values, without loss of generality we can suppose $\delta_k(t,x)$ to be such that the images $\boldsymbol{B}(V_j^k)$ are bounded in Y and so the sets $\overline{\text{co}} \boldsymbol{B}(V_j^k)$ are compact. Denote by $\overline{\Phi}_k(t, x(\cdot))$ the closure of $\Phi_k(t, x(\cdot))$. Then one can easily see that $\overline{\Phi}_k : [0,T] \times \tilde{\Omega} \to Y$ is a Hausdorff continuous set-valued mapping with compact convex values.

Define $\Psi_k : [0,T] \times \tilde{\Omega} \to Y$ by the formula $\Psi_k(t, x(\cdot)) = \Phi_k(t, f_t x(\cdot))$ and consider the set-valued mapping $\overline{\Psi}_k(t, x(\cdot))$. Since f_t is continuous, every $\overline{\Psi}_k$

is a Hausdorff continuous set-valued mapping with compact convex values and by construction it satisfies Condition 4.12.

The pair $(t, f_t x(\cdot))$ belongs to a finite collection of neighborhoods $V_{j_i}^k$ with centers at $(t_{j_i}^k, x_{j_i}^k(\cdot))$, $i = 1, \ldots, n$, and so by construction $\boldsymbol{B}(t, x(\cdot)) = \boldsymbol{B}(t, f_t x(\cdot)) \subset \boldsymbol{B}(V_{j_i}^k)$ for each i. Hence $\boldsymbol{B}(t, x(\cdot)) = \boldsymbol{B}(t, f_t x(\cdot)) \subset \Psi_k(t, x(\cdot))$ for every pair $(t, x(\cdot))$.

Let l be the index from the collection of indices j_i above such that $\eta_k(t_l^k, x_l^k)$ takes the greatest value among $\eta_k(t_{j_i}^k, x_{j_i}^k)$. Then all $(t_{j_i}^k, x_{j_i}^k(\cdot))$ are contained in the $2\eta_k(t_l^k, x_l^k)$-neighborhood of $(t_l^k, x_l^k(\cdot))$ and so every $V_{j_i}^k$ is contained in the $3\eta_k(t_l^k, x_l^k)$-neighborhood of $(t_l^k, x_l^k(\cdot))$ that is contained in the $\delta_k(t_l^k, x_l^k(\cdot))$-neighborhood of (t_l^k, x_l^k) by construction. Then, also by construction, $\Psi_k(t, x(\cdot))$ belongs to the $\frac{\varepsilon_k}{2}$-neighborhood of $\boldsymbol{B}(t_l^k, x_l^k(\cdot))$. Since both $\Psi_k(t, x(\cdot))$ and $\boldsymbol{B}(t_l^k, x_l^k(\cdot))$ are convex, this means that $\overline{\Psi}_k(t, x(\cdot))$ also belongs to the $\frac{\varepsilon_k}{2}$-neighborhood of $\boldsymbol{B}(t_l^k, x_l^k(\cdot))$. Notice that this is true for each k.

Since $\boldsymbol{B}(t, x(\cdot)) \subset \Psi_k(t, x(\cdot)) \subset \overline{\Psi}_k(t, x(\cdot))$, for the Hausdorff submetric $\bar{H}$ we have

$$\bar{H}\left(\boldsymbol{B}(t, x(\cdot)), \overline{\Psi}_k(t, x(\cdot))\right) = 0.$$

Hence for the Hausdorff metric H we obtain that

$$H\left(\overline{\Psi}_k(t, x(\cdot)), \boldsymbol{B}(t, x(\cdot))\right) = \bar{H}\left(\overline{\Psi}_k(t, x(\cdot)), \boldsymbol{B}(t, x(\cdot))\right).$$

Since $\varepsilon_k \to 0$, for $(t, x(\cdot))$ there exists an integer $\theta = \theta(t, x(\cdot)) > 0$ such that $\varepsilon_{k+\theta} < \delta_k(t, x(\cdot))$. Without loss of generality we can suppose that $\theta \geq 1$.

Thus $\boldsymbol{B}(t_l^{k+\theta}, x_l^{k+\theta}(\cdot))$ belongs to the $\frac{\varepsilon_k}{2}$-neighborhood of $\boldsymbol{B}(t, x(\cdot))$ and so

$$\bar{H}\left(\boldsymbol{B}\left(t_l^{k+\theta}, x_l^{k+\theta}(\cdot)\right), \boldsymbol{B}(t, x(\cdot))\right) < \frac{\varepsilon_k}{2}.$$

Since $\overline{\Psi}_{k+\theta}(t, x(\cdot))$ belongs to the $\frac{\varepsilon_{k+\theta}}{2}$-neighborhood of $\boldsymbol{B}(t_l^{k+\theta}, x_l^{k+\theta}(\cdot))$ (see above), we obtain that

$$\bar{H}\left(\overline{\Psi}_{k+\theta}(t, x(\cdot)), \boldsymbol{B}\left(t_l^{k+\theta}, x_l^{k+\theta}(\cdot)\right)\right) < \frac{\varepsilon_{k+\theta}}{2}.$$

Thus

$$\begin{aligned} H\left(\overline{\Psi}_{k+\theta}(t, x(\cdot)), \boldsymbol{B}(t, x(\cdot))\right) &= \bar{H}\left(\overline{\Psi}_{k+\theta}(t, x(\cdot)), \boldsymbol{B}(t, x(\cdot))\right) \\ &\leq \bar{H}\left(\overline{\Psi}_{k+\theta}(t, x(\cdot)), \boldsymbol{B}\left(t_l^{k+\theta}, x_l^{k+\theta}(\cdot)\right)\right) \\ &\quad + \bar{H}\left(\boldsymbol{B}\left(t_l^{k+\theta}, x_l^{k+\theta}(\cdot)\right), \boldsymbol{B}(t, x(\cdot))\right) \\ &< \frac{\varepsilon_{k+\theta}}{2} + \frac{\varepsilon_k}{2} < \varepsilon_k. \end{aligned}$$

So, at each $(t, x(\cdot))$ we have that $H(\overline{\Psi}_k(t, x(\cdot)), \boldsymbol{B}(t, x(\cdot))) \to 0$ as $k \to \infty$ and $\boldsymbol{B}(t, x(\cdot)) \subset \Psi_k(t, x(\cdot))$ for all k.

Consider the minimal selector $B_k(t, x(\cdot))$ of $\overline{\Psi}_k(t, x(\cdot))$, i.e., $B_k(t, x(\cdot))$ is the closest point to the origin in $\overline{\Psi}_i(t, x(\cdot))$. We again refer the reader to [5] for a complete description of minimal selectors and, in particular, recall that the minimal selectors we shall be considering are all continuous. One can easily see that all B_k satisfy Condition 4.12.

By construction the minimal selectors $B_k(t, x(\cdot))$ of $\overline{\Psi}_k(t, x(\cdot))$ point-wise converge to the minimal selector $B(t, x(\cdot))$ of $\boldsymbol{B}(t, x(\cdot))$ as $k \to \infty$ since at any $(t, x(\cdot))$ we have that $H(\overline{\Psi}_k(t, x(\cdot)), \boldsymbol{B}(t, x(\cdot))) \to 0$ as $k \to \infty$ and $\boldsymbol{B}(t, x(\cdot)) \subset \Psi_k(t, x(\cdot))$ for all k (see above). It is a well-known fact that the point-wise limit B of the sequence of continuous mappings B_k is measurable with respect to the Borel σ-algebras in Y and in $[0, T] \times \tilde{\Omega}$ (see [194]). The latter coincides with the product σ-algebra of the Borel σ-algebra on $[0, T]$ and $\tilde{\mathcal{F}}$ on $\tilde{\Omega}$ (see [208]). Properties (iii) and (iv) immediately follow from the construction. $\square$

Remark 4.15. Unlike $\bar{\Psi}_k(t, x(\cdot))$, the set-valued mapping $\bar{\Phi}_k(t, x(\cdot))$ may not satisfy Condition 4.12 since two different curves $x_1(\cdot)$ and $x_2(\cdot)$ coinciding on $[0, t]$ may have different neighborhoods V_j^k to which they belong, and so the values $\bar{\Phi}_k(t, x_1(\cdot))$ and $\bar{\Phi}_k(t, x_2(\cdot))$ may be different. On the other hand, it follows from [79] that $\bar{\Phi}_k$ is an ε_k-approximation of $\boldsymbol{B}$ while the same is not true for $\bar{\Psi}_k$.

Chapter 5
Analysis on Groups of Diffeomorphisms

5.1 General Concepts

By H^s we denote the Sobolev space of functions such that the functions and their generalized derivatives up to order s belong to the functional space L^2. A detailed description of Sobolev spaces can be found, e.g., in [62]. An introduction to the manifold structure in functional sets can be found in [64]. The reader may wish to consult [61] for details on the remaining material of this section.

The case of a compact manifold without boundary

Let M, $\dim M = n$, be a compact oriented manifold without boundary equipped with a Riemannian metric $\langle\cdot,\cdot\rangle$ and its Levi-Civitá connection H.

Denote by $H^s = H^s(M,M)$ the set of Sobolev H^s-mappings from M to M with $s > \frac{1}{2}n + 1$. Recall that for $s > \frac{1}{2}n + k$ the maps from H^s are C^k-smooth. Sometimes we shall also deal with sets $H^s(M,N)$ where N is a manifold with the same dimension as M. There is an infinite-dimensional manifold structure on $H^s(M,M)$ and $H^s(M,N)$ (see [61]).

Consider the open neighborhood $\mathcal{D}^s(M)$ of the identical mapping in $H^s(M,M)$ that consists of all H^s-diffeomorphisms. Consider also its subset $\mathcal{D}^s_\mu(M)$ comprising the H^s-diffeomorphisms which preserve the Riemannian volume form.

$\mathcal{D}^s(M)$ and $\mathcal{D}^s_\mu(M)$ have the structures of smooth (and separable) Hilbert manifolds as well as the natural multiplicative group structures (with respect to composition). A detailed description of these structures and their interconnections can be found in [61]. The tangent space $T_e\mathcal{D}^s(M)$ at the unit $e = \mathrm{id}$ is the space of all vector fields on M belonging to H^s, and $T_e\mathcal{D}^s_\mu(M)$ is the space of all divergent-free vector fields on M belonging to H^s. The tangent space $T_g\mathcal{D}^s_\mu(M)$, $g \in \mathcal{D}^s_\mu(M)$, consists of the compositions of the fields from $T_e\mathcal{D}^s_\mu(M)$ with g, i.e., $T_g\mathcal{D}^s_\mu(M) = \{X \circ g | X \in T_e\mathcal{D}^s_\mu(M)\}$. This

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_5,

means that $Y \in T_g\mathcal{D}^s_\mu(M)$ is a map $Y : M \to TM$ such that $\pi Y(m) = g(m)$ where $\pi : TM \to M$ is the natural projection. For $X \in T_e\mathcal{D}^s_\mu(M)$ we have $\pi X(m) = m$. $\mathcal{D}^s(M)$ also possesses the same features, i.e., $T_g\mathcal{D}^s(M) = \{X \in H^s(M, TM) | \pi \circ X = g\}$, etc.

Let N be a compact oriented Riemannian manifold without boundary. Consider the mappings $\alpha_g : H^s(M,N) \to H^s(M,N)$ of the form $\alpha_g(f) = f \circ g$, and $\omega_h : \mathcal{D}^s(M) \to H^s(M,N)$ of the form $\omega_h(r) = h \circ r$.

Lemma 5.1 (α-lemma) *α_g is C^∞-smooth and its derivative is also of the form α_g.*

Lemma 5.2 (ω-lemma) *ω_h is continuous. If $h \in H^{s+k}$, $\omega_h : \mathcal{D}^s \to H^s(M,N)$ is a C^k-mapping with derivative of the form ω_{Th}. In particular, if $h \in C^\infty$, ω_h is C^∞-smooth.*

The right translation $R_f : \mathcal{D}^s_\mu(M) \to \mathcal{D}^s_\mu(M)$, $R_f \circ \theta = \theta \circ f$, where $\theta, f \in \mathcal{D}^s_\mu(M)$, is C^∞-smooth and thus one may consider right-invariant vector fields on $\mathcal{D}^s_\mu(M)$. The tangent to right translation takes the form $TR_f X = X \circ f$ for $X \in T\mathcal{D}^s_\mu(M)$. For $\mathcal{D}^s(M)$ we have analogous properties.

Convention 5.3 *We shall consider $TR_g : T_\eta\mathcal{D}^s(M) \to T_{\eta\circ g}\mathcal{D}^s(M)$ for all $\eta, g \in \mathcal{D}^s(M)$ as a right action of $\mathcal{D}^s(M)$ on $T\mathcal{D}^s(M)$.*

Theorem 5.4 *Let $X \in T_e\mathcal{D}^s(M)$ be a vector field on M and $\bar{X}$ be the corresponding right-invariant vector field on $\mathcal{D}^s(M)$, $\bar{X}_g = X \circ g$. The vector field $\bar{X}$ on $\mathcal{D}^s(M)$ is C^k-smooth if and only if the vector field X on M belongs to the class H^{s+k}. In particular, $\bar{X}$ is C^∞-smooth if and only if X is C^∞-smooth. The same property holds for right-invariant vector fields on $\mathcal{D}^s_\mu(M)$.*

This fact is a consequence of the ω-lemma 5.2 and is also true for more complicated fields (for example, for right-invariant tensor fields).

The left translation $L_g f = g \circ f$ is continuous. If $g \in \mathcal{D}^{s+k}(M)$, L_g is a C^k-mapping. In particular, in this case $TL_g(X) = Tg \circ X$ where $Tg : TM \to TM$ is the tangent mapping of g and $X \in T\mathcal{D}^s(M)$. The mapping $g \mapsto g^{-1}$ is continuous on $\mathcal{D}^s(M)$. If $g \in \mathcal{D}^{s+k}(M)$, this is a C^k-mapping from $\mathcal{D}^{s+k}(M)$ to $\mathcal{D}^s(M)$. Thus, from the point of view of the standard finite-dimensional definition, $\mathcal{D}^s(M)$ is not a Lie group.

Theorem 5.5 *Let $s > \frac{n}{2} + 1$ and let the right-invariant vector field $\bar{X}$ on $\mathcal{D}^s(M)$ be C^1-smooth. Then:*

(i) *for every $g \in \mathcal{D}^s(M)$ there exists a unique integral curve $\gamma_g(t)$ of this field, well-defined for all $t \in (-\infty, \infty)$, such that $\gamma_g(0) = g$;*
(ii) *$\gamma_e(t)$ is the flow of the vector field $X = \bar{X}_e$ on M, $\gamma_g(t) = \gamma_e(t) \circ g$;*
(iii) *if $s > \frac{n}{2} + 2$, assertion* (i) *is valid for a continuous right-invariant vector field $\bar{X}$ on $\mathcal{D}^s(M)$.*

The same results are true for right-invariant vector fields of $\mathcal{D}^s_\mu(M)$.

For $g \in \mathcal{D}^s(M)$ consider the tangent space $T_g\mathcal{D}^s(M)$ (see above). Define an inner product $(\cdot,\cdot)$ in $T_g\mathcal{D}^s(M)$ by the formula

$$(X,Y)_g = \int_M \langle X(m), Y(m)\rangle_{g(m)}\mu(\mathrm{d}m), \tag{5.1}$$

where X, $Y \in T_g\mathcal{D}^s(M)$ and μ is the Riemannian volume form. Recall that $\pi X(m) = \pi Y(m) = g(m)$ so that the vectors $X(m)$ and $Y(m)$ belong to the tangent space at $g(m)$ and so the product is found with respect to $\langle\cdot,\cdot\rangle_{g(m)}$. Since the metric tensor $\langle\cdot,\cdot\rangle$ is C^∞-smooth, from the ω-lemma 5.2 it follows that (5.1) is C^∞-smooth in $g \in \mathcal{D}^s(M)$.

Clearly this metric introduces the topology of the functional space $L^2 = H^0$ in the tangent spaces, which is weaker than the initial topology on H^s. This is why $(\cdot,\cdot)$ is called a *weak Riemannian metric*.

One can easily check that the second tangent bundle $TT\mathcal{D}^s(M)$ consists of H^s maps from M to TTM with the additional properties that they are projected into maps from $\mathcal{D}^s(M)$. Consider the connector $K : TTM \to TM$ of the Levi-Civitá connection H on M.

Define the mapping $\bar{K} : TT\mathcal{D}^s(M) \to T\mathcal{D}^s(M)$ by the equality

$$\bar{K}(Y) = K \circ Y. \tag{5.2}$$

Theorem 5.6 *$\bar{K}$ is invariant with respect to right shifts on $\mathcal{D}^s(M)$.*

$\bar{K}$ is the connector of a connection $\bar{\mathsf{H}}$ that is proved to be the Levi-Civitá connection of the metric (5.1).

For vector fields X, Y on $\mathcal{D}^s(M)$ and for a vector field $X(t)$ along a certain smooth curve $g(t)$ in $\mathcal{D}^s(M)$ define the covariant derivatives $\bar{\nabla}_X Y$ and $\frac{\bar{\mathrm{D}}}{\mathrm{d}t}X(t)$, respectively, by the usual formulae (cf. Definitions 2.22 and 2.25)

$$\begin{aligned} \bar{\nabla}_X Y &= \bar{K} \circ TY(X) = K \circ TY(X), \\ \frac{\bar{\mathrm{D}}}{\mathrm{d}t}X(t) &= \bar{K} \circ \frac{\mathrm{d}}{\mathrm{d}t}X(t) = K \circ \frac{\mathrm{d}}{\mathrm{d}t}X(t). \end{aligned} \tag{5.3}$$

On $\mathcal{D}^s(M)$ with connection $\bar{\mathsf{H}}$ a vector field X along a curve $g(t)$ is parallel if $\frac{\bar{\mathrm{D}}}{\mathrm{d}t}X(t) = 0$. The curve $g(t)$ is geodesic if it satisfies the equation $\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\dot{g}(t) = 0$, where $\dot{g}(t) = \frac{\mathrm{d}}{\mathrm{d}t}g(t)$.

The geodesic spray $\bar{\mathcal{Z}}$ of $\bar{\mathsf{H}}$ is described as follows:

$$\bar{\mathcal{Z}}(X) = \mathcal{Z} \circ X \tag{5.4}$$

for $X \in T\mathcal{D}^s(M)$, where $\mathcal{Z}$ is the geodesic spray of the connection H on M. Since $\mathcal{Z}$ is C^∞-smooth, from the ω-lemma 5.2 it follows that $\bar{\mathcal{Z}}$ is C^∞-smooth on $T\mathcal{D}^s(M)$. From (5.4) it evidently follows that $\bar{\mathcal{Z}}$ is $\mathcal{D}^s(M)$-right-invariant.

Proposition 5.7 *If $g(t)$ is a geodesic of the connection $\bar{\mathsf{H}}$ on $\mathcal{D}^s(M)$, then for every $f \in \mathcal{D}^s(M)$ the curve $R_f g(t)$ is also a geodesic.*

This follows from the fact that the geodesic spray $\bar{\mathcal{Z}}$ is right-invariant.

Remark 5.8. The exponential mapping $\overline{\exp}$ of the Levi-Civitá connection on $\mathcal{D}^s(M)$ is well-defined. It is a C^∞-mapping of a neighborhood of the zero cross-section in $T\mathcal{D}^s(M)$ to $\mathcal{D}^s$. This follows from the existence, uniqueness and smooth dependence on initial values of the local solution of the Cauchy problem for integral curves of C^∞-smooth vector fields. The fact that $\overline{\exp}$ sends a neighborhood in $T_e\mathcal{D}^s(M)$ onto a neighborhood of e in $\mathcal{D}^s$ follows from general properties of smooth exponential mappings.

In what follows the tangent space $T_e\mathcal{D}^s(M)$ will be called the *algebra* of the group $\mathcal{D}^s(M)$ in analogy with finite-dimensional Lie groups. In fact the everywhere dense linear submanifold of C^∞-vector fields in $T_e\mathcal{D}^s(M)$ really is an infinite-dimensional Lie algebra, on which the bracket coincides with the ordinary Lie bracket of vector fields. However, the entire tangent space $T_e\mathcal{D}^s(M)$ is not an infinite-dimensional Lie algebra since the Lie bracket can give values which are not in $T_e\mathcal{D}^s(M)$.

The restriction of (5.1) to $T\mathcal{D}^s_\mu(M)$ is evidently right-invariant. It is a weak Riemannian metric on $\mathcal{D}^s_\mu(M)$.

Recall the Hodge decomposition for M [61]

$$H^s(TM) = G^s \oplus E^s \oplus \ker \Delta = G^s \oplus T_e\mathcal{D}^s_\mu(M) \tag{5.5}$$

where G^s is the space of gradients of all H^{s+1} functions on M, E^s is the space of all H^s-co-gradients on M, $\ker \Delta$ is the space of all harmonic (i.e., both gradient and co-gradient) vector fields on M and $\oplus$ denotes the orthogonal direct sum with respect to the L^2-inner product (5.1) in $T_e\mathcal{D}^s(M)$. By a *co-gradient* we mean a vector field corresponding to a co-exact form on M with respect to the Riemannian metric $\langle\cdot,\cdot\rangle$. $\ker \Delta$ is a finite-dimensional space and consists of C^∞-smooth vector fields.

Denote by $P_e : T_e\mathcal{D}^s(M) = H^s(TM) \to E^s \oplus \ker \Delta = T_e\mathcal{D}^s_\mu(M)$ the $(\cdot,\cdot)_e$-orthogonal projection in (5.5). Consider the mapping $\bar{P} : T\mathcal{D}^s(M)_{|\mathcal{D}^s_\mu(M)} \to T\mathcal{D}^s_\mu(M)$ determined for each $\eta \in \mathcal{D}^s_\mu(M)$ by the formula

$$\bar{P}_\eta = TR_\eta \circ P_e \circ TR_\eta^{-1}. \tag{5.6}$$

It is obvious that $\bar{P}$ is $\mathcal{D}^s_\mu(M)$-right-invariant. There is an important and rather complicated result (see [61]) that $\bar{P}$ is C^∞-smooth. Notice the consequence of (5.5) and of the definition of P_e: the relation

$$P_e(Y) = Y - \mathrm{grad} p \tag{5.7}$$

holds for every $Y \in T_e\mathcal{D}^s(M)$ where p is an H^{s+1}-function on M, unique to within an additive constant.

Since $\mathcal{D}^s_\mu(M)$ is a submanifold in $\mathcal{D}^s(M)$, according to the standard constructions of differential geometry there is a corresponding connection $\tilde{\mathsf{H}}$ on $\mathcal{D}^s_\mu(M)$ whose connector $\tilde{K}$ and covariant derivatives $\tilde{\nabla}$ and $\frac{\tilde{\mathrm{D}}}{\mathrm{d}t}$ are described by the formulae

$$\begin{aligned}
\tilde{K} &= \bar{P} \circ K, \\
\tilde{\nabla}_X Y &= \bar{P} \circ \bar{\nabla}_X Y = \bar{P} \circ K \circ TY(X) = \tilde{K} \circ TY(X), \\
\frac{\tilde{\mathrm{D}}}{\mathrm{d}t} X(t) &= \bar{P} \circ \frac{\bar{\mathrm{D}}}{\mathrm{d}t} X(t) = \bar{P} \circ K \frac{\mathrm{d}}{\mathrm{d}t} X(t) = \tilde{K} \frac{\mathrm{d}}{\mathrm{d}t} X(t)
\end{aligned} \tag{5.8}$$

where X, Y are vector fields on $\mathcal{D}^s_\mu(M)$ and $X(t)$ is a vector field along a smooth curve $g(t)$ in $\mathcal{D}^s_\mu(M)$. So, a vector field X along a curve $g(t)$ in $\mathrm{D}^s_\mu(M)$ is parallel if $\frac{\tilde{\mathrm{D}}}{\mathrm{d}t} X(t) = 0$ and the curve $g(t)$ is geodesic if $\frac{\tilde{\mathrm{D}}}{\mathrm{d}t} \dot{g}(t) = 0$.

Theorem 5.9 *The geodesic spray $\mathcal{S}$ of the Levi-Civitá connection $\tilde{\mathsf{H}}$ of the metric (5.1) on $\mathcal{D}^s_\mu(M)$ is a C^∞-smooth right-invariant vector field on $T\mathcal{D}^s_\mu(M)$ of the form $\mathcal{S} = T\bar{P}(\bar{\mathcal{Z}})$ where $\bar{\mathcal{Z}}$ is the geodesic spray (5.4) on $T\mathcal{D}^s(M)$.*

Indeed, $\bar{P}$ and $\bar{\mathcal{Z}}$ are $\mathcal{D}^s_\mu(M)$-right-invariant and C^∞-smooth on $T\mathcal{D}^s_\mu(M)$, hence so is $\mathcal{S}$. Denote by $\widetilde{\exp}$ the corresponding exponential map of a neighborhood of the zero section in $T\mathcal{D}^s_\mu(M)$ onto $\mathcal{D}^s_\mu(M)$. Clearly the map $\widetilde{\exp}$ is C^∞-smooth and $\mathcal{D}^s_\mu(M)$-right-invariant.

Theorem 5.10 *There exists a neighborhood W of the unit e in $\mathcal{D}^s_\mu(M)$ that is covered by the image of $T_e\mathcal{D}^s_\mu(M)$ under the exponential mapping of the Levi-Civitá connection on $\mathcal{D}^s_\mu(M)$.*

This follows from the smoothness of $\mathcal{S}$.

The case of M with boundary

Let, as above, $s > \frac{n}{2} + 1$ and let M be a compact oriented manifold with boundary ∂M. Denote by $\mathcal{D}^s(M)$ the set of C^1-diffeomorphisms of M belonging to the Sobolev class H^s and by $\overset{\circ}{\mathcal{D}^s}(M)$ the set in $\mathcal{D}^s(M)$ consisting of diffeomorphisms coinciding with the identity on ∂M. In this case we cannot use $H^s(M, M)$ to introduce the smooth manifold structure on $\mathcal{D}^s(M)$ and on $\overset{\circ}{\mathcal{D}^s}(M)$ since $H^s(M, M)$ has infinite-dimensional corners.

Consider an arbitrary compact oriented manifold N without boundary that has the same dimension n as M and is such that M is embedded into N. We can take, say, the double of M as N (see Section 1.1) with Riemannian metric smoothly expanded over the boundary. Consider the Hilbert manifold $H^s(M, N)$.

Theorem 5.11 ([61]) *$\mathcal{D}^s(M)$ and $\overset{\circ}{\mathcal{D}^s}(M)$ are smooth sub-manifolds in $H^s(M, N)$. For $e = \mathrm{id} \in \mathcal{D}^s(M)$ the tangent space $T_e\mathcal{D}^s(M)$ is the space of H^s-vector fields on M tangent to the boundary ∂M, and $T_e \overset{\circ}{\mathcal{D}^s}(M)$ is the space of H^s-vector fields on M equal to zero on ∂M.*

The description of tangent bundles, group structures as well as of smooth properties of right and left shifts are completely analogous to those described above. For convenience of reference we summarize the properties of right-invariant vector fields in the following Remark.

Remark 5.12. If $X \in T_e\mathcal{D}^s(M)$ is an H^{s+k}-vector field on M, the corresponding right-invariant vector field $\bar{X}$ on $\mathcal{D}^s(M)$ is C^k-smooth. Nevertheless, generally speaking, the converse is not true. If $\bar{X}$ is C^k on $\mathcal{D}^s(M)$, X is H^{s+k} in the interior of M and in the directions tangent to the boundary ∂M, but it may not be H^{s+k}-smooth in the directions normal to the boundary.

On the manifold $H^s(M,N)$ we introduce a weakly Riemannian metric in the same way as (5.1). For this metric the analogs of the above-mentioned theorems are obviously fulfilled. It is important to mention that this metric can be considered at the points of the manifold $\mathcal{D}^s(M) \subset H^s(M,N)$. It is also possible to define a geodesic spray $\mathcal{Z} \circ X$ at $X \in T\mathcal{D}^s(m)$ but in this case the geodesics may not exist ($\overline{\exp}$ is not well-defined) since the boundary ∂M, generally speaking, is not a completely geodesic manifold in N.

Definition 5.13. We say that a k-form α on M is *tangent* (*normal*) to the boundary ∂M if the restriction to ∂M of the form $*\alpha$ (form α, respectively) is the identically zero form.

There are several versions of the Hodge decomposition on a manifold with boundary. We shall mainly deal with the following one (see [61]):

$$H^s(\wedge^k) = \mathrm{d}H^{s+1}(\wedge^{k-1}) \oplus \mathcal{E}^s(\wedge^k_t), \tag{5.9}$$

where $\oplus$ is the orthogonal direct sum with respect to the H^0-inner product $(\cdot,\cdot)_e$ (5.1) and $\mathcal{E}^s(\wedge^k_t)$ denotes the co-closed H^s-fields tangent to the boundary ∂M.

From (5.9) we obtain the following important statement:

Theorem 5.14 ([61, 170]) *For every H^s-vector field X, $s \geq 0$, on M with boundary ∂M there exists a unique divergence-free H^s-vector field Y tangent to the boundary ∂M and unique to within an additive constant.*

For M with boundary, $\mathcal{D}^s_\mu(M)$ is a smooth submanifold in $\mathcal{D}^s(M)$ and consequently in $H^s(M,N)$. The space $T_e\mathcal{D}^s_\mu(M) = \mathcal{E}^s(\wedge^k_t)$ is the space of all divergence-free H^s-vector fields on M tangent to the boundary ∂M. Let $P_e : T_eH^s(M,N) \to T_e\mathcal{D}^s_\mu(M)$ be the orthogonal projector in (5.9). The corresponding morphism $\bar{P}$, defined by analogy with (5.6), is C^∞-smooth and right-invariant. Thus we can introduce the covariant derivatives $\tilde{\nabla}$ and $\frac{\tilde{\mathrm{D}}}{\mathrm{d}t}$ by formulae (5.8), the geodesic spray $\mathcal{S}$ as in Theorem 5.9, etc. From Theorem 5.14 it follows that

$$Y = P_eX = X - \operatorname{grad} p. \tag{5.10}$$

Remark 5.15. In this case the following modification of Remark 5.12 holds. If a divergence-free vector X tangent to the boundary belongs to the class H^{s+k}, the corresponding right-invariant vector field $\bar{X}$ on $\mathcal{D}^s_\mu(M)$ is C^k-smooth. However, if a right-invariant vector field $\bar{X}$ on $\mathcal{D}^s_\mu(M)$ is C^k-smooth, the field $X = \bar{X}_e$ on M belongs to H^{s+k} in the interior of M and in the directions tangent to boundary, but it may not belong to this class in the directions normal to the boundary.

Note that in the case of a manifold M with boundary direct analogs of Theorems 5.9 and 5.10 remain true.

Strong Riemannian metrics

Let M be a manifold without boundary as above. Consider $g \in \mathcal{D}^s(M)$ and also X_g and $Y_g \in T_g\mathcal{D}^s(M)$ with $X_g = X \circ g$ and $Y_g = Y \circ g$ where $X, Y \in T_e\mathcal{D}^s(M)$ (see the definition of $T_g\mathcal{D}^s(M)$ above). Introduce on $T_g\mathcal{D}^s(M)$ a “strong” inner product $(\cdot,\cdot)^{(s)}_g$ by the formula

$$\begin{aligned}(X_g, Y_g)^{(s)}_g = \int_M (&\langle X_g(m), Y_g(m)\rangle_{g(m)} \\ &+ \langle (\mathrm{d}+\delta)^s X \circ g(m), (\mathrm{d}+\delta)^s Y \circ g(m)\rangle_{g(m)})\mu(\mathrm{d}m),\end{aligned} \tag{5.11}$$

where d is the differential, δ is the co-differential and $(\mathrm{d}+\delta)^2 = (\mathrm{d}\delta+\delta\mathrm{d}) = \Delta$ is the Laplace-de Rham operator. Since the Riemannian metric is given on M, we do not distinguish between 1-forms and vector fields.

We shall also use another strong right-invariant Riemannian metric:

$$((X_g, Y_g))^{(s)}_g = (TR_g^{-1}X_g, TR_g^{-1}Y_g)^{(s)}_e. \tag{5.12}$$

5.2 The Group of Diffeomorphisms of a Flat Torus

Consider the constructions, introduced above, in the particular case where M is a flat n-dimensional torus $\mathcal{T}^n$, i.e., $\mathcal{T}^n = \mathbb{R}^n/Z^n$ and the Riemannian metric on $\mathcal{T}^n$ is inherited from the Euclidean space $\mathbb{R}^n$ via factorization with respect to the integral lattice Z^n (see Remark 1.18).

Recall that the tangent bundle $T\mathcal{T}^n$ is trivial, i.e., there is a canonical identification of $T\mathcal{T}^n$ with $\mathcal{T}^n \times \mathbb{R}^n$ that is also inherited from $T\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$ (see Remark 1.40). Note that here the connection $\bar{\mathsf{H}}$ is generated by the flat connection on $\mathcal{T}^n$ inherited from the ordinary flat connection on $\mathbb{R}^n$ (the Levi-Civitá connection of the Euclidean inner product).

Definition 5.16. We introduce the operators:

(i) $\mathsf{B} : T\mathcal{T}^n \to \mathbb{R}^n$, the projection onto the second factor in $\mathcal{T}^n \times \mathbb{R}^n$;

(ii) $\mathsf{A}(m) : \mathbb{R}^n \to T_m\mathcal{T}^n$, the inverse to B (see (i)) sending $\mathbb{R}^n$ onto the tangent space $T_m\mathcal{T}^n$ to $\mathcal{T}^n$ at $m \in \mathcal{T}^n$;

(iii) $\mathsf{Q}_g = \mathsf{A}(g(m)) \circ \mathsf{B}$, the linear isomorphism $\mathsf{Q}_g : T_m\mathcal{T}^n \to T_{g(m)}\mathcal{T}^n$ where $g \in \mathcal{D}^s$ and $m \in \mathcal{T}^n$.

Note that if we specify a vector $X \in \mathbb{R}^n$, $\mathsf{A}(X) : \mathcal{T}^n \to T\mathcal{T}^n$ is a vector field on $\mathcal{T}^n$ with constant coordinates with respect to the standard basis $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}$ in tangent spaces. In particular, this vector field is C^∞-smooth and divergence-free with respect to the above-mentioned metric on $\mathcal{T}^n$.

By construction, for every $f \in \mathcal{D}^s(\mathcal{T}^n)$ and $X \in T_f\mathcal{D}^s(\mathcal{T}^n)$ the vector Q_gX lies in $T_g\mathcal{D}^s(\mathcal{T}^n)$. In particular, $\mathsf{Q}_eX \in T_e\mathcal{D}^s(\mathcal{T}^n)$. Note that even for $f \in \mathcal{D}^s_\mu(\mathcal{T}^n)$, the operator Q_e does not send $T_f\mathcal{D}^s_\mu(\mathcal{T}^n)$ to $T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$, but $P_e\mathsf{Q}_e(T_f\mathcal{D}^s_\mu(\mathcal{T}^n)) = T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$.

We describe the action of operator Q_g on tangent vectors to $\mathcal{D}^s(\mathcal{T}^n)$ as on mappings from $\mathcal{T}^n$ to $T\mathcal{T}^n$ and compare this action with the right shift. Recall that a vector $X \in T_e\mathcal{D}^s(\mathcal{T}^n)$, i.e., a vector field on $\mathcal{T}^n$, sends a point $m \in \mathcal{T}^n$ to the vector $(m, X(m))$. The right shift TR_g on $T_e\mathcal{D}^s(\mathcal{T}^n)$ sends the latter vector to $(g(m), X(g(m))$, and $\mathsf{Q}_g(m, X(m)) = (g(m), X(m))$.

Lemma 5.17 *The following relations hold:*

$$TR_{g^{-1}}(\mathsf{Q}_gX) = \mathsf{Q}_e(TR_{g^{-1}}X); \tag{5.13}$$

$$TR_g(\mathsf{Q}_{g^{-1}}X) = \mathsf{Q}_e(TR_gX). \tag{5.14}$$

Proof. According to the above formulae, $\mathsf{Q}_e(TR_{g^{-1}}X)$ sends a point $m \in \mathcal{T}^n$ to $(m, X(g^{-1}(m)))$. On the other hand, $\mathsf{Q}_gX = (g(m), X(m))$ and $TR_{g^{-1}}(\mathsf{Q}_gX) = (m, X(g^{-1}(m)))$. From this (5.13) follows. Formula (5.14) is obtained from (5.13) by replacing g with g^{-1}. □

Theorem 5.18 $\mathsf{Q}_g : T_\eta\mathcal{D}^s(\mathcal{T}^n) \to T_g\mathcal{D}^s(\mathcal{T}^n)$ *is the parallel translation in* $\mathcal{D}^s(\mathcal{T}^n)$ *with respect to* $\bar{\mathsf{H}}$.

Indeed, since (roughly speaking) the connectors on $\mathcal{T}^n$ and on $\mathcal{D}^s(\mathcal{T}^n)$ coincide, the parallel translations coincide as well.

For a specified $x \in \mathbb{R}^n$ we introduce the diffeomorphism $L_x : \mathcal{T}^n \to \mathcal{T}^n$ by the formula $L_x(m) = m + x$ modulo factorization with respect to the integral lattice. Evidently L_x is C^∞-smooth and preserves the volume. Note that $L_xg(m) = g(m) + x$ is in fact the left shift of $g \in \mathcal{D}^s(\mathcal{T}^n)$ by L_x.One can easily see that $TL_x = \mathsf{Q}_{L_x}$.

Theorem 5.19 *Let $g(t)$ be a geodesic of the flat connection $\bar{\mathsf{H}}$ on $\mathcal{D}^s(\mathcal{T}^n)$ (i.e., $\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\dot{g}(t) = 0$). Then $L_xg(t)$ is also a geodesic.*

Proof. Note that the derivative $\dot{g}(t)$ sends the point $m \in \mathcal{T}^n$ to the vector $(g(t)(m), \dot{g}(t)(m))$. Since $g(t)$ is a geodesic of a flat connection, it satisfies the relation $(g(t)(m), \ddot{g}(t)(m)) = (g(t)(m), 0)$. Then, since x is constant, $\frac{\mathrm{d}}{\mathrm{d}t}(L_xg(t))$ sends m to $(g(t)(m)+x, \dot{g}(t)(m))$ and for the covariant derivative we obtain $(g(t)(m) + x, \ddot{g}(t)(m)) = (g(t)(m) + x, 0)$. □

Part II
Stochastic Analysis

Chapter 6
Essentials from Stochastic Analysis in Linear Spaces

6.1 Some Definitions from Probability Theory and the Theory of Stochastic Processes

In this section we describe some facts and constructions from probability theory and the theory of stochastic processes, some of which are not generally included in standard university courses on these subjects. This is done mainly for convenience of reference, but if necessary this material can be used as an introduction to the subject. Nevertheless the reader is assumed to be familiar with the main notions of probability theory including the notions of a σ-algebra (in particular, a Borel σ-algebra), measure and independent random variables.

We consider random variables (measurable mappings) given on a complete probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and taking values in a finite-dimensional linear space $\mathbb{R}^n$ equipped with the Borel σ-algebra. For random variables with values in non-linear manifolds and in infinite-dimensional spaces there are analogous constructions.

A detailed exposition of this material can be found, e.g., in [175, 176, 194, 204, 208].

We say that a σ-subalgebra $\mathcal{B}_\xi$ in $\mathcal{F}$ is generated by a random variable $\xi : \Omega \to \mathbb{R}^n$ if $\mathcal{B}_\xi$ is the minimal σ-algebra containing the pre-images of all Borel sets in $\mathbb{R}^n$ under the mapping ξ or, equivalently, $\mathcal{B}_\xi$ is the minimal σ-algebra with respect to which ξ is measurable.

6.1.1 Stochastic processes. Cylinder sets

A *stochastic process* is a random variable, given on a probability space and taking values in $\mathbb{R}^n$, that depends on time. A process $\eta(t)$, $t \in [0, \infty)$, has *almost surely (a.s.) continuous sample paths* (or *trajectories*) P-a.s. if for

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_6,

$\omega \in \Omega$ the curve $\eta(t,\omega)$ is continuous in t. The space of continuous curves $C^0([0,\infty),\mathbb{R}^n)$ in this case is called the *space of (sample) paths* (or *trajectories*).

Specify a finite number $t_1,\dots,t_k \in [0,\infty)$ of time instants and a finite collection of Borel sets $B_1,\dots,B_k \subset \mathbb{R}^n$. A *cylinder set* in $C^0([0,\infty),\mathbb{R}^n)$, corresponding to the above collections of times and Borel sets, is the set of curves

$$J_{t_1,\dots,t_k} \times (B_1,\dots,B_k) = \{x(\cdot) \in C^0([0,\infty),R^n) \mid x(t_i) \in B_i\},$$

i.e., the set of curves that at time t_i take a value in the set B_i, and arbitrary values at the other times. The minimal σ-algebra that contains all cylinder sets for a finite time interval $t \in [0,T] \subset \mathbb{R}$ coincides with the Borel σ-algebra on the Banach space $C^0([0,T],\mathbb{R}^n)$ of continuous curves in $\mathbb{R}^n$ given on $[0,T]$, with the usual uniform norm. A stochastic process with a.s. continuous sample paths is usually considered as a random variable with values in $C^0([0,\infty),\mathbb{R}^n)$ equipped with the σ-algebra generated by cylinder sets, i.e., a measurable mapping from Ω to $C^0([0,\infty),\mathbb{R}^n)$ with respect to the σ algebra generated by cylinder sets in $C^0([0,\infty),\mathbb{R}^n)$ and the σ-algebra $\mathcal{F}$ in Ω.

Denote $C^0([0,\infty),\mathbb{R}^n)$ by $\tilde{\Omega}$ and the σ-algebra generated by cylinder sets by $\tilde{\mathcal{F}}$. Then, since a process $\xi(\cdot)$ can be considered as a measurable mapping from $(\Omega,\mathcal{F})$ to $(\tilde{\Omega},\tilde{\mathcal{F}})$, it generates a probability measure μ_ξ on $(\tilde{\Omega},\tilde{\mathcal{F}})$ in the usual way: $\mu_\xi(A) = \mathsf{P}(\xi^{-1}(A))$ for $A \in \tilde{\mathcal{F}}$. This measure is called the *measure generated by* $\xi(t)$ on $(\tilde{\Omega},\tilde{\mathcal{F}})$ or the *distribution* of $\xi(t)$.

For any probability measure μ given on $(\tilde{\Omega},\tilde{\mathcal{F}})$, one can construct a stochastic process $\eta_\mu(t)$ on the probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\mu)$ with values in $\mathbb{R}^n$ as follows: $\eta_\mu(t,\omega) = \omega(t)$ where the elementary event $\omega \in C^0([0,\infty),\mathbb{R}^n)$ is by definition a continuous curve $\omega : [0,\infty) \to R^n$. The process $\xi_\mu(t)$ is called the *coordinate process* on the probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\mu)$.

Note that for $\mu = \mu_\xi$ generated by a stochastic process $\xi(t)$ with continuous path in $\mathbb{R}^n$, the corresponding coordinate process has, by construction, the same distribution as $\xi(t)$.

Every stochastic process $\xi(t)$ in $\mathbb{R}^n$, $t \in [0,T]$, given on a probability space $(\Omega,\mathcal{F},\mathsf{P})$, determines three families of σ-subalgebras of the σ-algebra $\mathcal{F}$:

(i) "past" $\mathcal{P}_t^\xi$, generated by pre-images of Borel sets in $\mathbb{R}^n$ by all mappings $\xi(s) : \Omega \to \mathbb{R}^n$ for $0 < s < t$;
(ii) "future" $\mathcal{F}_t^\xi$, generated by pre-images of Borel sets in $\mathbb{R}^n$ by all mappings $\xi(s) : \Omega \to R^n$ for $t < s < T$;
(iii) "present" ("now") $\mathcal{N}_t^\xi$, generated by the mapping $\xi(t)$.

We suppose that all of these families are complete, i.e., contain all sets of probability zero: $\mathsf{P} = 0$.

Let $\mathcal{B}_t$ be a non-decreasing family of σ-subalgebras of the σ-algebra $\mathcal{F}$.

Definition 6.1. A random process $A(t)$ is said to be *non-anticipative* with respect to a *filtration* $\mathcal{B}_t$ if $A(t)$ is measurable with respect to $\mathcal{B}_t$ for every t.

Consider the measure space $(\tilde{\Omega}, \tilde{\mathcal{F}})$ introduced above. Denote by $\tilde{\mathcal{P}}_t$ the σ-algebra generated by cylinder sets with bases over $[0,t]$. Note that for any probability measure μ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ the coordinate process $\xi_\mu(t)$ on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$ is non-anticipative with respect to $\tilde{\mathcal{P}}_t$ and moreover, $\tilde{\mathcal{P}}_t$ is its "past".

6.1.2 Conditional expectation

Consider the Hilbert space $L^2(\Omega, \mathcal{F}, \mathsf{P})$ of square integrable random variables. Let $\mathcal{F}_0$ be a σ-subalgebra in $\mathcal{F}$. Consider $L^2(\Omega, \mathcal{F}_0, \mathsf{P})$, the Hilbert space of square integrable random variables that are measurable with respect to $\mathcal{F}_0$. It is clear that $L^2(\Omega, \mathcal{F}_0, \mathsf{P})$ is a closed subspace in $L^2(\Omega, \mathcal{F}, \mathsf{P})$. Denote by $Q : L^2(\Omega, \mathcal{F}, \mathsf{P}) \to L^2(\Omega, \mathcal{F}_0, \mathsf{P})$ the orthogonal projector. The projector Q extends to the projector in the corresponding space L^1 of integrable random variables.

Definition 6.2. For every $\xi \in L^1(\Omega, \mathcal{F}, \mathsf{P})$ the random variable $Q\xi \in L^1(\Omega, \mathcal{F}_0, \mathsf{P})$ is called the *conditional expectation* of ξ with respect to $\mathcal{F}_0$ and is denoted by $E(\xi|\mathcal{F}_0)$.

It is important to point out that (up to sets with probability zero) $E(\xi|\mathcal{F}_0)$ is the unique random variable in $L^1(\Omega, \mathcal{F}, \mathsf{P})$ such that for every set $A \in \mathcal{F}_0$ the equality $\int\limits_A \xi d\mathsf{P} = \int\limits_A E(\xi|\mathcal{F}_0)d\mathsf{P}$ holds. Recall that the existence of such a function follows from the Radon-Nikodym theorem. This description of $E(\xi|\mathcal{F}_0)$ is equivalent to Definition 6.2 and is often used as the definition in the probabilistic literature (see e.g., [208]). The details of the approach based on Definition 6.2 are given, e.g., in [194].

It is not hard to see that the usual mathematical expectation is the conditional expectation with respect to the trivial σ-algebra comprising two sets: $\emptyset$ and Ω.

Let $A \in \mathcal{F}$ and χ_A be the indicator of A. The value $\mathsf{P}(A|\mathcal{F}_0) = E(\chi_A|\mathcal{F}_0)$ is called a *conditional probability.*

Let us describe some properties of conditional expectation. These properties generally follow from the properties of projectors. A detailed presentation of this material can be found in [194, 208]

Theorem 6.3

(i) *Conditional expectation is a linear operator.*
(ii) *If η is measurable with respect to $\mathcal{F}_0$, $E(\eta|\mathcal{F}_0) = \eta$.*
(iii) *If $\mathcal{F}_1 \subset \mathcal{F}_0$, $E(E(\eta|\mathcal{F}_0)|\mathcal{F}_1) = E(\eta|\mathcal{F}_1)$. In particular, the equality $E(E(\eta|\mathcal{F}_0)) = E\eta$ holds.*

(iv) *If* $\mathcal{F}_1 \subset \mathcal{F}_0$, $E(E(\eta|\mathcal{F}_1)|\mathcal{F}_0) = E(\eta|\mathcal{F}_1)$.
(v) *If* η *does not depend on* $\mathcal{F}_0$, $E(\eta|\mathcal{F}_0) = E\eta$.
(vi) *Let* ξ *and* η *be random variables with values in* $\mathbb{R}$. *If* ξ *is measurable with respect to* $\mathcal{F}_0$, *then for every* η *the equality* $E(\xi\eta|\mathcal{F}_0) = \xi E(\eta|\mathcal{F}_0)$ *holds. In the case of vector-valued random variables this property is valid for inner products.*

Let ξ and η be random variables given on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and taking values in $\mathbb{R}^n$. By $E(\xi|\eta)$ we denote the conditional expectation of ξ with respect to the σ-algebra generated by pre-images of Borel sets in $\mathbb{R}^n$ under the mapping $\eta : \Omega \to \mathbb{R}^n$. One can easily show (see, e.g., [194]) that there exists a unique Borel measurable mapping $Y : \mathbb{R}^n \to \mathbb{R}^n$ such that $E(\xi \mid \eta) = Y(\eta)$. The mapping Y is called the *regression* of ξ with respect to η and is usually denoted by the expression $Y(x) = E(\xi \mid \eta = x)$.

6.1.3 Markov processes

Let a non-decreasing family $\mathcal{B}_t$ of σ-subalgebras of a σ-algebra $\mathcal{F}$, $t \in [0, \infty)$, be given. A random process $\xi(t)$ is called a *Markovian* (or *Markov*) *process* with respect to $\mathcal{B}_t$ if P-a.s. $\mathsf{P}(B \cap F \mid \mathcal{N}_t^\xi) = \mathsf{P}(B \mid \mathcal{N}_t^\xi) \cdot \mathsf{P}(F \mid \mathcal{N}_t^\xi)$ for every $t \in [0, \infty)$, $B \in \mathcal{B}_t$ and $F \in \mathcal{F}_t^\xi$ (see the Definition of "past", "future" and "present" in Section 6.1.1).

A process $\xi(t)$ is called a *simply Markovian* (or *simple Markov*) *process* if it is Markovian with respect its own "past" $\mathcal{P}_t^\xi$.

The next two conditions are equivalent to each other and to the fact that the process is Markovian with respect to $\mathcal{B}_t$.

1) For every $t \in [0, \infty)$ and every bounded $\mathcal{F}_t^\xi$-measurable random variable φ with values in $\mathbb{R}$ the relation $E(\varphi \mid \mathcal{B}_t) = E(\varphi \mid \mathcal{N}_t^\xi)$ holds.
2) For $t \geq s \geq 0$ and every (measurable) function $f(x)$, for which $\sup\limits_{x \in \mathbb{R}^n} |f(x)| < \infty$, the equality $E(f(\xi(t)) \mid \mathcal{B}_s) = E(f(\xi(t)) \mid \mathcal{N}_s^\xi)$ holds.

A random variable $\tau(\omega)$ taking values in $[0, \infty)$ is called a *random time.* A random time is called a *Markov time* if for every $t \geq 0$ the inclusion $\{\omega \mid \tau(\omega) \geq t\} \in \mathcal{B}_t$ holds. If $\mathsf{P}(\tau(\omega) < \infty) = 1$, the Markov time is called a *stopping time.*

6.1.4 Martingales and semi-martingales

A stochastic process $\eta(t)$ is called a *martingale* with respect to a non-decreasing family of σ-algebras $\mathcal{B}_t$, $t \in [0, \infty)$, if for every t the variable

$\eta(t)$ is measurable with respect to $\mathcal{B}_t$ (i.e., non-anticipative with respect to $\mathcal{B}_t$) and for every $t \geq s \geq 0$ the equality $E(\eta(t) \mid \mathcal{B}_s) = \eta(s)$ holds.

Directly from the definition of martingale and property (iii) of conditional expectation (see Section 6.1.2) it follows that the mathematical expectation of a martingale is constant.

A process $\eta(t)$ is called a *local martingale* if there exists a non-decreasing sequence of Markov times (stopping times) τ_n such that $\lim \tau_n = \infty$ and for every $\tau_n(\omega)$ the process $\eta(t \wedge \tau_n)$ is a martingale, where $t \wedge \tau_n = \min(t, \tau_n(\omega))$.

A stochastic process $\eta(t)$ is called a *semi-martingale* if $\eta(t) = A(t) + M(t)$ where $M(t)$ is a local martingale and $A(t)$ is a process whose sample paths a.s. have bounded variation in t (i.e., the Stieltjes integral of integrable functions is well-defined with those paths as integrators).

It is clear that a martingale is a local martingale and that a local martingale is a semi-martingale. Note that there is a construction of the integration of random functions with respect to semi-martingales (see [176]), a particular case of which is the stochastic integral with respect to the Wiener process that is described below in Section 6.2.

Under a smooth change of coordinates in $\mathbb{R}^n$ martingales and local martingales do not transform into analogous processes, but semi-martingales are transformed into semi-martingales. A decomposition into the sum of a local martingale and a process of bounded variation is possible but in this decomposition the summands are not the results of transformation of corresponding summands in the previous coordinate system. From this property it follows that the notion of a semi-martingale with values on a manifold is well-defined.

A stochastic process $\xi(t)$ is called a *backward martingale* with respect to a non-increasing family of σ-algebras $\mathcal{F}_t$ if for every t the variable $\xi(t)$ is measurable with respect to $\mathcal{F}_t$ and for every $s \geq t \geq 0$ the equality $E(\eta(t)|\mathcal{F}_s) = \eta(s)$ holds.

6.1.5 Weak convergence of probability measures

A detailed presentation of this material can be found in, e.g., [25, 194, 208]. Everywhere in this Section the symbol $\mathfrak{X}$ denotes a separable complete metric space and $\mathcal{B}$ is its Borel σ-algebra (i.e., the minimal σ-algebra generated by open sets). Recall that a measure μ on $(\mathfrak{X}, \mathcal{B})$ is called a *probability measure* if $\mu(\mathfrak{X}) = 1$.

Definition 6.4. A sequence of probability measures μ_n on $(\mathfrak{X}, \mathcal{B})$ weakly converges to a probability measure μ_0 if $\int\limits_{\mathfrak{X}} f \, \mathrm{d}\mu_n \to \int\limits_{\mathfrak{X}} f \, \mathrm{d}\mu_0$ for every continuous bounded function $f : \mathfrak{X} \to \mathbb{R}$.

Definition 6.5. A family of probability measures $\{\mu_\alpha\}$ on $(\mathfrak{X}, \mathcal{B})$ is called *weakly relatively compact* if every sequence of measures from $\{\mu_\alpha\}$ has a weakly convergent subsequence.

The measure to which the subsequence converges in Definition 6.5 may not belong to the set $\{\mu_\alpha\}$. For short we shall often call weakly relatively compact set of measures *weakly compact*, omitting the word "relatively".

Theorem 6.6 (Prokhorov's theorem) *A family of probability measures $\mathcal{M}$ on $(\mathfrak{X}, \mathcal{B})$ is weakly relatively compact if and only if for every $\varepsilon > 0$ there exists a compact $K_\varepsilon \subset \mathfrak{X}$ such that $\mu_\alpha(K_\varepsilon) > 1 - \varepsilon$ for every $\mu_\alpha \in \mathcal{M}$.*

6.2 A Survey on Stochastic Integrals and Equations

In this section we briefly describe the basic facts from the theory of stochastic integrals and stochastic differential equations necessary for understanding their geometric properties below. We consider only integrals with respect to a Wiener process since they play the main role in the forthcoming sections. A complete and detailed exposition of this material can be found in many monographs and textbooks (see, e.g., [76, 83, 84, 162, 165, 175, 176, 216]). We should particularly highlight the excellent introductory paper [50] which illuminates those aspects of the theory that are especially important for our approach.

6.2.1 White noise and Wiener processes

We begin this section with a physically motivated observation that leads to an intuitive introduction to Wiener processes (for details, see [160]).

Consider an ordinary differential equation $\dot{x}(t) = F(t, x(t))$ in $\mathbb{R}^n$ and suppose that its right-hand side is subjected to an additive random influence that satisfies the following physically natural assumptions:

(a) the mechanism that produces the randomness is the same at all times;
(b) the randomness occurs at any time t independently of the other times;
(c) the mathematical expectation of the random variable equals 0 while the dispersion equals 1.

We can interpret (a) as saying that the process has independent values at different times and (b) as saying that the distributions of the values at all times are the same. Assumption (c) is given for simplicity; one can consider more general processes satisfying (a) and (b).

The above-mentioned process is denoted by $\dot{w}(t)$ and is called *white noise*. It turns out that this process takes values in generalized functions and so it is rather difficult to deal with. We shall avoid the use of generalized functions in the usual way: we introduce the process $w(t) = \int_0^t \dot{w}(s)\mathrm{d}s$. From the properties (a)–(c) of $\dot{w}(t)$ we intuitively derive that $w(t)$ must have a.s. continuous sample paths and independent increments such that for a given difference $t - s = \delta$ all increments $w(t_1) - w(s_1)$ with $t_1 - s_1 = \delta$ have the same distribution. Finally, for all such increments, $E(w(t) - w(s)) = 0$ and $E(w(t) - w(s))^2 = t - s$.

The differential equation we started with, with the above random influence, has the form $\dot{x}(t) = F(t, x(t)) + \dot{w}(t)$. Avoiding the use of generalized functions, we transform it into the integral equation $x(t) = x_0 + \int_0^t F(s, x(s))\mathrm{d}s + w(t)$. For a differential equation without random influence this transformation yields the equivalent integral equation. In the presence of random influence the transformation makes the equation easier to work with since it is given in terms of processes with continuous sample paths.

$w(t)$ has all the properties of a certain process, called a Wiener process, that we want to formally introduce in this Section. The precise definition requires the following abstract scheme.

Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space and $\mathcal{B}_t$, $t \in [0, \infty)$, be a nondecreasing family of σ-subalgebras of the σ-algebra $\mathcal{F}$. In what follows, we assume that the σ-algebras $\mathcal{B}_t$ are complete, i.e., they contain all sets from $\mathcal{F}$ of measure zero.

Here we consider only stochastic processes (random variables) on $(\Omega, \mathcal{F}, \mathsf{P})$ with values in a Euclidean space with an inner product $(\cdot, \cdot)$. Specifying a basis, we shall describe its vectors by columns of coordinates, i.e., we identify this space with $\mathbb{R}^n$.

Definition 6.7. A stochastic process $w(t)$ is called a *Wiener process* (relative to the family $\mathcal{B}_t$) if:

1) the sample paths of $w(t)$ are almost surely (a.s.) continuous in t;
2) $w(t)$ is a square integrable martingale with respect to $\mathcal{B}_t$;
3) $w(0) = 0$ and $E((w(t) - w(s))^2 | \mathcal{B}_s) = t - s$ for $t \geq s$.

In this case it is said that the Wiener process $w(t)$ is *adapted to* $\mathcal{B}_t$.

From Definition 6.7 we deduce that the Wiener process has the (intuitive) properties that we listed the beginning of this Section:

Theorem 6.8 (Levi, see, e.g., [175]) *If $w(t)$ is a Wiener process, then it has stationary independent Gaussian increments. Furthermore, $w(t)$ satisfies the following conditions: $E(w(t) - w(s)) = 0$ and $E((w(t) - w(s))^2) = t - s$ for $t \geq s$.*

In other words, for $t \geq s$, the increment $w(t) - w(s)$ is independent of $\mathcal{B}_s$ and has the same probability distribution as $w(t - s)$.

The distribution density $\rho^w(t, x)$ of a Wiener process in $\mathbb{R}^n$ is described by the formula (see, e.g., [162])

$$\rho^w(t,x) = \frac{1}{(2\pi t)^{\frac{n}{2}}}\, \mathrm{e}^{-\frac{x^2}{2t}}. \tag{6.1}$$

One can easily see that $\mathcal{P}_t^w \subset \mathcal{B}_t$ where $\mathcal{P}_t^w$ is the "past" of $w(t)$ (see Section 6.1.1).

Consider the space $\tilde{\Omega} = C^0([0,\infty),\mathbb{R}^n)$ of continuous curves in $\mathbb{R}^n$ given on the semi-infinite interval $[0,\infty)$. Introduce in $\tilde{\Omega}$ the σ-algebra $\tilde{\mathcal{F}}$ generated by cylinder sets (see Section 6.1.1). As for the other processes with continuous sample paths, every Wiener process can be regarded as a mapping of the measure space $(\Omega,\mathcal{F})$ into the measure space $(\tilde{\Omega},\tilde{\mathcal{F}})$. Thus, P gives rise to a measure ν on $(\tilde{\Omega},\tilde{\mathcal{F}})$ according to the construction given in Section 6.1.1. The measure ν, called the *Wiener measure*, depends only on the inner product on $\mathbb{R}^n$, not on a specific Wiener process $w(t)$. The Wiener measure enables one to introduce the probability distributions of $w(t)$ in $\tilde{\Omega} = C^0([0,\infty),\mathbb{R}^n)$, i.e., the conditional probability distributions of the random variables $w(t_1),\dots,w(t_k)$ in $\mathbb{R}^n$ for all collections $t_1,\dots,t_k$.

Consider $(\tilde{\Omega},\tilde{\mathcal{F}},\nu)$ as a probability space and define the coordinate process $\bar{w}(t)$ on $(\tilde{\Omega},\tilde{\mathcal{F}},\nu)$ (see Section 6.1.1) via $\tilde{w}(t,\omega) = \omega(t)$ (here the elementary event $\omega \in C^0([0,\infty),\mathbb{R}^n)$ is by definition a continuous curve $\omega : [0,\infty) \to \mathbb{R}^n$). Consider the σ-algebra $\bar{\mathcal{B}}_t$ generated by the cylinder sets with base over $[0,t]$, i.e., $\bar{\mathcal{B}}_t = \mathcal{P}_t^{\bar{w}}$. Clearly, $\bar{w}(t)$ is a Wiener process relative to the family $\bar{\mathcal{B}}_t$. The process $\bar{w}(t)$ is called a *Brownian motion process* or a *standard Wiener process*.

Remark 6.9. Often in the probability literature one finds that the Wiener process is described as unique. This phrase means only that the standard Wiener process is unique since the Wiener measure ν (as well as the coordinate process) is unique on $(\tilde{\Omega},\tilde{\mathcal{F}})$. But there are plenty of concrete realizations of Wiener processes on various probability spaces. In particular different Wiener processes can be independent.

To facilitate further references, we summarize here some results on Wiener processes.

Theorem 6.10 *Any Wiener process $w(t)$ has the following properties:*

1) *A sample path of $w(t,\omega)$ is a.s. (i.e., with probability 1) non-differentiable for all t and has unbounded variation on any arbitrarily small interval.*
2) *The coordinates $w^j(t)$ of $w(t)$ are one-dimensional Wiener processes that are mutually independent and the orthogonal projection of $w(t)$ to any k-dimensional subspace of $\mathbb{R}^n$ is a k-dimensional Wiener process.*
3) *Let a be an orthogonal operator in $\mathbb{R}^n$. Then $a \circ w(t)$ is a Wiener process. In particular, if $\bar{w}(t)$ is a standard Wiener process, then so is $a \circ \bar{w}(t)$, i.e., the Wiener measure is invariant under the action of orthogonal operators on $\mathbb{R}^n$.*

Note that assertion 1) of Theorem 6.10 clarifies the fact that white noise, the "derivative" of a Wiener process, takes values only in generalized functions.

6.2.2 Stochastic integrals

Our goal in this section is to define the stochastic integral with respect to a Wiener process. For the sake of simplicity, we restrict our attention to the construction based on a Riemann integral. An approach involving a Lebesgue integral can be found in [76, 83, 84, 162, 175].

Specify a positive constant $l < \infty$. Let $A\colon [0,l] \times \Omega \to L(\mathbb{R}^k, \mathbb{R}^n)$ be a random operator function, i.e., $A(t)$ is a random linear operator from $\mathbb{R}^k$ to $\mathbb{R}^n$ for every $t \in [0,l]$. In what follows the main role will be played by the particular case $k = n$, i.e., where $A(t)$ is a random linear operator in $\mathbb{R}^n$.

Consider a Wiener process $w(t)$ with respect to $\mathcal{B}_t$ with values in $\mathbb{R}^k$. To define the Itô integral of $A(t)$, pick a partition $q = (0 = t_0 < t_1 < \ldots < t_q = l)$ of the interval $[0,l]$ and consider the integral sum

$$\sum_{i=0}^{q-1} A(t_i)\big(w(t_{i+1}) - w(t_i)\big). \tag{6.2}$$

Note that we have selected the argument in $A(\cdot)$ as the left end of $[t_i, t_{i+1}]$ in the i-th summand. The limit (if it exists) of such sums as $\operatorname{diam} q \to 0$ (usually in the space $L^2((\Omega, \mathcal{F}, \mathsf{P}), \mathbb{R}^n)$ but possibly with respect to some other type of convergence of random variables) is called the *Itô integral* of $A(t)$ and is denoted by $\int_0^l A(t)\mathrm{d}w(t)$. Since the trajectories of $w(t)$ have a.s. unbounded variation, the Itô integral cannot be defined as the Stieltjes integral along every trajectory.

It turns out that under certain boundedness hypotheses, the Itô integral does exist as the L^2-limit of the integral sums when $A(t)$ is non-anticipative with respect to $\mathcal{B}_t$. In particular, it exists (as a Lebesgue type integral) if the entries $A_i^j(t)$ of $A(t)$ satisfy the equality

$$\mathsf{P}\left\{\omega \,\middle|\, \int_0^l (A_i^j)^2(t,\omega)\mathrm{d}t < \infty\right\} = 1. \tag{6.3}$$

The Itô integral with varying upper limit is the process defined by

$$\int_0^t A(\tau)\,\mathrm{d}w(\tau) = \int_0^l \chi_t A(\tau)\,\mathrm{d}w(\tau),$$

where χ_t is the characteristic function of $[0,t]$. Note that $\int_0^t A(\tau)\,\mathrm{d}w(\tau)$ is linear in A and $\mathrm{d}w$. Some other important properties of the integral are given in the next theorem.

Theorem 6.11 *The process $\int_0^t A(\tau)\,\mathrm{d}w(\tau)$ has the following properties:*

(1) *it is non-anticipative with respect to $\mathcal{B}_t$;*
(2) *it is a martingale relative to $\mathcal{B}_t$;*
(3) *its sample paths are a.s. continuous in t.*

Let us outline the proof of Theorem 6.11(3) for future reference. Consider processes with continuous trajectories of the form

$$\sum\nolimits_I^q(t) = \sum_{i=1}^{q-1} \chi_t(t_{i+1})A(t_i)\big(w(t_{i+1}) - w(t_i)\big) + A(t_k)\big(w(t) - w(t_k)\big),$$

where $k = \max\{i\,|\chi_t(t_i) = 1\}$. Under certain hypotheses, the sequence $\sum_I^q(t)$ contains a subsequence that converges a.s. uniformly to $\int_0^t A(\tau)\,\mathrm{d}w(\tau)$. This yields assertion (3).

Note the following discrepancy between the ordinary Riemann and stochastic Itô integrals. The former calculated for, say, a bounded continuous function $f(t)$ with respect to some power $\alpha > 1$ of $\mathrm{d}t$ always equals zero since the integral sum $\sum_{i=1}^{q-1} f(t^*)(t_{i+1} - t_i)^\alpha$, $t^* \in [t_i, t_{i+1}]$, tends to zero as $\operatorname{diam} \mathrm{q}$ tends to zero. However, this is not the case for the latter. One can define the multiple stochastic integral $\int_0^t a(\tau)\,\mathrm{d}w^1(\tau)\cdots \mathrm{d}w^k(\tau)$ of a given stochastic process $a(\cdot)$ to be the limit of the integral sums $\sum_{i=1}^{q-1} a(t_i)\big(w^1(t_{i+1}) - w^1(t_i)\big)\cdots\big(w^k(t_{i+1}) - w^k(t_i)\big)$ which may not be equal to zero.

Later on, we shall use the following result on the existence and properties of multiple integrals.

Theorem 6.12 *Let $\alpha(t)$ be a random real-valued function and $w(t)$ be a Wiener process with values in $\mathbb{R}^n$, i.e., $w(t) = (w^1(t), \dots, w^n(t))$, where the coordinates $w^i(t)$, $i = 1, \dots, n$ are mutually independent one-dimensional Wiener processes. Then:*

(i) $\int_0^t \alpha(\tau)\mathrm{d}w^i(\tau)\mathrm{d}w^j(\tau) = 0 \qquad i \neq j;$

(ii) $\int_0^t \alpha(\tau)(\mathrm{d}w^i(\tau))^2 = \int_0^t \alpha(\tau)\mathrm{d}\tau;$

(iii) $\int_0^t \alpha(\tau)\mathrm{d}\tau\mathrm{d}w^i(\tau) = 0;$

(iv) $\int_0^t \alpha(\tau)(\mathrm{d}w^i(\tau))^3 = 0;$

(v) *all integrals of higher order in $\mathrm{d}\tau$ and $\mathrm{d}w^i(\tau)$ exist and are equal to zero.*

Here we just outline the proof. Assertion (i) follows immediately from the hypothesis that $w^i(t)$ and $w^j(t)$ are independent. To prove (ii), it suffices to observe that for any Wiener process $w(t)$ we have $E\Big(\big(w(t)-w(s)\big)^2\Big) = |t-s|$. Assertions (iii)–(v) result from the fact that the multiple Riemann integral with respect to $(\mathrm{d}t)^k$, $k > 1$, is equal to zero.

Remark 6.13. It is sometimes useful to use white noise (the "derivative" of a Wiener process, see Section 6.2.1) to give a better physical interpretation of solutions of certain equations but, as mentioned in Section 6.2.1, it is difficult to use in practice since it takes values in generalized functions. The use of the Itô integral allows one to avoid dealing with white noise by using integral (rather than differential) equations and taking into account that $\int_0^t A(t)\dot{w}(\tau)\,\mathrm{d}\tau = \int_0^t A(t)\,\mathrm{d}w(\tau)$. Then, if an ordinary differential equation $\dot{x}(t) = a(t, x(t))$ is subjected to a random perturbation and takes the form $\dot{x}(t) = a(t, x(t)) + A(t, x(t))\dot{w}(t)$ where $A(t, x)$ is a linear operator, a mathematically exact description of the perturbed equation is given by transition to the integral equation $x(t) = x_0 + \int_0^t a(\tau, x(\tau))\mathrm{d}\tau + \int_0^t A(\tau, x(\tau))\mathrm{d}w(\tau)$. The latter equation is called the *stochastic differential equation in Itô form* or *Itô stochastic differential equation*. Such equations are considered in detail below.

It should be pointed out that the value of a stochastic integral depends on the point in $[t_{i+1}, t_i]$ that is substituted as the argument value of the integrand in the summands of the integral sum. Recall that in integral sums of Itô integrals we choose the left end of $[t_{i+1}, t_i]$.

Alternative choices yield some other versions of stochastic integrals. In particular, we introduce the so-called backward (or anticipative) integral $\int_0^t A(\tau)\,\mathrm{d}_* w(\tau)$ [152] as the limit of the following integral sums

$$\sum_{i=0}^{q-1} A(t_{i+1})\big(w(t_{i+1}) - w(t_i)\big), \tag{6.4}$$

(i.e., choosing the right end of $[t_i, t_{i+1}]$) if, of course, the limit exists. This integral differs, in general, from the Itô integral. For example, the backward integral with varying upper limit is not a martingale relative to $\mathcal{B}_t$.

Considering the integral sums

$$\Sigma_S^q = \sum_{i=0}^{q-1} \frac{A(t_{i+1}) + A(t_i)}{2}(w(t_{i+1}) - w(t_i)) \tag{6.5}$$

we arrive at the *Stratonovich integral* $\int_0^t A(\tau) \circ \mathrm{d}w(\tau)$ (or $\int_0^t A(\tau)\,\mathrm{d}_S w(\tau)$) defined as the limit of these sums (if it exists) (see [50, 216].) It is easy to see that the Stratonovich integral is equal to half of the sum of the Itô and backward integrals, provided that all three integrals exist. The Stratonovich integral with varying upper limit can be defined in the standard way. Note, however, that this integral (like the anticipative integral and unlike the Itô integral with varying upper limit) is not a martingale with respect to $\mathcal{B}_t$.

Under an extra hypothesis, the Stratonovich integral may be defined as the limit of the integral sums $\sum_{i=0}^{q-1} A\left(\frac{t_{i+1}-t_i}{2}\right)\left(w(t_{i+1})-w(t_i)\right)$ (i.e., where the value of A is taken at the mid-point of the segment $[t_i, t_{i+1}]$, see [177]).

Remark 6.14. The idea to use the mid-point, as in the definition of the Stratonovich integral sums (6.5), is due to Richard Feynman.

The differentials $\mathrm{d}w$, $\mathrm{d}_* w$, and $\circ\,\mathrm{d}w = \mathrm{d}_S w$ (appearing in the definitions of the Itô, anticipative and Stratonovich integrals) are conveniently called the *forward*, *backward* and *symmetric* differentials, respectively, in reference to the location of the point t in $[t_i, t_{i+1}]$ at which the value of A is evaluated. The terminology we introduce here is actively used below.

Let us now turn to the formulas relating the values of the three integrals. By the definition of the Stratonovich integral, we have

$$\sum\nolimits_S^q = \sum\nolimits_I^q + \frac{1}{2}\sum_{i=0}^{q-1}\left(A(t_{i+1}) - A(t_i)\right)\left(w(t_{i+1}) - w(t_i)\right)$$

where $\sum_I^q$ is the Itô integral sum (6.2). The limit of the second sum on the right-hand side is a second order integral in $\mathrm{d}A$ and $\mathrm{d}w$, which can naturally be denoted by $\int_0^t \mathrm{d}A\,\mathrm{d}w$. Thus,

$$\int_0^t A(\tau)\circ \mathrm{d}w(\tau) = \int_0^t A(\tau)\,\mathrm{d}w(\tau) + \frac{1}{2}\int_0^t \mathrm{d}A(\tau)\,\mathrm{d}w(\tau). \tag{6.6}$$

Similarly, one can show that

$$\int_0^t A(\tau)\,\mathrm{d}_* w(\tau) = \int_0^t A(\tau)\,\mathrm{d}w(\tau) + \int_0^t \mathrm{d}A(\tau)\,\mathrm{d}w(\tau). \tag{6.7}$$

An *Itô process* is a process $\xi(t)$ of the form

$$\xi(t) = \xi_0 + \int_0^t a(s)\mathrm{d}s + \int_0^t A(s)\mathrm{d}w(s), \tag{6.8}$$

where $a(t)$ is a process with sample paths a.s. having bounded variation.

Let $f(t,x)$ be a continuously differentiable mapping from $\mathbb{R}\times\mathbb{R}^n$ to $\mathbb{R}^n$. Consider its Taylor decomposition at a neighborhood of the point (t_0, x_0):

$$f(t,x) = f(t_0,x_0) + \frac{\partial f}{\partial t}\Delta t + f'\Delta x + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(\Delta t)^2 + \frac{1}{2}f''(\Delta x, \Delta x) + \dots,$$

where the primes denote derivatives of f in x at x_0 (recall that f' and f'' are linear and bilinear operators respectively). Having replaced in this formula

$\varDelta x$ by the increment $\varDelta\xi(s) = a(s)\varDelta s + A(s)\varDelta w(s)$ of the Itô process $\xi(t)$, we obtain

$$\begin{aligned}
f(s,\xi(s)) &= f(t_0,x_0) + \frac{\partial f}{\partial s}\varDelta s + f'(a(s)\varDelta s + A(s)\varDelta w(s)) + \frac{1}{2}\frac{\partial^2 f}{\partial s^2}(\varDelta s)^2 \\
&\qquad + \frac{1}{2}f''(a(s)\varDelta s + A(s)\varDelta w(s), a(s)\varDelta s + A(s)\varDelta w(s)) + \ldots \\
&= f(t_0,x_0) + \frac{\partial f}{\partial s}\varDelta s + f'a(s)\varDelta s + f'A(s)\varDelta w(s) \\
&\qquad + \frac{1}{2}(f''(a(s),a(s))(\varDelta s)^2 + f''(a(s)\varDelta s, A(a)\varDelta w(s)) \\
&\qquad + f''(A(s)\varDelta w(s), a(s)\varDelta s) \\
&\qquad + f''(A(s)\varDelta w(s), A(s)\varDelta w(s)) + \ldots
\end{aligned} \tag{6.9}$$

After integrating formula (6.9) we obtain that if $f(t,x)$ is C^1-smooth in t and C^2-smooth in x, the so-called classical *Itô formula* (see [50, 76, 83, 84, 162, 175], for example) holds for the process $f(\xi(t))$:

$$\begin{aligned}
f(\xi(t)) = f(\xi_0) &+ \int_0^t \left[\frac{\partial f}{\partial t} + f'(a(s)) + \frac{1}{2}\mathrm{tr} f''(A(s),A(s))\right]\mathrm{d}s \\
&+ \int_0^t f'(A(s))\mathrm{d}w(s),
\end{aligned} \tag{6.10}$$

where

$$\mathrm{tr} f''(A(s),A(s)) = \sum_{i=1}^{n} f''(A(s)e_i, A(s)e_i), \tag{6.11}$$

and $e_1,...,e_k$ is an arbitrary orthonormal frame in $\mathbb{R}^k$. Indeed, by formulae (i) and (ii) from Theorem 6.12, exactly one second order integral

$$\frac{1}{2}\int_0^t f''(A(s)\mathrm{d}w(s), A(s)\mathrm{d}w(s)) = \int_0^t \frac{1}{2}\mathrm{tr} f''(A(s),A(s))\mathrm{d}s$$

is not equal to zero, which yields (6.10) (note that for the ordinary Riemann integral in a non-random integrator, all integrals of the second and higher orders equal zero and so all summands with derivatives of f of order higher than 1 vanish upon integration).

Remark 6.15. It is a well-known result in linear algebra that the trace, introduced by formula (6.11), does not depend on the choice of orthonormal frame $e_1,\ldots,e_k$ in $\mathbb{R}^k$.

A *backward Itô process* is a process of the form

$$\xi(t) = \xi_0 + \int_0^t a(s)\mathrm{d}s + \int_0^t A(s)\mathrm{d}_* w(s). \tag{6.12}$$

For it the so-called *backward Itô formula* holds:

$$f(\xi(t)) = f(\xi_0) + \int_0^t \left[\frac{\partial f}{\partial t} + f'(a(s)) - \frac{1}{2}\mathrm{tr} f''(A(s), A(s))\right] \mathrm{d}s \\ + \int_0^t f'(A(s))\mathrm{d}_* w(s)$$

Indeed, in the Taylor expansion with respect to the right end of an interval the summands with even derivatives change sign. Represent $\int_0^t A(s)\mathrm{d}_* w(s)$ by formula (6.7) and substitute into the latter expansion. It is not hard to see that the integral of higher order, in which $\mathrm{d}A\mathrm{d}w$ is an argument in f'', equals zero. Then (6.13) is proved by the same argument as (6.10).

Let us call a *Stratonovich process* a process of the form

$$\xi(t) = \xi_0 + \int_0^t a(s)\mathrm{d}s + \int_0^t A(s) \circ \mathrm{d}w(s). \tag{6.13}$$

If $f(t, x)$ is a smooth mapping as above,

$$f(\xi(t)) = f(\xi_0) + \int_0^t \left[\frac{\partial f}{\partial t} + f'(a(s))\right] \mathrm{d}s + \int_0^t f'(A(s)) \circ \mathrm{d}w(s). \tag{6.14}$$

Formula (6.14) is proved by the same arguments as (6.10) and (6.13) modified by the fact that the Taylor expansion with respect to the mid-point does not contain summands of even order at all. Note that the form of formula (6.14) coincides with that for the transformation of non-random smooth curves.

Definition 6.16. An Itô process $\xi(t)$ is called a *diffusion type process* if both $a(t)$ and $A(t)$ are not anticipative with respect to the "past" filtration $\mathcal{P}_t^\xi$ of $\xi(\cdot)$ and the Wiener process $w(t)$ is adapted to $\mathcal{P}_t^\xi$.

Diffusion type processes exist, say, as solutions of the so-called Itô diffusion type equations (see Definition 6.21 below).

Definition 6.17. A diffusion type process $\xi(t)$ is called a *diffusion process* if $\beta(t) = a(t, \xi(t))$ and $A(t) = A(t, \xi(t))$ where $a(t, x)$ and $A(t, x)$ are Borel measurable mappings of $\mathbb{R} \times \mathbb{R}^n$ to $\mathbb{R}^n$ and to the space of linear operators $L(\mathbb{R}^k, \mathbb{R}^n)$, respectively.

6.2.3 Stochastic differential equations

Let on the space $\mathbb{R}^n$ a non-autonomous vector field $a(t, x)$ and a non-autonomous field of linear operators $A(t, x)$ be given (i.e., $A(t, x) : \mathbb{R}^k \to \mathbb{R}^n$ is a linear operator depending on $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$). A *Stochastic differential* *equation* (SDE) *in Itô form* or an *Itô stochastic differential equation* is an integral equation

$$\xi(t) = \xi(0) + \int_0^t a(\tau, \xi(\tau))\mathrm{d}\tau + \int_0^t A(\tau, \xi(\tau))\mathrm{d}w(\tau) \tag{6.15}$$

where the second summand on the right-hand side is a Lebesgue integral. Here we do not discuss the conditions under which all integrals in this expression are well-defined. An interpretation of Itô stochastic differential equations as ordinary differential equations with random perturbations is described in Remark 6.13.

Equation (6.15) is usually written in the following formal differential form (identical in meaning to (6.15))

$$\mathrm{d}\xi(t) = a(t, \xi(t))\mathrm{d}t + A(t, \xi(t))\mathrm{d}w(t). \tag{6.16}$$

Equations of the form (6.15) are often called diffusion equations since diffusion processes are described by equations of this sort (see below).

A *Stochastic differential equation in Stratonovich form* (or *Stratonovich stochastic differential equation*) is the integral equation with Stratonovich integral

$$\xi(t) = \xi(0) + \int_0^t a(\tau, \xi(\tau))\mathrm{d}\tau + \int_0^t A(\tau, \xi(\tau)) \circ \mathrm{d}w(\tau), \tag{6.17}$$

which is usually written in the reduced differential form as follows:

$$\mathrm{d}\xi(t) = a(t, \xi(t))\mathrm{d}t + A(t, \xi(t)) \circ \mathrm{d}w(t). \tag{6.18}$$

Definition 6.18. In a stochastic differential equation the coefficient $a(t, x)$ is called the *drift* and the bilinear form ((2, 0)-tensor field) $A(t, x) \circ A^*(t, x)$ is called the *diffusion coefficient*.

There are more general types of SDEs called diffusion type SDEs. In these equations the coefficients $a(t, \cdot)$ and $A(t, \cdot)$ depend not on $x \in \mathbb{R}^n$ but on a curve $x(\cdot)$ given on the interval $[0, t]$, i.e., they are equations with delayed argument.

The exact description is as follows. Specify an interval $[0, l] \subset [0, \infty)$ and consider the mappings $a : [0, l] \times C^0([0, l], \mathbb{R}^n) \to \mathbb{R}^n$ and $A : [0, l] \times C^0([0, l], \mathbb{R}^n) \to L(\mathbb{R}^k, \mathbb{R}^n)$ where $L(\mathbb{R}^k, \mathbb{R}^n)$ is the space of linear operators from $\mathbb{R}^k$ to $\mathbb{R}^n$ and $C^0([0, l], \mathbb{R}^n)$ is equipped with the σ-algebra of cylinder sets. We always assume that $a(t, x(\cdot))$ and $A(t, x(\cdot))$ satisfy the following conditions:

Condition 6.19

(i) *The mappings $a(t, x(\cdot))$ and $A(t, x(\cdot))$ are jointly measurable.*

(ii) *For every* $t \in [0, l]$ *the mappings* $a(t, \cdot) : C^0([0, l], \mathbb{R}^n) \to \mathbb{R}^n$ *and* $A(t, \cdot) : C^0([0, l], \mathbb{R}^n) \to L(\mathbb{R}^k, \mathbb{R}^n)$ *are measurable with respect to the* σ*-algebra generated by cylinder sets with bases over* $[0, t]$ *in* $C^0([0, l], \mathbb{R}^n)$*, and the Borel* σ*-algebras in* $\mathbb{R}^n$ *and* $L(\mathbb{R}^k, \mathbb{R}^n)$*.*

Remark 6.20. Condition 6.19(ii) (cf. Condition 4.12) is equivalent to the fact that if for any $t \in [0, l]$ two different curves $x_1(\cdot)$ and $x_2(\cdot)$ coincide on the interval $[0, t]$, then $a(t, x_1(\cdot)) = a(t, x_2(\cdot))$ and $A(t, x_1(\cdot)) = A(t, x_2(\cdot))$ (cf. Remark 4.13). For details, see [83] .

Definition 6.21. An equation of Itô type

$$\mathrm{d}\xi(t) = a(t, \xi(\cdot))\mathrm{d}t + A(t, \xi(\cdot))\mathrm{d}w(t) \tag{6.19}$$

is called a *diffusion type stochastic differential equation.*

Equation (6.19) is a reduced form of the integral expression

$$\xi(t) = \xi_0 + \int_0^t a(s, \xi(\cdot))\mathrm{d}s + \int_0^t A(s, \xi(\cdot))\mathrm{d}w(s). \tag{6.20}$$

It is clear that equation (6.16) is a particular case of (6.19).

We shall often require that the coefficients of equation (6.19) in addition satisfy the following:

Condition 6.22 *The mappings* $a(t, x(\cdot))$ *and* $A(t, x(\cdot))$ *are jointly continuous.*

Sometimes one also needs to consider the equations with random coefficients (i.e., coefficients explicitly depending on $\omega \in \Omega$).

In the theory of stochastic differential equations one distinguishes between two types of solution: strong and weak.

Definition 6.23. Equation (6.16) ((6.18) or (6.19), respectively) has a *strong solution* $\xi(t)$ if for every Wiener process $w(t)$ on a probability space, and adapted to a filtration $\mathcal{B}_t$, there exists a stochastic process $\xi(t)$ on the same probability space as $w(t)$ and non-anticipative with respect $\mathcal{B}_t$, such that for $\xi(t)$ and $w(t)$ a.s. for every t in some interval, equality (6.15) ((6.17) or (6.20), respectively) is fulfilled.

Definition 6.24. Equation (6.16) ((6.18) or (6.19), respectively) has a *weak solution* $\xi(t)$ if there exist a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, a non-decreasing family $\mathcal{B}_t$ of σ-subalgebras of the σ-algebra $\mathcal{F}$, a process $\xi(t)$ in $\mathbb{R}^n$ non-anticipative with respect to $\mathcal{B}_t$, and a Wiener process $w(t)$ in $\mathbb{R}^k$ adapted to $\mathcal{B}_t$, such that for $\xi(t)$ and $w(t)$ a.s. for every t in some interval, the equation (6.15) ((6.17) or (6.20), respectively) is fulfilled.

It should be emphasized that a strong solution is well-defined on every probability space on which a Wiener process is well-defined, and it is non-anticipative with respect to the Wiener process. A weak solution $\xi(t)$ must be well-defined on at least one probability space and in some sense the corresponding Wiener process is non-anticipative with respect to $\xi(t)$.

For strong solutions the "past" $\mathcal{P}_t^w$ of a Wiener process is usually taken as $\mathcal{B}_t$ and it turns out that $\mathcal{P}_t^w = \mathcal{P}_t^\xi$. On the other hand, for weak solutions it is often the case that $\mathcal{B}_t = \mathcal{P}_t^\xi$ where $\mathcal{P}_t^\xi$ is the "past" of $\xi(t)$. Thus, in both cases one may suppose that $w(t)$ is adapted to $\mathcal{P}_t^\xi$.

A strong solution is said to be *strongly unique* if any two strong solutions coincide a.s. A weak solution is called *weakly unique* if for any two weak solutions the measures corresponding to them on the path space coincide (see Section 6.1.1).

From Definitions 6.17, 6.23 and 6.24 it follows that the solutions of (6.16) are diffusion processes. It also follows that the strong solutions of (6.16) are Markov processes. The solutions of (6.19) are evidently diffusion type processes (see Definition 6.16) and, generally speaking, they are not Markov processes.

Definition 6.25. The coefficients of (6.16) are said to satisfy the *Itô condition* (have linear growth) if there exists a constant $K > 0$ such that for all $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$ the inequality

$$\|a(t,x)\| + \|A(t,x)\| < K(1 + \|x\|), \tag{6.21}$$

is satisfied where $\|A\|$ is the operator norm of A.

The Itô condition for a Stratonovich equation has the same form as for an Itô equation. For a diffusion type equation it takes the form

$$\|a(t,x(\cdot))\| + \|A(t,x(\cdot))\| < K(1 + \|x(\cdot)\|_{C^0([0,l],\mathbb{R}^n)}). \tag{6.22}$$

Existence theorems for local solutions of SDEs assert that a solution exists up to the (random) hitting time of the boundary of a neighborhood of the initial value. It is clear that for local existence, conditions of type (6.21) are not required.

Conditions (6.21) or (6.22) guarantee the global in time existence of SDE solutions (if local solutions exist). We discuss some generalizations of Condition (6.21) in Section 7.3 and conditions of another type in Section 7.4.

Existence theorems for strong solutions are mainly proved by the contraction mapping principle, e.g., if (6.21) is satisfied and the coefficients are in some sense Lipschitz continuous (in this case the solution is well-defined for $t \in [0,\infty)$). Theorems of this sort exist for a broad class of equations with random coefficients (see, e.g., [83, 84, 162]). For example, for equation (6.19), the existence of a strong solution of the Cauchy problem for $t \in [0,\infty)$ is proved if the coefficients are Lipschitz continuous or smooth and (6.22) is satisfied. There are existence of strong solution theorems based on some other principles, and theorems which derive the existence of a strong solution from that of a weak solution. Note that there are examples of equations that have weak solutions but no strong solutions (see, e.g., [84]).

The notion of weak solution is due to A. Skorokhod. He also proved the following classical existence theorem for weak solutions of equation (6.19) in $\mathbb{R}^n$ with an n-dimensional Wiener process (see, e.g., [83, 84, 162]):

Theorem 6.26 *Let the coefficients $a(t,x(\cdot))$ and $A(t,x(\cdot))$ in (6.20) satisfy Conditions 6.19 and 6.22 and the Itô condition (6.22). Then for every deterministic initial condition $\xi(0)=x_0\in\mathbb{R}^n$ equation (6.20) has a weak solution.*

Theorem 6.26 is proved in [83, Theorem III.2.4]. Note that in some sense this theorem is a natural analog of the existence of solution theorem for ordinary differential equations with continuous right-hand sides.

The proof of Theorem 6.26 is based on the following technical statements which we shall make use of later in the book.

Lemma 6.27 *For a solution of the diffusion type stochastic differential equation $\xi(t)=\xi_0+\int\limits_0^t a(s,\xi(\cdot))\mathrm{d}s+\int\limits_0^t A(s,\xi(\cdot))\mathrm{d}w(s)$ in $\mathbb{R}^n$, $t\in[0,T]$, whose coefficients satisfy (6.22) for some $K>0$, for any integer $p\geq 2$ there exists a constant $C_p>0$, depending only on p, K and T, such that the inequality $E(\sup\limits_{t\leq T}\|\xi(t)\|^p)<C_p$ holds.*

Lemma 6.27 follows from [83, Lemma III.2.1] and the remark after it.

Lemma 6.28 *Let $\{\mu_\xi\}$ denote the measures on the path space $(\tilde{\Omega},\tilde{\mathcal{F}})$ (in the notation from Section 6.1.1) corresponding to the solutions of the equations (6.20) with various $a(t,x(\cdot))$ and $A(t,x(\cdot))$ that satisfy Conditions 6.19 and 6.22 and the Itô condition (6.22) with the same K. Then $\{\mu_\xi\}$ is weakly compact.*

Lemma 6.28 is a Corollary to [83, Lemma III.2.2].

Consider a sequence of diffusion type equations with continuous coefficients $\mathrm{d}\xi_k(t)=a_k(t,\xi(\cdot))\mathrm{d}t+A_k(t,\xi(\cdot))\mathrm{d}w(t)$ satisfying the hypothesis of Lemma 6.27 with the same K and T for all k. Let there exist (weak) solutions ξ_k of the above equations and let μ_k be the corresponding measures on $(\tilde{\Omega},\tilde{\mathcal{F}})$.

Lemma 6.29 *Suppose the measures μ_k weakly converge to a measure μ. Then for every integer $p\geq 1$ the measures ν_k defined by the relations $\mathrm{d}\nu_k=(1+\|x(\cdot)\|^p)\mathrm{d}\mu_k$ weakly converge to the measure ν defined by the relation $\mathrm{d}\nu=(1+\|x(\cdot)\|^p)\mathrm{d}\mu$.*

Proof. Let $f:\tilde{\Omega}\to\mathbb{R}$ be an arbitrary bounded continuous function. The random elements $f(\xi_k(\cdot))(1+\|\xi_k(\cdot)\|^p)$ are uniformly integrable. This follows from the facts that $f(x(\cdot))$ is bounded, that by Lemma 6.27

$$\int\limits_{\tilde{\Omega}} \|x(\cdot)\|^{p+1}\mathrm{d}\mu_k \leq \sup_t E\left(\|\xi_k(t)\|^{p+1}\right) < C_{p+1}$$

and that

$$\int\limits_{\|x_i(\cdot)\|>c} \|x_i(\cdot)\|^p \mathrm{d}\mu_i < \frac{1}{c} \int\limits_{\|x_i(\cdot)\|>c} \|x_i(\cdot)\|^{p+1}\mathrm{d}\mu_i$$

(see [25]). Then, since $f(x(\cdot))(1 + \|x(\cdot)\|^p)$ is a continuous map from $\tilde{\Omega}$ to $\mathbb{R}$, the weak convergence of μ_k to μ yields $E(f(\xi_k)(1+\|\xi_k\|)) \to E(f(\xi)(1+\|\xi\|))$ as $k \to \infty$ (see [25]). Thus $\lim\limits_{k\to\infty} \int_\Omega f(x(\cdot))(1 + \|x(\cdot)\|^p)\mathrm{d}\mu_k = \int_\Omega f(x(\cdot))(1 + \|x(\cdot)\|^p)\mathrm{d}\mu$ and so $\lim\limits_{k\to\infty} \int_\Omega f(x(\cdot))\mathrm{d}\nu_k = \int_\Omega f(x(\cdot))\mathrm{d}\nu$. □

There are existence of weak solution theorems that require the linear operator $A(t, x)$ to be non-degenerate at all (t, x) and which therefore have no analogs for ordinary differential equations. For example, if $A(t, x)$ is non-degenerate and continuous, $a(t, x)$ is measurable, and they satisfy (6.21), then for every initial condition there exists a weak solution of (6.16) well-defined for $t \in [0, \infty)$ [83, Theorem III.3.3].

N.V. Krylov's theorem [165, Theorem II.6.1] proves the existence of a weak solution for the diffusion equation (6.16) in the case when both coefficients are only measurable but uniformly bounded and $A(t, x)$ is positive definite and satisfies a qualified non-degeneracy condition.

We refer the reader to, say, [162], which contains a detailed survey of existence theorems for strong and weak solutions of SDEs.

If the coefficient A is smooth, a solution of an equation in Itô form is also a solution of an equation in Stratonovich form with different drift and vice versa. Indeed, apply the Itô formula (6.10) to $A(t, \xi(t))$ as to a mapping. Then

$$\mathrm{d}A(t, \xi(t)) = \frac{\partial A}{\partial t}\mathrm{d}t + A'(a(t, \xi(t)))\mathrm{d}t + \frac{1}{2}\mathrm{tr}A''(A, A)\mathrm{d}t + A'(A(t, \xi(t)))\mathrm{d}w(t).$$

Now substitute this expression for $\mathrm{d}A$ into the integral $\int_0^t \mathrm{d}A\mathrm{d}w(\tau)$. By Theorem 6.12 only one summand in the expression for $\int_0^t \mathrm{d}A\mathrm{d}w(\tau)$ is not equal to zero, so that $\int_0^t \mathrm{d}A\mathrm{d}w(\tau) = \int_0^t \mathrm{tr}A'(A(\tau, \xi(\tau)))\mathrm{d}\tau$. Substitute this expression into (6.6). Then

$$\int_0^t A \circ \mathrm{d}w(\tau) = \frac{1}{2}\int_0^t \mathrm{tr}A'(A(\tau, \xi(\tau)))\mathrm{d}\tau + \int_0^t A(\tau, \xi(\tau))\mathrm{d}w(\tau), \tag{6.23}$$

i.e., a solution $\xi(t)$ of the equation in Itô form (6.16) satisfies the equation in Stratonovich form

$$\mathrm{d}\xi(t) = \left[a(t, \xi(t)) - \frac{1}{2}\mathrm{tr}A'(A(t, \xi(t))\right]\mathrm{d}t + A(t, \xi(t)) \circ \mathrm{d}w(t). \tag{6.24}$$

On the other hand, if $\xi(t)$ is a solution of the equation in Stratonovich form (6.17), it satisfies the equation in Itô form

$$\mathrm{d}\xi(t) = \left[a(t,\xi(t)) + \frac{1}{2}\mathrm{tr}\, A'(A(t,\xi(t))\right]\mathrm{d}t + A(t,\xi(t))\mathrm{d}w(t). \tag{6.25}$$

Note that in equivalent equations in Itô and Stratonovich forms the diffusions are the same while the drifts are different.

So, if the coefficients of an equation in Stratonovich form are smooth enough, by using formula (6.24) one can derive existence theorems for such equations from those for equations in Itô form. If the coefficients are not smooth enough, independent proofs are required. Existence of solution problem for equations in Stratonovich form have not been as deeply investigated as for equations in Itô form.

By analogy with formula (6.24) (now by the use of formula (6.7)) one can show that a solution $\xi(t)$ of the equation in Itô form (6.16) satisfies the following equation with anticipating integral:

$$\begin{aligned}\xi(t) = \xi_0 &+ \int_0^t a(s,\xi(s))\mathrm{d}s - \int_0^t \mathrm{tr} A'(A(s,\xi(s)))\mathrm{d}s \\ &+ \int_0^t A(s,\xi(s))\mathrm{d}_* w(s),\end{aligned} \tag{6.26}$$

There are various methods of approximating the Wiener process by processes with smooth or piecewise smooth sample paths. Consequently a stochastic differential equation is approximated by ordinary differential equations with coefficients depending on a parameter $\omega \in \Omega$. It turns out that the solutions of these approximating equations tend to a solution of an equation with initial coefficients in Stratonovich form rather than in Itô form.

Let q be a partition of the interval $[0,l]$. Consider the piecewise linear approximations of the Wiener process $w(t)$ in the form $w_q(t) = \frac{1}{t_{i+1}-t_i}((t_{i+1} - t)w(t_i) + (t-t_i)w(t_{i+1})),\;\; t_i \le t \le t_{i+1}$, and the ordinary differential equations

$$\frac{\mathrm{d}x^q(t,\omega)}{\mathrm{d}t} = a(t,x^q(t,\omega)) + A(t,x^q(t,\omega))\frac{\mathrm{d}w_q}{\mathrm{d}t},\;\; \omega \in \Omega. \tag{6.27}$$

Note that by replacing $w(t)$ by $w_q(t)$ we obtain the same equation (6.27) both from the Itô and from the Stratonovich equations with equal coefficients. The classical *Wong-Zakai theorem* asserts that under some conditions that guarantee the existence of solutions on the entire interval $[0,l]$, the solutions of (6.27) with initial conditions $x^q_{m_0}(0,\omega) = m_0(\omega)$ tend with uniform probability on $[0,l]$ to a solution of the Stratonovich equation (6.18) (not of the Itô equation (6.16)) with initial condition $x_{m_0}(0,\omega) = m_0(\omega)$ as $\mathrm{diam}\, q \to 0$. It is clear that one can choose a sequence of partitions q_i such that $x^{q_i}_{m_0}$ tends to ξ_{m_0} a.s. uniformly on $[0,l]$.

Several versions of this statement, due to McShane, can be found in [66]. We should also mention a theorem of P. Malliavin (see, e.g., [177]) where a Wiener process is approximated by a process with averaged paths and the solutions of the corresponding ordinary differential equations converge to a solution of a Stratonovich SDE.

6.3 Stochastic Flows and their Generators

By analogy with the case of ordinary differential equations (see Section 1.1) the general solution of a stochastic differential equation with smooth coefficients is called a *stochastic evolution family* or, in the autonomous case, a *stochastic flow*. For simplicity of presentation we shall use the term *stochastic flow* for general solutions in both autonomous and non-autonomous cases.

Denote by $\xi(t, s)$, $s \geq t \geq 0$ the flow generated by a stochastic differential equation with smooth coefficients. This equation can be given in Itô or in Stratonovich form. Since the coefficients are smooth, one can pass from Itô to Stratonovich form and vice versa (see formulae (6.24) and (6.25)). For $x \in \mathbb{R}^n$ and $t \geq 0$ the Markov diffusion process $\xi_{t,x}(s)$ $(s \geq t)$, the solution of the above-mentioned equation with initial condition $\xi_{t,x}(t) = x$, is called the *orbit* of the flow $\xi(t, s)$. Generally speaking, x can be a random variable with values in $\mathbb{R}^n$, but if the contrary is not stated, we shall suppose x to be a non-random point in $\mathbb{R}^n$.

Denote by $(\Omega, \mathcal{F}, \mathsf{P})$ the probability space on which the solutions of the above-mentioned equation are given.

In general, it is not assumed that the orbit $\xi_{t,x}(s)$ exists for all $s \geq t$.

Definition 6.30. If for all $t \in [0, +\infty)$ and $x \in \mathbb{R}^n$ (or, in the general case, $x \in M$, where M is a smooth manifold) the orbits exist a.s. for all $s \in [t, +\infty)$, the flow is said to be *complete*.

For an arbitrary $\omega \in \Omega$ the corresponding sample path $\xi_{t_1,x_1}(s)_\omega$ of an orbit $\xi_{t_1,x_1}(s)$ may exist for all $s \geq t$, yet $\xi_{t_2,x_2}(s)_\omega$ may not exist for another orbit $\xi_{t_2,x_2}(s)$.

Definition 6.31. Denote by $\breve{\Omega}$ the set of those $\omega \in \Omega$ for which, for all $t \in [0, +\infty)$ and $x \in \mathbb{R}^n$ (or, in the general case, $x \in M$ where M is a smooth manifold), the sample path $\xi_{t,x}(s)_\omega$ exists for $s \in [t, +\infty)$. The flow is said to be *strongly* (or *strictly*) *complete* if $\mathsf{P}(\breve{\Omega}) = 1$.

Some completeness criteria for the general case of stochastic flows on manifolds will be considered in Section7.4 below.

In the general case the orbit exists on a random time interval.

Definition 6.32. Let $[0, \tau(\omega))$ be the maximal random time interval on which a solution of a stochastic differential equation (in particular, an orbit of flow) exists. The random time $\tau(\omega)$ is called the *explosion time*.

An important role in the investigation of stochastic flows is played by the so-called *infinitesimal generators* of flows (for short we shall often refer to infinitesimal generators simply as *generators*). Let the equation describing the flow be given in Itô form: (6.16). We introduce the following notation: $(q^1, \ldots, q^n)$ denote coordinates in $\mathbb{R}^n$ and A^* denotes the matrix of the operator adjoint to A. Take a relatively compact open neighborhood $U \subset \mathbb{R}^n$ of $x \in \mathbb{R}^n$ and let τ be the Markov time that the process $\xi_{t,x}(s)$ first hits the boundary of U. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real-valued function that has continuous bounded derivatives up to the second order.

Theorem 6.33 (see, e.g., [76]) *The equality*

$$\lim_{\Delta t \to +0} \frac{1}{\Delta t}(E(f(\xi_{t,x}((t+\Delta t) \wedge \tau))) - f(x)) = \mathcal{A}(t,x)f \tag{6.28}$$

holds where $(t + \Delta t) \wedge \tau = \min((t + \Delta t), \tau)$, $\Delta t \to +0$ *means that* $\Delta t \to 0$ *and* $\Delta t > 0$, *the differential operator* $\mathcal{A}(t,x)$ *is given by the relation*

$$\mathcal{A}(t,x)f = \frac{1}{2}\sum_{i,j} \sigma^{ij}(t,x)\frac{\partial^2 f}{\partial q^i \partial q^j} + a(t,x)f, \tag{6.29}$$

the matrix $(\sigma^{ij}) = A \circ A^*$ *is the diffusion coefficient of the equation, and* $a(t,x)f$ *denotes the derivative of* f *in the direction of the vector field* $a(t,x)$.

Proof. It is known that $\int_t^{(t+\Delta t)\wedge\tau} A(s, \xi_{t,x}(s))\mathrm{d}w(s)$ is a martingale. Since for $\Delta t = 0$ this integral equals zero, $E(\int_t^{(t+\Delta t)\wedge\tau} A(s, \xi_{t,x}(s))\mathrm{d}w(s)) = 0$ (see Section 6.1.4). Taking into account this equality, one can derive from the Itô formula that

$$\begin{aligned} &E(f(\xi_{t,x}(t+\Delta t) \wedge \tau))) - f(x) \\ &= E\left(\int_t^{(t+\Delta t)\wedge\tau} f'(a(s, \xi_{t,x}(s)))\mathrm{d}s \right. \\ &\quad \left. + \frac{1}{2}\int_t^{(t+\Delta t)\wedge\tau} \operatorname{tr} f''(A(s,\xi_{t,x}(s)), A(s,\xi_{t,x}(s))\mathrm{d}s\right). \end{aligned}$$

Direct calculations in coordinates show that $f'(a)$ is the derivative $a(t,x)f$ of the function f in the direction of the vector field $a(t,x)$ and that $\operatorname{tr} f''(A,A) = \sum_{i,j} \sigma^{ij}\frac{\partial^2 f}{\partial q^i \partial q^j}$. So, the latter equality can be presented in the form:

$$E(f(\xi_{t,x}((t+\Delta t) \wedge \tau))) - f(x) = E\left(\int_t^{(t+\Delta t)\wedge\tau} (\mathcal{A}f)(s, \xi_{t,x}(s))\mathrm{d}s\right).$$

Since for $s \in [t, (t+\Delta t)\wedge\tau)$ the process $\xi_{t,x}(s)$ belongs to the compact set U and since f is smooth, the values of f and of its derivatives are bounded. Hence we can apply Lebesgue's theorem on limits under integrals (expectations) and so there exists a random value $s' \in [t, (t+\Delta t)\wedge\tau]$ such that:

$$\int_t^{(t+\Delta t)\wedge\tau} (\mathcal{A}f)(s, \xi_{t,x}(s)))\mathrm{d}s = (\mathcal{A}f)(s', \xi_{t,x}(s'))((t+\Delta t)\wedge\tau - t).$$

Evidently $(t+\Delta t)\wedge\tau - t = ((t+\Delta t)-t)\wedge(\tau - t) = \Delta t \wedge (\tau - t)$ and

$$\lim_{\Delta t\to+0}\frac{1}{\Delta t}E((\mathcal{A}f)(s', \xi_x(s'))((t+\Delta t)\wedge\tau - t))$$
$$= E\lim_{\Delta t\to+0}((\mathcal{A}f)(s', \xi_x(s'))\frac{\Delta t\wedge(\tau-t)}{\Delta t}. \tag{6.30}$$

Since at time t the process under consideration has value x, i.e., t is not the first time that the sample path hits the boundary of the compact set U, we see that a.s. $\tau - t > 0$. In addition the expression $\tau - t$ does not depend on Δt. Hence $\lim\limits_{\Delta t\to+0}\frac{\tau-t}{\Delta t} = +\infty$. From the last expression it follows that

$$\lim_{\Delta t\to+0}\frac{1}{\Delta t}(\Delta t\wedge(\tau-t)) = 1\wedge\lim_{\Delta t\to+0}\frac{\tau-t}{\Delta t} = 1.$$

Since $s' \in [t, (t+\Delta t)\wedge\tau]$, by Lebesgue's theorem $s' \to t$ as $\Delta t \to +0$. Passing to the limit in (6.30) completes the proof. □

Definition 6.34. The operator $\mathcal{A}$ defined by (6.29) is called the *infinitesimal generator* (or simply *generator*) of the flow ξ.

If a and A satisfy some regularity conditions, the generator $\mathcal{A}$ determines the diffusion process uniquely: any two processes having the same generator and having the same initial value induce the same measure on the path space (see Section 6.1.1). This means the process is weakly unique, see Section 6.2.3.

It is easy to see that the generator of the Wiener process is $\frac{1}{2}\Delta$ where Δ is the Laplace operator in $\mathbb{R}^n$.

We should point out the following classical result of A.V. Skorokhod:

Theorem 6.35 (see, e.g., [81]) *Let the coefficients of equation* (6.16) *be continuous and bounded together with their first and second derivatives in x on the product $\mathbb{R}\times\mathbb{R}^n$. Suppose that the function $f : \mathbb{R}^n \to \mathbb{R}$ has continuous partial derivatives up to the second order and that it, and its first and second derivatives, is bounded on $\mathbb{R}^n$. Choose an arbitrary interval $[0, l] \subset \mathbb{R}$. Then the function $v(t, x) = E(\xi_{t,x}(l))$ on $[0, l]\times\mathbb{R}^n$ is C^1-smooth in t and C^2-smooth in x and it satisfies the equation $\frac{\partial v}{\partial t} + \mathcal{A}v = 0$.*

Thus, solutions of stochastic differential equations play for parabolic equations a role similar to that of characteristics for first order partial differential equations.

By analogy with the case of ordinary differential equations (see Section 1.1 and Section 3.1), the mapping $\xi_{t,(\cdot)}(s)$ is a *random diffeomorphism* (*random homeomorphism* – in the case of continuous coefficients of the equation) of $\mathbb{R}^n$ onto its image. If the image coincides with $\mathbb{R}^n$, $\xi(t,s)$ is said to be a *stochastic flow of diffeomorphisms* (*homeomorphisms*, respectively).

Note that unlike the case of ordinary flows (see Section 1.1), without additional constructions the evolution (or group) property makes no sense for stochastic flows since for such properties the flow must be defined for all $t, s \in (-\infty, \infty)$. Thus, we need to define a solution $\xi_{t,x}(s)$ for $s \le t$ and prove the solution's existence on the entire real line.

If $\xi(t,s)$ is a flow of diffeomorphisms, there exists a backward flow $\hat{\xi}(s,t)$ consisting of diffeomorphisms inverse to those in the flow $\xi(t,s)$.

Theorem 6.36 [167] *If a flow of diffeomorphisms $\xi(t,s)$ is given by equation (6.16), the backward flow $\hat{\xi}(s,t)$ is given by the Itô equation of the form*

$$\mathrm{d}\hat{\xi}(t) = [-a(t,\xi(t)) + \mathrm{tr}\, A'(A(t,\xi(t))]\mathrm{d}t - A(t,\xi(t))\mathrm{d}w(t). \tag{6.31}$$

If both forward and backward flows are strictly complete, it is not hard to see that having defined for $s \le t$ a value of $\xi(t,s)$ equal to $\hat{\xi}(s,t)$, we obtain a family of random diffeomorphisms for which the evolution property $\xi_{s_1,\xi_{t,x}(s_1)}(s_2) = \xi_{t,x}(s_1+s_2)$ holds for all $t, s_1, s_2 \in (-\infty, +\infty)$.Note that equation (6.31) is well-defined even if $\xi(t,s)$ is not a flow of diffeomorphisms.

Definition 6.37. If the flow $\xi(t,s)$ is not a flow of diffeomorphisms, the flow generated by equation (6.31) is called the *backward flow* to $\xi(t,s)$.

Definition 6.38. The generator $\hat{\mathcal{A}}$ of the backward flow generated by (6.31) is called the *backward generator* of the flow $\xi(t,s)$.

One can easily see that $\hat{\mathcal{A}}$ is given by the relation

$$\hat{\mathcal{A}}(t,x)f = \frac{1}{2}\sum_{i,j}\sigma^{ij}(t,x)\frac{\partial^2 f}{\partial q^i \partial q^j} - a(t,x)f + \mathrm{tr}\, A'(A(t,x))f, \tag{6.32}$$

where $\mathrm{tr}\, A'(A(t,x))f$ is the derivative of f along the vector field $\mathrm{tr}\, A'(A(t,x))$.

If a flow is given by an equation in Stratonovich form (6.18), it is not hard to derive from formula (6.25) that in terms of the coefficients of equation (6.18) the generators $\mathcal{A}$ and $\hat{\mathcal{A}}$ have the following presentations:

$$\mathcal{A}(t,x) = \frac{1}{2}\sum_{i,j}\sigma^{ij}(t,x)\frac{\partial^2}{\partial q^i \partial q^j} + (a(t,x) + \frac{1}{2}\mathrm{tr}\, A'(A(t,x)), \tag{6.33}$$

$$\hat{\mathcal{A}}(t,x) = \frac{1}{2}\sum_{i,j}\sigma^{ij}(t,x)\frac{\partial^2}{\partial q^i \partial q^j} + (-a(t,x) + \frac{1}{2}\mathrm{tr}\, A'(A(t,x)). \tag{6.34}$$

We refer the reader to, e.g., [1, 167] for a more detailed discussion of stochastic flows.

Chapter 7
Stochastic Analysis on Manifolds

The purpose of this chapter is to describe and investigate the main features of stochastic analysis on smooth manifolds. Our principal focus shall be on stochastic differential equations. A monographic presentation of various alternative aspects of and approaches to stochastic analysis on manifolds can be found in [23, 66, 69, 147, 179, 180, 190, 205].

7.1 Stochastic Differential Equations in Stratonovich Form on a Manifold

7.1.1 General construction

The theory of stochastic differential equations in Stratonovich form can be generalized to the case of manifolds in a very simple way. The reason for this is that by formula (6.14) a solution of an equation in Stratonovich form (6.18) under the coordinate change $\varphi_{\beta\alpha}$ transforms into

$$\mathrm{d}\varphi_{\beta\alpha}(\xi(t)) = \varphi'_{\beta\alpha}[a(t,\xi(t)) + A(t,\xi(t)) \circ \mathrm{d}w(t)] \tag{7.1}$$

which coincides with formula (1.1) for the transformation of a tangent vector. Equations in Itô form are transformed by formula (7.6), i.e., they are cross-sections of a special bundle, and so they require special constructions for their description. This is why Stratonovich's approach to stochastic differential equations on manifolds is used so extensively in the literature.

In this section we describe an approach to stochastic differential equations on manifolds which is based on that of Stratonovich. Itô's approach will be described in later sections.

Let M be a smooth finite dimensional manifold with dimension n.

Definition 7.1. The pair $(a(t,m), A(t,m))$, where $a(t,m)$ is a vector field on M and $A(t,m)$ is a field of linear operators $A(t,m) : \mathbb{R}^k \to T_mM$ sending

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_7, © Springer-Verlag London Limited 2011

a certain Euclidean space $\mathbb{R}^k$ to the tangent spaces to M, is called an *Itô vector field*.

For Itô vector fields we often use the notation (a, A). The value at the point m is denoted by (a_m, A_m) or $(a(m), A(m))$ (in the non-autonomous case the notation $(a(t, m), A(t, m))$ is also used).

Since the space $\mathbb{R}^k$ is specified (i.e., it is not subjected to coordinate changes), one can easily see that under a change of coordinates $\varphi_{\beta\alpha}$ in a chart on M the coordinate descriptions $A^\alpha(t, m)$ and $A^\beta(t, m)$ are connected by formula $A^\beta(t, m) = \varphi'_{\beta\alpha} A^\alpha(t, m)$, i.e., under changes of coordinates both components of the pair $(a(t, m), A(t, m))$ are transformed by formula (1.1). This is the reason for using the term "Itô vector field".

Lemma 7.2 *Let $A(t, m)$ be as above. Then AA^* is a symmetric $(2, 0)$-tensor field on M.*

Proof. If the matrix of the operator $A(t, m)$ is calculated with respect to the standard basis in $\mathbb{R}^k$ and the basis in T_mM is generated by the coordinate system of some chart, the transposed matrix $A^*(t, m)$ is the matrix of the conjugate operator calculated with respect to the dual basis in the cotangent space T^*_mM and the standard basis in $\mathbb{R}^k$ (we identify $\mathbb{R}^k$ with its conjugate space in terms of the standard inner product). Note that $A(t, m)A^*(t, m)$ sends T^*_mM to T_mM, i.e., a pair of cotangent vectors is transformed into a pair comprising a tangent vector and a cotangent vector whose "coupling" yields a real number that linearly depends on both cotangent vectors from the initial pair. Thus $A(t, m)A^*(t, m)$ is a bilinear form on cotangent vectors, i.e., it is a $(2, 0)$-tensor field. The fact that the matrix of AA^* is symmetric in any coordinate system is obvious. □

Let $w(t)$ be a Wiener process in $\mathbb{R}^k$ and $(a(t, m), A(t, m))$ be an Itô vector field on M. The expression

$$d\xi(t) = a(t, \xi(t))dt + A(t, \xi(t)) \circ dw(t) \tag{7.2}$$

is called a *stochastic differential equation in Stratonovich form on M, given by Itô vector field* $(a(t, m), A(t, m))$. This means that in every chart on M the solution $\xi(t)$ satisfies equation (6.17). As said above, (7.2) has the correct transformation rule under changes of coordinates, i.e., it is well-defined.

Remark 7.3. Let M be embedded into a Euclidean space $\mathbb{R}^N$ as a submanifold. Since the Itô vector field (a, A) on M is transformed as a vector under changes of coordinates, the phrase "*an Itô vector field is tangent to the submanifold M in $\mathbb{R}^N$*" is well-defined. This means that $a(t, m) \in T_mM \subset T_m\mathbb{R}^N$ and $A(t, m) : \mathbb{R}^k \to T_mM \subset T_m\mathbb{R}^N$ at every point $m \in M$; notice that these relations remain true under changes of coordinates. For Itô equations the property of being tangent to M is ill-defined.

We consider some examples of stochastic differential equations in Stratonovich form on manifolds.

Example 7.4. Embed the manifold M by Whitney's Theorem (Theorem 1.2) into a Euclidean space $\mathbb{R}^k$ for some sufficiently large k. Denote by $P_m : \mathbb{R}^k \to T_mM$ the operator of orthogonal projection. Let $a(t,m)$ be a vector field and $B(t,m)$ be a $(1,1)$-tensor field on M. Define the field of linear operators $\mathbf{A}(t,m) : \mathbb{R}^k \to T_mM$ by the formula $\mathbf{A}(t,m) = B(t,m) \circ P_m$. It is clear that $(a(t,m), \mathbf{A}(t,m))$ is an Itô vector field on M that generates the equation

$$\mathrm{d}\xi(t) = a(t,\xi(t))\mathrm{d}t + \mathbf{A}(t,\xi(t)) \circ \mathrm{d}w(t). \tag{7.3}$$

Example 7.5. Let M be a Riemannian manifold with $\dim M = n$. In this case we set $k = n$. Consider the total space OM of the bundle of orthonormal frames over M and the Levi-Civitá connection H on OM (see Section 2.7). Let $a(t,m)$ be a vector field and $B(t,m)$ be a $(1,1)$-tensor field on M. Consider their horizontal lifts $a^T(t,b)$ and $B^T(t,b)$ onto OM. Recall that $a^T(t,b) = T\pi^{-1}(a(t,\pi b)_{|\mathsf{H}_b}$ and $B^T(t,b)$ is defined analogously. Consider also the basic vector field $\mathsf{E}(w(t))$ on OM (see Definition 2.68), constructed from the vector $w(t)$ in $\mathbb{R}^n$. The pair $(a^T(t,b), B^T(t,b) \circ \mathsf{E}_b)$ is an Itô vector field on OM and together with $w(t)$ from $\mathbb{R}^n$ it determines the equation

$$\mathrm{d}\xi(t) = a^T(t,\xi(t))\mathrm{d}t + B^T(t,\xi(t)) \circ \mathsf{E}_{\xi(t)}(\circ\, \mathrm{d}w(t)). \tag{7.4}$$

on OM.

Example 7.6. We present an example of a stochastic differential equation with random coefficients. Let the following objects be given: a non-decreasing family of σ-subalgebras $\mathcal{B}_t$ of the σ-algebra $\mathcal{F}$ ($t \in [0,l]$, $l > 0$) to which a Wiener process $w(t)$ in $\mathbb{R}^k$ is adapted; a stochastic process $a(t)$ in $\mathbb{R}^n$ and a stochastic process $A(t)$ with values in the space of linear mappings from $\mathbb{R}^k$ to $\mathbb{R}^n$ that are non-anticipating with respect to $\mathcal{B}_t$ and having a.s. continuous sample trajectories; the field of linear operators $E_m : \mathbb{R}^n \to T_mM$, smooth in $m \in M$. The pair $(E_m a(t), E_m A(t))$ on M is a random Itô vector field and it generates the stochastic differential equation

$$\mathrm{d}\xi(t) = E_{\xi(t)}a(t)\mathrm{d}t + E_{\xi(t)}A(t) \circ \mathrm{d}w(t)).$$

The notions of strong and weak solutions continue to be meaningful for stochastic differential equations on manifolds.

Theorem 7.7 *Let $F : M \to N$ be a smooth mappings of manifolds and $\xi(t)$ be a solution of equation (7.2) on M. Then $F(\xi(t))$ satisfies the equation*

$$\mathrm{d}F(\xi(t)) = TFa(t,\xi(t))\mathrm{d}t + TFA(t,\xi(t)) \circ \mathrm{d}w(t),$$

on N where TF is the tangent mapping of F.

Theorem 7.7 follows from formula (6.14).

Local existence of solution theorems for equation (7.2) with smooth enough coefficients (a, A) can be proved via the following argument. Specify an initial condition $\xi(0) = m_0 \in M$ and consider a relatively compact chart $\mathfrak{U}_\alpha \ni m_0$. Since $A(t, m)$ is smooth enough, in this chart it is possible to pass from equation (6.17) (whose solution in the chart is $\xi(t)$ by definition) to equation (6.25). If the coefficients (a, A) are smooth enough, applying the existence theorems for linear spaces from Section 6.2.3 (applicable in every chart) one can easily see that the latter equation has a strong solution with given initial condition. This solution exists on a random time interval $[0, \tau_\omega)$ where τ_ω is the time that a sample trajectory $\xi_\omega(t)$ first hits the boundary of the domain $\mathfrak{U}_\alpha$ (the stopping time, see Section 6.1.3).

We describe two tricks that from the very beginning have been applied in this theory for proving the global existence of solutions (alternative techniques will be considered later). The first trick embeds M into a Euclidean space, as in Example 7.4 (but it is applicable in a much broader setting). Embed M into a Euclidean space $\mathbb{R}^N$ of sufficiently large dimension. Consider the normal bundle $N(M)$ of M in $\mathbb{R}^N$ with fiber N_m. Take a tubular neighborhood Θ of M. Recall that Θ is retracted onto M by the fibers of N_m (see [172]). Denote by $\hat{r} : \Theta \to M$ the corresponding retraction.

The neighborhood Θ is represented as a direct product $M \times W$ where W is an open ball in $\mathbb{R}^{N-n}$ that at every point $m \in M$ we can identify with a ball of the normal space N_m to M. From the presentation of Θ as $M \times W$ it follows that the tangent space $T_{(m,x)}\Theta = T_m M \times T_x W_m$. It is obvious that $T\hat{r}$ (where $\hat{r} : \Theta \to M$ is the retraction introduced above) sends the subspaces $T_x W_m$ into zero vectors in $T_m M$. Thus on the subspaces $T_m M$ the mapping $T\hat{r}$ is a linear isomorphism. Introduce vectors $a(t, (m, x)) = T\hat{r}^{-1} a(t, m)$ and linear operators $A(t, (m, x)) = T\hat{r}^{-1} A(t, m)$ at $x \in \Theta$. By Remark 7.3 it is clear that $(a(t, x), A(t, x))$ is a well-defined Itô vector field on Θ.

Let $\mathcal{O}$ be a neighborhood of M such that $\bar{\mathcal{O}} \subset \Theta$, where $\bar{\mathcal{O}}$ is the closure of $\mathcal{O}$. Let $\varphi(y) : \mathbb{R}^N \to \mathbb{R}$ be a smooth function that meets the conditions: $0 \le \varphi \le 1$, $\varphi(y) = 1$ for $x \in \bar{\mathcal{O}}$ and $\varphi(y) = 0$ for $x \notin \Theta$ (Urysohn's function). Consider the Itô vector field $(\bar{a}(t, x), \bar{A}(t, x))$ on $\mathbb{R}^N$ $(\bar{a}(t, x), \bar{A}(t, x))$ of the form

$$
\bar{a}(t,x) = \begin{cases} \varphi(x) a(t,x) & x \in \Theta, \\ 0 \quad x \notin \Theta, \end{cases}
$$

$$
\bar{A}(t,x) = \begin{cases} \varphi(x) A(t,x) & x \in \Theta, \\ 0 \quad x \notin \Theta. \end{cases}
$$

On $\mathbb{R}^N$ this vector field has the same smoothness as $(a(t, m), A(t, m))$ on M.

Specify an arbitrary initial condition $\xi(0) = m_0 \in M$.

Theorem 7.8 *Let the above-mentioned Itô vector field* $(\bar{a}(t, x), \bar{A}(t, x))$ *on* $\mathbb{R}^N$ *be smooth and* C^1*-uniformly bounded. Then equation* (7.2) *on* M *with initial condition* $\xi(0) = m_0 \in M$ *has a unique strong solution on* M *that exists for all* $t \in [0, +\infty)$.

Proof. Consider the following equation in Stratonovich form in $\mathbb{R}^N$

$$\mathrm{d}\bar{\xi}(t) = \bar{a}(t, \bar{\xi}(t))\mathrm{d}t + \bar{A}(t, \bar{\xi}(t)) \circ \mathrm{d}w(t).$$

Pass by formula (6.25) to the corresponding equation in Itô form. Since by hypothesis the coefficients $(\bar{a}(t,x), \bar{A}(t,x))$ and their first derivatives are smooth and uniformly bounded (from this it in particular follows that $\bar{a}(t,x)$, $\frac{1}{2}\mathrm{tr}\,\bar{A}'(\bar{A}(t,x))$ and $\bar{A}(t,x)$ are uniformly bounded), the corresponding equation in Itô form that we have obtained has a unique strong solution that exists for $t \in [0, +\infty)$ (see Section 6.2.3). Since by construction the Itô vector field $(\bar{a}(t,x), \bar{A}(t,x))$ at the points $x \in M$ coincides with $(a(t,x), A(t,x))$, (i.e., it is tangent to M), it is easy to show that the solution $\bar{\xi}(t)$ with initial conditions $\bar{\xi}(0) = m_0$ a.s. lies in M. □

7.1.2 Riemannian uniform atlases

The second traditional trick that allows one to prove the existence of global in time solutions of stochastic differential equations is based on the existence of a certain covering, uniform in the following sense:

(i) for every point there exists a chart that together with the point contains a ball of some specified radius, this radius being independent of the point, the chart, etc.;

(ii) on every ball as described in (i), the coefficients of the equation are bounded by a constant and that constant is independent of the ball, the point, the chart etc.

The technique of proving the existence of a solution for $t \in [0, +\infty)$ is as follows. Applying theorems of existence of strong solutions in vector spaces (see Section 6.2.3), we can prove the existence and uniqueness of a solution in the chart with center at m_0 on a random interval $[0, \tau_1]$ where $\tau_1(\omega)$ is the Markov time of the first hit of a sample trajectory $\xi(t,\omega)$ on the boundary of the above-mentioned ball with specified radius and center at m_0 in this chart (if $\xi(t,\omega)$ does not hit the boundary at all, $\tau(\omega) = \infty$). Then in an analogous manner we start a solution from $\xi(\tau_1,\omega)$ given for $[\tau_1, \tau_2]$ where $\tau_2(\omega)$ is the Markov time of the first hit to the boundary of the ball centered at $\xi(\tau_1,\omega)$ in the corresponding charts, etc. Some estimates are proved for the probability of hitting the boundary of the ball for small t, based on boundedness of the equation coefficients. Then it is derived from those estimates that a.s. $\sup\limits_n \tau_n = \infty$.

Such a trick was used for the first time by Itô for equations in Itô form in [149] (this is perhaps the first paper in the literature devoted to stochastic differential equations on manifolds). For equations in Stratonovich form we refer the reader to the paper by Clark [40] (see also [66]). The same trick was used in [23] for Itô equations in Belopolskaya-Daletskii form (see below). In

[66] the global existence of solutions for the equation from Example 7.6 was proved by this method.

Usually the radii of balls are estimated in terms of Euclidean distance in charts. This is why it is not clear whether such a covering exists on a given manifold for a given equation.

Following [95], here we present a modification of such conditions in which the radii of balls are measured with respect to a certain Riemannian metric on the manifold, i.e., we introduce a special class of Riemannian metrics. We prove an important result that such metrics exist on every finite-dimensional manifold.

Let M be a connected finite-dimensional Riemannian manifold with Riemannian metric $\langle\cdot,\cdot\rangle$ and let ρ be the Riemannian distance function on M.

Definition 7.9. An atlas on M is called a *uniform Riemannian atlas* if for every point $m \in M$ there exists a chart $(\mathcal{V}, \varphi)$, $m \in \mathcal{U} = \varphi(\mathcal{V})$, of this atlas such that $\mathcal{U}$ contains the metric ball $V_m(r)$ centered at m and having a specified radius $r > 0$ independent of m and $\mathcal{U}$, where $V_m(r)$ is taken with respect to the Riemannian distance ρ.

Note that, in general, the metric ball $V_m(r) = \{m' \in M | \rho(m, m') < r\}$ may not be homeomorphic to a ball in the model space and may have a complicated topological structure.

Obviously, a Riemannian metric possessing a uniform Riemannian atlas is complete.

Theorem 7.10 [95] *For any Riemannian metric on M, there exists a Riemannian metric conformal to it that possesses a uniform Riemannian atlas.*

To prove Theorem 7.10, we refine the methods developed in [140, 191] for the investigation of convex neighborhoods and complete Riemannian metrics. Theorem 7.10 is a generalization of the main result from [191] proving that any Riemannian metric is conformal to a complete Riemannian metric.

Pick a Riemannian metric $\langle\cdot,\cdot\rangle$ on M, i.e., let $\langle\cdot,\cdot\rangle_m$ be an inner product in the tangent space T_mM, and let ρ be the Riemannian distance on M corresponding to $\langle\cdot,\cdot\rangle$.

It is known (see [161]) that for any point $m \in M$, there exists a number $a(m) > 0$ such that the ρ-metric ball $V_m\big(a(m)\big)$ lies in a normal coordinate neighborhood (chart) of any point $m' \in V_m\big(a(m)\big)$. Let $r(m)$ be the least upper bound of such $a(m)$.

If $r(m^*) = \infty$ for a point $m^* \in M$, the proof is clear. Assume that $r(m) < \infty$ for all $m \in M$.

Lemma 7.11 *For any two points $m, m' \in M$, the following inequality holds:*

$$\big|r(m) - r(m')\big| \leq \rho(m, m'). \tag{7.5}$$

Proof. First, consider the case where $m' \in V_m(r(m))$. Then $V_{m'}(r(m) - \rho(m, m')) \subset V_m(r(m))$ and, by the definition of $r(m')$, we have $r(m') \geq r(m) - \rho(m, m')$. If $r(m) \geq r(m')$, then (7.5) follows. On the other hand, $m \in V_m(r(m'))$ if $r(m') > r(m)$, and, therefore, $r(m) \geq r(m) - \rho(m, m')$, which proves (7.5). The case where $m \in V_{m'}\big(r(m')\big)$ can be dealt with in the same manner. In the remaining case, the inequalities $r(m) \leq \rho(m, m')$ and $r(m') \leq \rho(m, m')$ follow from $m' \notin V_m(r(m))$ and, at the same time, $m \notin V_{m'}(r(m'))$. Therefore, $|r(m) - r(m')| < \rho(m, m')$, which completes the proof of the lemma. □

Without loss of generality, we may assume that the function $r(m)$ is smooth. If $r(m)$ is not smooth, it can be approximated by a smooth function $r^*(m)$ such that $0 < r^*(m) < r(m)$ and $r^*(m)$ satisfies (7.5).

Let us introduce a new metric $\langle \cdot, \cdot \rangle^*$ on M by the formula

$$\langle \cdot, \cdot \rangle^*_m = \frac{1}{r^2(m)} \langle \cdot, \cdot \rangle_m \,.$$

Denote by ρ^* the Riemannian distance on M corresponding to $\langle \cdot, \cdot \rangle^*$.

Lemma 7.12 *If $\rho(m, m') \geq r(m)$ then $\rho^*(m, m') \geq 1/2$.*

Proof. Let $\gamma(t)$ be an arbitrary piecewise smooth curve such that $\gamma(a) = m$ and $\gamma(b) = m'$. Denote its length in the metric $\langle \cdot, \cdot \rangle$ by L, i.e.,

$$L = \int_a^b \|\dot{\gamma}(t)\| \, \mathrm{d}t.$$

The length of γ in the metric $\langle \cdot, \cdot \rangle^*$ can be found by the formula

$$L^* = \int_a^b \frac{\|\dot{\gamma}(t)\|}{r\left(\gamma(t)\right)} \, \mathrm{d}t.$$

Using the classical mean value theorem, we obtain

$$L^* = \frac{1}{r\left(\gamma(\tau)\right)} \int_a^b \|\dot{\gamma}(t)\| \, \mathrm{d}t = \frac{L}{r\left(\gamma(\tau)\right)},$$

where $\tau \in [a, b]$. Then

$$L^* = \frac{L}{r(\gamma(\tau)) - r(m) + r(m)}$$

and, by Lemma 7.11,

$$L^* > \frac{L}{r(m) + \rho\big(m, \gamma(\tau)\big)}.$$

By assumption, $L > r(m)$. Moreover, $\rho\big(m, \gamma(\tau)\big)$ is not greater than the length of γ on the interval $[\,a, \tau)$, which, in turn, is not greater than L, i.e., $L \geq \rho\big(m, \gamma(\tau)\big)$. Thus

$$L^* > \frac{L}{L+L} = \frac{1}{2} \ .$$

Since (2.2) holds for an arbitrary γ, $\rho(m, m') > 1/2$ and the lemma is proved. $\square$

Proof. [of Theorem 7.10] By construction, the metric $\langle \cdot, \cdot \rangle^*$ is conformal to the original metric $\langle \cdot, \cdot \rangle$. By definition, a normal chart of the metric $\langle \cdot, \cdot \rangle$ at m contains the metric ball $V_m\big(r(m)\big)$ with respect to the distance ρ. It follows from Lemma 7.12 that $\rho(m, m') < r(m)$ when $\rho(m, m') < 1/2$. Thus, at every point $m \in M$, the normal chart of the metric $\langle \cdot, \cdot \rangle$ contains the ball centered at m and having the radius $1/2$ with respect to the metric ρ^*. Therefore, $\langle \cdot, \cdot \rangle^*$ is the desired metric. The theorem is proved. $\square$

Theorem 7.13 *Let there exist a Riemannian metric on M possessing a uniform Riemannian atlas such that a smooth Itô vector field $(a(t, m), A(t, m))$ is uniformly bounded with respect to the norm of the functional space C^1 generated by this metric. Then for every initial condition $\xi(0) = m_0 \in M$ equation (7.2) with this Itô vector field has a unique solution well-defined for all $t \in [0, +\infty)$.*

Theorem 7.13 is proved by the argument presented at the beginning of this Section, modified by replacing distances with respect to the Euclidean norm by Riemannian distances.

7.2 The Itô Bundle and Itô Equations on a Manifold

The investigation of equations in Itô form on manifolds was initiated by Itô's paper [149] and yielded interesting constructions clarifying the geometric nature of stochastic differential equations (see, e.g., [21, 23, 77]).

According to the Itô formula (6.10), under a coordinate change $\varphi_{\beta\alpha}$ a solution of the equation in Itô form (6.16) transforms into the equation

$$\begin{aligned} \mathrm{d}\varphi_{\beta\alpha}(\xi(t)) = \varphi'_{\beta\alpha}[a(t, \xi(t))\mathrm{d}t + A(t, \xi(t))\mathrm{d}w(t)] \\ + \frac{1}{2}\mathrm{tr}\, \varphi''_{\beta\alpha}(A(t, \xi(t)), A(t, \xi(t)))\mathrm{d}t, \end{aligned} \tag{7.6}$$

i.e., the solutions are cross-sections of a special fiber bundle. In order to describe this bundle precisely, according to Definition 1.32, we first describe its structure group.

Let M be a smooth manifold of dimension n. Denote by $L(\mathbb{R}^k, \mathbb{R}^n)$ the space of linear operators sending $\mathbb{R}^k$ to $\mathbb{R}^n$, by $GL(n, \mathbb{R})$ the group of invertible $n \times n$ matrices (or invertible linear operators acting on $\mathbb{R}^n$) and by

$\mathbf{L}^2(\mathbb{R}^n)$ the set of bilinear mappings $\alpha : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ (cf. the notation of Section 1.2).

Definition 7.14. The *Itô group* G_I is the set of pairs (B, β) where $B \in GL(n, \mathbb{R})$ and $\beta \in \mathbf{L}^2(\mathbb{R}^n)$, with the operation defined by the following equality:

$$(B, \beta) \cdot (C, \gamma) = (B \circ C, B \circ \gamma(\cdot, \cdot) + \beta(C(\cdot), C(\cdot))). \tag{7.7}$$

Theorem 7.15 *G_I with operation (7.7) is indeed a group.*

Proof. The associativity of (7.7) is verified by direct calculation. The unit in G_I is the pair $(I, 0)$ where I is the unit operator and 0 is the zero bilinear mapping. For a pair (B, β) the inverse element with respect to (7.7) is the pair $(B, \beta)^{-1} = (B^{-1}, -B^{-1} \circ \beta(B^{-1}(\cdot), B^{-1}(\cdot)))$. □

Remark 7.16. As for every Lie group, the tangent space $T_{(I,0)}G_I$ at the unit $(I, 0)$ of G_I has the structure of a Lie algebra. Note that $T_{(I,0)}G_I$ consists of the pairs $\{(D, \delta)\}$, where $D \in L(\mathbb{R}^n)$, the group of all linear operators in $\mathbb{R}^n$ ($n \times n$ matrices) and $\delta \in \mathbf{L}^2(\mathbb{R}^n)$. Direct calculations with left-invariant vector fields on G_I according to formula (1.7) determining the Lie bracket yields the following formula for the bracket of vectors (D, δ) and (E, ϵ) from $T_{(I,0)}G_I$:

$$[(D, \delta), (E, \epsilon)] = (DE - ED, E(\delta(I, D)) - D(\epsilon(I, E))). \tag{7.8}$$

We call the Lie algebra with bracket (7.8) the *Itô algebra*.

Recall that for a bilinear operator $\Psi(\cdot, \cdot)$ on an n-dimensional Euclidean space taking values in the same space, its trace is the vector defined by the formula

$$\mathrm{tr}\Psi = \sum_{i=1}^{n} \Psi(e_i, e_i) \tag{7.9}$$

where $e_1, \dots, e_n$ is an orthonormal frame and the trace does not depend on the choice of orthonormal frame.

Let Ψ be a bilinear operator on the tangent space T_mM of a Riemannian manifold which takes values in T_mM. Denote by Ψ^k_{ij} its coefficients with respect to the basis $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}$ in a chart and by g^{ij} the components of the metric tensor (see Notation 1.51 and Remark 1.52). Then it is easy to see that in local coordinates the trace is described by the formula

$$\mathrm{tr}\Psi = g^{ij}\Psi^k_{ij}. \tag{7.10}$$

Define the left action of group G_I on the product $\mathbb{R}^n \times L(\mathbb{R}^k, \mathbb{R}^n)$ by the formula

$$(B, \beta) \cdot (X, A) = \left(BX + \frac{1}{2}\mathrm{tr}\beta(A(\cdot), A(\cdot)), B \circ A\right). \tag{7.11}$$

Definition 7.17. The *Itô bundle* $I(M)$ over a manifold M is the bundle described in Definition 1.32 with standard fiber $\mathbb{R}^n \times L(\mathbb{R}^k, \mathbb{R}^n)$ and structure group G_I that acts on $\mathbb{R}^n \times L(\mathbb{R}^k, \mathbb{R}^n)$ from the left by formula (7.11).

It is clear that the remaining two elements from the Definition 1.32 of the bundle (i.e., the total space and projection), are determined here by the standard fiber and structure group. We emphasize that over every chart $\mathcal{U}_\alpha$ on M the Itô bundle is presented as a direct product $\mathcal{U}_\alpha \times (\mathbb{R}^n \times L(\mathbb{R}^k, \mathbb{R}^n))$ and under the change of coordinates $\varphi_{\beta\alpha}$ from the chart $\mathcal{U}_\alpha$ to another chart $\mathcal{U}_\beta$ an arbitrary point $(m^\alpha, (a^\alpha, A^\alpha))$ is transformed according to the rule

$$(m^\alpha, (a^\alpha, A^\alpha)) \mapsto \left(\varphi_{\beta\alpha} m^\alpha, \left(\varphi'_{\beta\alpha} a^\alpha + \frac{1}{2} \operatorname{tr} \varphi''_{\beta\alpha}(A^\alpha, A^\alpha), \varphi'_{\beta\alpha} A^\alpha\right)\right). \quad (7.12)$$

It is not hard to describe the principal Itô bundle, i.e., the principal bundle with G_I as the structure group. In every chart $\mathcal{U}_\alpha$ it is presented as a direct product $\mathcal{U}_\alpha \times G_I$ and under the change of coordinates $\varphi_{\beta\alpha}$ to a chart $\mathcal{U}_\beta$ it transforms according to the rule

$$(m^\alpha, (B^\alpha, \gamma^\alpha)) \mapsto \left(\varphi_{\beta\alpha} m^\alpha, \left(\varphi'_{\beta\alpha} B^\alpha, \varphi'_{\beta\alpha} \gamma^\alpha(\cdot, \cdot) + \varphi''_{\beta\alpha}\left(B^\alpha(\cdot), B^\alpha(\cdot)\right)\right)\right).$$

It is easy to see that $I(M)$ is a bundle associated with the principal Itô bundle where the left action on the fiber is given by formula (7.11).

Definition 7.18. The cross-sections of the Itô bundle $I(M)$ are called the *Itô equations.*

We introduce the notation $(\hat{a}, A)$ for an Itô equation, the value at a point m is denoted by $(\hat{a}_m, A_m)$ or $(\hat{a}(m), A(m))$ (in the non-autonomous case $(\hat{a}(t, m), A(t, m))$). This notation makes sense in every chart. Notice that the second element of the pair is well-defined as a linear operator $A_m : \mathbb{R}^k \to T_m M$ (under changes of coordinates A transforms as a linear operator of this sort, see formula (7.12)). Taking a trivialization in a chart, $\hat{a}$ can be identified with a vector from $T_m M$, but this identification depends on the choice of chart and trivialization (the transformation rule for $\hat{a}$ under changes of coordinates depends on A). Convenient coordinate and invariant descriptions of Itô equations will be given below.

Let $(\hat{a}, A)$ be an Itô equation and $w(t)$ be a Wiener process in $\mathbb{R}^k$. In a given chart $\mathcal{U}_\alpha$ we consider the following stochastic differential equation in Itô form

$$\mathrm{d}\xi(t) = \hat{a}(t, \xi(t))\mathrm{d}t + A(t, \xi(t))\mathrm{d}w(t). \quad (7.13)$$

Comparing the Itô formula (7.6) and formula (7.12), one can easily see that equation (7.13) has the correct transformation rule under a change of coordinates, i.e., (7.13) can be considered on the entire manifold M.

A solution of (7.13) is a diffusion process according to Definition 6.17 (see also Section 6.2.3) and so it is a semi-martingale since $\xi(t)$ in a chart is the sum of an Itô integral $\int_0^t A(t,\xi(t))\mathrm{d}w(t)$ (i.e., an ordinary martingale) and a Lebesgue integral $\int_0^t \hat{a}(t,\xi(t))\mathrm{d}t$ (i.e., a process with bounded variation of sample paths). Recall (see Section 6.1.4) that under coordinate changes semi-martingales are transformed into semi-martingales, i.e., the notion of a semi-martingale is well-defined on manifolds.

However the notion of a martingale is ill-defined on manifolds since under a coordinate change a martingale is transformed to a semi-martingale. A natural generalization of the notion of a martingale to the case of processes on manifolds is the notion of a martingale with respect to a connection that arises in the works of L. Schwartz and P.A. Meyer. We refer the reader to [69, 179, 180, 204, 205] where a detailed description of this material in the general case can be found. Here we introduce it only for the particular case of diffusion processes.

The construction of the generator described in Section 6.3 is valid for a solution of (7.13). Recall that a generator is a second order vector field on M (see Section 2.9). Introduce a connection H on M and consider the corresponding field of fiber-wise linear operators $\mathcal{H} : \tau M \to TM$ defined by formula (2.45).

Definition 7.19. A diffusion process with generator $\mathcal{A}$ is called a *martingale with respect to a connection* H if $\mathcal{H}\mathcal{A} = 0$.

For specialists we mention that in the general case a semi-martingale is called a martingale with respect to a connection H if the mapping $\mathcal{H}$ sends its quadratic characteristic to zero.

The notions of strong and weak solutions is naturally transferred to stochastic equations (7.13) as well as to equations (7.2). The existence of local solution theorems (in charts) also remain valid. Let us present a theorem of existence of a global solution.

Use the natural trivialization of $I(M)$ in every chart and for an Itô equation $(\hat{a}, A)$ determine in this trivialization the norm $\|\hat{a}_m\|$ as the Riemannian norm of the vector in T_mM that corresponds to the first element of the pair $(\hat{a}, A)$ with respect to the given trivialization, and the norm $\|A_m\|$ as the norm of the linear operator A sending $\mathbb{R}^n$ to T_mM where the latter is equipped with the Riemannian inner product. Recall that a Riemannian metric possessing a uniform Riemannian atlas exists on every finite-dimensional manifold (see Section 7.1.2).

Theorem 7.20 *Let an Itô equation $(\hat{a}(t,m), A(t,m))$ be smooth in $m \in M$ and continuous in $t \in [0,\infty)$. Let on M there exist a Riemannian metric possessing a uniform Riemannian atlas and let in the charts of that atlas, in every ball $V_m(r)$, the estimates $\|\hat{a}(t,m')\| < C$ and $\|A(t,m')\| < C$ hold for all $t \in [0,\infty)$ and $m' \in V_m(r)$ where the constant $C > 0$ does not depend on the choice of chart and ball. Then for every initial condition $\xi(0) = m_0 \in M$ there exists a unique strong solution $\xi(t)$ of equation (7.13) that is well-defined for $t \in [0,\infty)$.*

Theorem 7.20 is proved by the same argument as Theorem 7.13 (see the beginning of Section 7.1.2). It is a modification of classical existence statements for solutions of stochastic differential equations on manifolds that are proved in various settings in [23, 40, 66, 149, 177]. The difference is that we use the charts of a uniform Riemannian atlas and postulate the uniform boundedness of the equation in the balls of that atlas with respect to that metric, while in the above-mentioned papers the existence of a special atlas and boundedness with respect to Euclidean norms in the balls in the charts are required (the exact formulation varies depending on the setting).

We should mention that the proof of Theorem 7.20 is quite analogous to that of [23, Theorem 2.2]. A modification is required only for proving the estimates (Propositions 2.1 and 2.2 of [23]): one should replace Euclidean norms in the tangent space by Riemannian norms and Euclidean distances in charts by Riemannian distances. The proof of [23, Theorem 2.2] remains valid without change, taking into account the obvious statement that for every $m' \in V_m(\frac{r}{2})$ the inclusion $V_{m'}(\frac{r}{2}) \subset V_m(r)$ holds.

Remark 7.21. Theorem 7.20 is a general statement on the existence of solutions for $t \in [0, \infty)$ (completeness of stochastic flow). Some known conditions of completeness of flow follow from this theorem as simple corollaries.

For example, let for an equation in Itô form (6.16) in $\mathbb{R}^n$ the following condition of Wintner type (see Theorem 3.39) be fulfilled:

$$\|a(t,m)\| + \|A(t,m)\| < L(\|m\|), \tag{7.14}$$

where $t \in [0, \infty)$; $m \in \mathbb{R}^n$; $L : [0, \infty) \to (0, \infty)$ is continuous and satisfies inequality (3.16) (the norms are with respect to the Euclidean metric in $\mathbb{R}^n$). For example if $L(u) = K(1 + u)$, $K > 0$, formula (7.14) turns into an Itô condition of linear growth (6.21). Without loss of generality we can suppose that L is smooth (see the proof of Corollary 3.42). Pass to a new Riemannian metric (3.17), with respect to which from (7.14) it follows that the equation is uniformly bounded for all $t \in [0, \infty)$ and $m \in \mathbb{R}^n$. The existence of a uniform Riemannian atlas for the metric (3.17) on $\mathbb{R}^n$ is obvious. Thus from Theorem 7.20 and Condition (7.14) it follows that the flow of the equation in $\mathbb{R}^n$ is complete.

Define another left action of G_I on $\mathbb{R}^n \times L(\mathbb{R}^k, \mathbb{R}^n)$ by the formula

$$(B, \beta) \cdot (X, A) = \left(BX - \frac{1}{2}\mathrm{tr}\beta(A(\cdot), A(\cdot)), B \circ A \right). \tag{7.15}$$

Definition 7.22. The *backward Itô bundle* $I_*(M)$ over a manifold M is a bundle (according to Definition 1.32) with standard fiber $\mathbb{R}^n \times L(\mathbb{R}^k, \mathbb{R}^n)$ and structure group G_I that acts on $\mathbb{R}^n \times L(\mathbb{R}^k, \mathbb{R}^n)$ from the left by formula (7.15).

Over every chart $\mathfrak{U}_\alpha$ on M the backward Itô bundle is presented as a direct product $\mathfrak{U}_\alpha \times (\mathbb{R}^n \times L(\mathbb{R}^k, \mathbb{R}^n))$ and under the change of coordinates

$\varphi_{\beta\alpha}$ from the chart $\mathfrak{U}_\alpha$ to another chart $\mathfrak{U}_\beta$ an arbitrary point $(m^\alpha, (a^\alpha, A^\alpha))$ is transformed according to the rule

$$(m^\alpha, (a^\alpha, A^\alpha)) \mapsto \left(\varphi_{\beta\alpha} m^\alpha, \left(\varphi'_{\beta\alpha} a^\alpha - \frac{1}{2} \operatorname{tr} \varphi''_{\beta\alpha}(A^\alpha, A^\alpha), \varphi'_{\beta\alpha} A^\alpha\right)\right). \quad (7.16)$$

Definition 7.23. The cross-sections of a backward Itô bundle $I_*(M)$ are called *backward Itô equations* and are denoted by $(\hat{a}_*, A)$.

7.3 Itô Equations in Belopolskaya-Daletskii Form

If a connection is specified on a manifold M, one can apply it to identify the Itô equations from Section 7.2 with the Itô vector fields from Section 7.1 so that this identification is invariant with respect to changes of coordinates, i.e., it is well-defined on the entire manifold.

Definition 7.24. An Itô vector field (a, A) and an Itô equation $(\hat{a}, \hat{A})$ are said to *canonically correspond to each other at a point* $m \in M$ with respect to the connection H if at m they coincide under the trivialization in the normal chart of H at m. If this identification is fulfilled at all points $m \in M$, (a, A) and $(\hat{a}, A)$ are said to *canonically correspond to each other* with respect to H on M.

Lemma 7.25 *An Itô vector field (a, A) and an Itô equation $(\hat{a}, \hat{A})$ canonically correspond to each other with respect to a connection H on M if and only if in every chart $\mathfrak{U}_\alpha$ the fields of linear operators A and $\hat{A}$ coincide and a and $\hat{a}$ are related by the formula*

$$\hat{a}(t, m) = a(t, m) - \frac{1}{2} \operatorname{tr} \boldsymbol{\Gamma}_m(A(t, m), A(t, m)), \quad (7.17)$$

where $\boldsymbol{\Gamma}_m(\cdot, \cdot)$ is the local connector of H in the chart.

Proof. The equality $A(t, m) = \hat{A}(t, m)$ trivially follows from the facts that at every point the linear operators coincide in the normal chart and have the same rule of transformation under changes of coordinates.

To prove (7.17), choose some $m \in \mathfrak{U}_\alpha$ and consider the normal chart $\mathfrak{U}_n$ of H at this point. Let $X, Y \in T_m M$. Consider the vector in $T_{(m,X)}TM$ which is described in $\mathfrak{U}_n$ by the quadruple $(m, X, Y, 0)$ (see Section 2.1). Then by formula (2.10) in another chart $\mathfrak{U}_\alpha$ this vector takes the form $(\varphi_{\alpha n} m, \varphi'_{\alpha n} X, \varphi'_{\alpha n} Y, \varphi''_{\alpha n}(X, Y))$. Since in $\mathfrak{U}_n$ the local connector of H at m equals zero, from formula (2.19) we obtain that $\boldsymbol{\Gamma}_m(X, Y) = -\varphi''_{\alpha n}(X, Y)$ in $\mathfrak{U}_\alpha$.

Since $a(t, m)$ and $\hat{a}(t, m)$ coincide after trivialization, in the chart $\mathfrak{U}_n$ and under the change of coordinates $\varphi_{\alpha n}$ they transform by formulae (1.1) and (7.12), respectively, and in the chart $\mathfrak{U}_\alpha$ they satisfy the relation

$$
\begin{aligned}
\hat{a}(t,m) &= a(t,m) + \frac{1}{2}\mathrm{tr}\,\varphi''_{\alpha n}(A(t,m),A(t,m)) \\
&= a(t,m) - \frac{1}{2}\mathrm{tr}\,\boldsymbol{\Gamma}_m(A(t,m),A(t,m)).
\end{aligned}
$$

The proof of sufficiency is based on the same formulae. □

The generator $\mathcal{A}$ of the flow of $(\hat{a}, A)$ is well-defined on a manifold, it is a second order tangent vector (see Section 2.9) that in local coordinates of some chart $\mathcal{U}_\alpha$ takes the form $\mathcal{A} = \hat{a}^k \frac{\partial}{\partial q^k} + \frac{1}{2} a^{ij} \frac{\partial^2}{\partial q^i \partial q^j}$ where $\hat{a}^k$ are coordinates of $\hat{a}$ and a^{ij} are elements of the matrix AA^*.

Lemma 7.26 *Let* (a, A) *be the Itô vector field canonically corresponding to the Itô equation* $(\hat{a}, A)$ *with respect to the connection* H*. Then* $a = \mathcal{H}(\mathcal{A})$ *where* $\mathcal{H} : \tau M \to TM$ *is the mapping generated by the connection* H *via formula* (2.45)*.*

Proof. Note that $\mathrm{tr}\boldsymbol{\Gamma}_m(A(t,m),A(t,m)) = \Gamma^k_{ij} a^{ij} \frac{\partial}{\partial q^k}$ where, as above, a^{ij} are the elements of the matrix AA^* in local coordinates. By Lemma 7.25 $\hat{a}(t,m) = a(t,m) - \frac{1}{2}\mathrm{tr}\boldsymbol{\Gamma}_m(A,A)$. Then $\mathcal{H}\mathcal{A} = a(t,m) - \frac{1}{2}\Gamma^k_{ij} a^{ij} \frac{\partial}{\partial q^k} + \frac{1}{2}\Gamma^k_{ij} a^{ij} \frac{\partial}{\partial q^k} = a(t,m)$. □

We now turn to the construction of Ya.I. Belopolskaya and Yu.L. Daletskii (see, e.g., [43, 23]), by means of which it is very easy to describe the Itô equations in terms of Itô vector fields and connections.

Definition 7.27. The *forward stochastic differential*

$$a(t,m)\mathrm{d}t + A(t,m)\mathrm{d}w(t)$$

at a point $m \in M$ given by an Itô vector field (a, A) is the class of stochastic processes in the tangent space $T_m M$ that consists of the solutions of all stochastic differential equations of the form

$$X(t+s) = \int_t^{t+s} \tilde{a}(\tau, X(\tau))\mathrm{d}\tau + \int_t^{t+s} \tilde{A}(\tau, X(\tau))\mathrm{d}w(\tau),$$

where $\tilde{a}(\tau, X)$ is a vector field on $T_m M$; $\tilde{A}(\tau, X) : \mathbb{R}^k \to T_m M$ is a linear operator depending on the parameters $\tau \in R$ and $X \in T_m M$; and the following conditions are satisfied: $\tilde{a}(\tau, X)$ and $\tilde{A}(\tau, X)$ are Lipschitz continuous, are equal to zero outside some neighborhood of the origin in $T_m M$ and such that for $\tau \geq t$ the equalities $\tilde{a}(\tau, 0) = a(t,m)$ and $\tilde{A}(\tau, 0) = A(t,m)$ hold.

Note that since $\tilde{a}(\tau, X)$ and $\tilde{A}(\tau, X)$ are Lipschitz continuous, the process $X(t+s)$ is a strong solution of the equation and so it is well-defined for every Wiener process in $\mathbb{R}^k$.

Let H be a connection and consider its exponential mapping exp.

Definition 7.28. We say that a process $\xi(t)$ satisfies the *Itô equation in Belopolskaya-Daletskii form*

$$\mathrm{d}\xi(t) = \exp_{\xi(t)}(a(t,\xi(t))\mathrm{d}t + A(t,\xi(t))\mathrm{d}w(t)) \tag{7.18}$$

if for every point $\xi(t)$ there exists a neighborhood of $\xi(t)$ in M such that before $\xi(t+s)$, $s \geq 0$, leaves this neighborhood, $\xi(t+s)$ a.s. coincides with a process from the class $\exp_{\xi(t)}(a(t,\xi(t))\mathrm{d}t + A(t,\xi(t))\mathrm{d}w(t))$.

Theorem 7.29 *Let (a, A) be the Itô vector field canonically corresponding to $(\hat{a}, A)$ with respect to the connection H. Equation (7.13) for a process $\xi(t)$ is fulfilled if and only if equality (7.18) holds for $\xi(t)$.*

Proof. Let $\boldsymbol{\Gamma}_m(\cdot,\cdot)$ be the local connector of H in a chart $\mathcal{U}_\alpha$. By formula (7.17) we have that $\hat{a}(t,m) = a(t,m) - \frac{1}{2}\mathrm{tr}\,\boldsymbol{\Gamma}_m(A(t,m),A(t,m))$. Thus equation (7.13) in $\mathcal{U}_\alpha$ can be equivalently written by means of the Itô vector field $(a(t,m),A(t,m))$ in the form

$$\mathrm{d}\xi(t) = a(t,\xi(t))\mathrm{d}t - \frac{1}{2}\mathrm{tr}\,\boldsymbol{\Gamma}_{\xi(t)}(A(t,\xi(t)),A(t,\xi(t)))\mathrm{d}t + A(t,\xi(t))\mathrm{d}w(t). \tag{7.19}$$

Recall that in the chart $\mathcal{U}_\alpha$ the exponential mapping of H for $m \in \mathcal{U}_\alpha$ and $X \in T_mM$ to within terms of order X^3 is presented via the Taylor expansion (see [63])

$$\exp_m X = m + X - \frac{1}{2}\boldsymbol{\Gamma}_m(X,X) + \dots . \tag{7.20}$$

From formula (7.20) and Theorems 6.10 and 6.12 it follows that to within terms of order higher than $\mathrm{d}t$ the equality

$$\begin{aligned}\exp_{\xi(t)}(a_{\xi(t)}\mathrm{d}t + A_{\xi(t)}\mathrm{d}w(t)) = \xi(t) &+ a(t,\xi(t))\mathrm{d}t\\ &- \frac{1}{2}\mathrm{tr}\,\boldsymbol{\Gamma}_{\xi(t)}(A(t,\xi(t)),A(t,\xi(t)))\mathrm{d}t\\ &+ A(t,\xi(t))\mathrm{d}w(t)\end{aligned} \tag{7.21}$$

holds. The assertion of the theorem follows by comparing equalities (7.19) and (7.21). □

Thus for equations of type (7.18) the existence of solution theorems from Section 7.2 are applicable.

Remark 7.30. In the literature the local expression (7.19) for (7.18) is known as the *Itô equation in Baxendale's form* after P. Baxendale who independently found this presentation of Itô equations in charts (see. [21]). We use this term strictly for the local description, retaining the term "Itô equation in Belopolskaya-Daletskii form" for global expressions of type (7.18). The first publication by Ya.I. Belopolskaya and Yu.L. Daletskii in this direction was [44]. Notice also Gangolli's paper [77] where local connectors were first used for the covariant description of diffusion processes on manifolds. This paper was the inspiration for further constructions.

By Theorem 7.29, having specified various connections, we can express an Itô equation in different Belopolskaya-Daletskii forms corresponding to those connections. Thus it is a reasonable idea to look for the "best" connection relative to a given problem. A "good" connection, well-defined in some special cases, is described in [43]. Equations in Belopolskaya-Daletskii form with respect to this connection are equations in Stratonovich form.

The next statement gives an example of another useful connection.

Theorem 7.31 *Let $(\hat{a}, A)$ be an Itô equation on a manifold M that is smooth, autonomous and such that $A(m)$ has rank equal to $\dim M$ at all $m \in M$. Then there exists a connection on M such that the corresponding equation in Belopolskaya-Daletskii form has no drift, i.e., it is described as*

$$\mathrm{d}\xi(t) = \exp_{\xi(t)}(A(\xi(t))\mathrm{d}w(t)). \tag{7.22}$$

Proof. Since $A(m)$ is smooth and has rank equal to $\dim M$ at all $m \in M$, the symmetric $(2,0)$-tensor field AA^* on M is smooth and positive definite. Hence it can be taken as a metric $(2,0)$-tensor on M. Denote its matrix in a chart by (a^{ij}). Then the $(0,2)$-tensor $(a_{ij}) = (a^{ij})^{-1}$ is a Riemannian metric on M. Note that a_{ij} and a^{ij} are the components of this metric tensor. Denote by H the Levi-Civitá connection of this metric, by $\boldsymbol{\Gamma}(\cdot,\cdot)$ its local connector and by Γ^k_{ij} the corresponding Christoffel symbols of the second kind. Thus the equation in Belopolskaya-Daletskii form corresponding to $(\hat{a}, A)$ with respect to H in a chart takes the form (7.19) with the above local connector where (a, A) is the Itô vector field canonically corresponding to $(\hat{a}, A)$ with respect to H.

In [148, Proposition V.4.3] it is shown that for every vector $a = a^k \frac{\partial}{\partial q^i}$ (in particular for a, the first term of the Itô vector field (a, A)) there exists a Riemannian connection $\bar{\mathsf{H}}$ of the metric (a_{ij}) with Christoffel symbols $\bar{\Gamma}^k_{ij}$ such that $a^k = \frac{1}{2}a^{ij}(\Gamma^k_{ij} - \bar{\Gamma}^k_{ij})$. Denote by $\bar{\boldsymbol{\Gamma}}(\cdot,\cdot)$ its local connector. Recall that $\operatorname{tr}\boldsymbol{\Gamma}(A, A) = a^{ij}\Gamma^k_{ij}$. Thus, replacing $a(t, \xi(t))$ in (7.19) by $\frac{1}{2}a^{ij}(\Gamma^k_{ij} - \bar{\Gamma}^k_{ij})$, we obtain that the equation takes the form $\mathrm{d}\xi(t) = -\frac{1}{2}\operatorname{tr}\bar{\boldsymbol{\Gamma}}_{\xi(t)}(A(\xi(t)), A(\xi(t)))\mathrm{d}t + A(\xi(t))\mathrm{d}w(t)$. Hence the equation in Belopolskaya-Daletskii form corresponding to $(\hat{a}, A)$ with respect to $\bar{\mathsf{H}}$ is (7.22). □

Corollary 7.32 *Let $\mathcal{A}(m)$ be a smooth autonomous second order vector field on M with invertible matrices of coefficients at second derivatives. Then there exists a connection on M such that $\mathcal{A}$ is the generator of a solution of a type* (7.22) *equation with respect to this connection.*

Remark 7.33. Theorem 7.31 and Corollary 7.32 are obtained by modifying to Itô equations a construction from [148, Section V.4] involving equations in Stratonovich form. The generalization of that construction to the case where

A has non-maximal but constant rank, for equations in Stratonovich form, is presented in [68, Theorem 2.1.1].

We describe the transformations of equations of type (7.18) under certain special mappings. Let M and N be manifolds equipped with connections. Denote the exponential mapping on N by $\exp^N$ and retain the notation exp for the exponential mapping on M. Let $F : M \to N$ be a C^2-mapping that sends geodesics of the connection on M to geodesics of the connection on N. This means that $F \circ \exp(X) = \exp^N(TF \circ X)$ for $X \in TM$ (see Section 2.4). From this we immediately obtain the statement that replaces Theorem 7.7 for equations (7.18) (see [23]):

Theorem 7.34 *Under the above-mentioned assumptions for a solution $\xi(t)$ of equation (7.18) on M the process $F(\xi(t))$ on N satisfies the equation*

$$\mathrm{d}F(\xi(t)) = \exp^N_{F(\xi(t))}(TF(a(t, \xi(t))\mathrm{d}t + TFA(t, \xi(t))\mathrm{d}w(t)).$$

Now we can present several existence theorems adapted to equation (7.18). For this we introduce the following notion.

Definition 7.35. We say that a connection H and a Riemannian metric $\langle \cdot, \cdot \rangle$ on M are *compatible* if $\langle \cdot, \cdot \rangle$ has a uniform Riemannian atlas in whose charts on the balls $V_m(r)$ the local connector $\boldsymbol{\Gamma}_{m'}(X, X)$ at all $m' \in V_m(r)$ is uniformly bounded in the norm generated by the metric, as a quadratic operator of X, by a certain constant $C_0 > 0$ independent of the choice of chart and ball.

It is obvious that on a compact manifold every Riemannian metric and every connection are compatible. Another class of examples exhibiting this behavior are the left-invariant (right-invariant) Riemannian metrics and connections on Lie groups. Indeed, a left-invariant metric $\langle \cdot, \cdot \rangle$ on a Lie group G has a uniform Riemannian atlas constructed by left shifts to points $g \in G$ of a specified chart in a neighborhood of the unit. The estimates for the local connector of a left-invariant connection in the charts of the obtained atlas remain the same as in the above-mentioned chart at the unit, i.e., they are independent of the choice of chart and ball. The same argument is valid for right-invariant metrics and connections.

Let a connection H be given on M. Denote by exp its exponential mapping.

Theorem 7.36 *Let an Itô vector field $(a(t, m), A(t, m))$ be smooth in $m \in M$ and continuous in $t \in [0, \infty)$. Let there exist a Riemannian metric on M, compatible with H, with respect to which $\|a(t, m)\| < C_1$ and $\|A(t, m)\| < C_1$ (here $C_1 > 0$ is a constant) for all t, m. Then for every initial condition $\xi(0) = m_0$ there exists a strong and strongly unique solution $\xi(t)$ of equation (7.18), well-defined for all $t \in [0, \infty)$.*

Using Theorem 7.29 and formula (7.19), Theorem 7.36 is reduced to Theorem 7.20.

Remark 7.37. The hypothesis of Theorem 7.36 is satisfied for an autonomous smooth Itô vector field on a compact manifold and any connection H.

Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space. All σ-subalgebras of the σ-algebra $\mathcal{F}$ that we use below in this section are assumed to be complete.

Theorem 7.38 *Let the following objects be given: a non-decreasing family of σ-subalgebras $\mathcal{B}_t$ of $\mathcal{F}$ $(t \in [0,l],\ l > 0)$ and a Wiener process $w(t)$ in $\mathbb{R}^n$ adapted to $\mathcal{B}_t$; a stochastic process $a(t)$ in $\mathbb{R}^n$ and a stochastic process $A(t)$ with values in the space $L(\mathbb{R}^n)$ of linear automorphisms of $\mathbb{R}^n$; a connection H on the manifold M, compatible with a Riemannian metric $\langle\cdot,\cdot\rangle$; and a field of linear operators $E_m : \mathbb{R}^n \to T_mM$ smooth in $m \in M$ and uniformly bounded in the norm generated by $\langle\cdot,\cdot\rangle$. Assume that the processes $a(t)$ and $A(t)$ are non-anticipating with respect to $\mathcal{B}_t$, a.s. have continuous sample paths and such that both $a(t)$ and $A(t)$ are a.s. bounded in norm, uniformly in t, by a certain constant. Then for every initial condition $\xi(0) = m_0 \in M$ there exists a strong and strongly unique solution $\xi(t)$ of the equation*

$$\mathrm{d}\xi(t) = \exp_{\xi(t)}(E_{\xi(t)}a(t)\mathrm{d}t + E_{\xi(t)}A(t)\mathrm{d}w(t)),$$

non-anticipating with respect to $\mathcal{B}_t$ and a.s. having continuous sample paths, that is well-defined for $t \in [0,l]$.

The equation (7.19) under consideration is reduced to an equation of type (7.13) with random coefficients. The remaining part of the proof is analogous to that of [23, Theorem 2.2], modified as described in Remark 7.21.

Note that the equation in Theorem 7.38 is an analog of the equation in Stratonovich form in Example 7.6. The proof of Theorem 7.38 can also be obtained from the results of [66] for equations in Stratonovich form.

Remark 7.39. The conditions formulated in the hypotheses of Theorems 7.36 and 7.38 can be slightly weakened. Indeed, from formula (7.19) and Theorem 7.20 it follows that in Theorem 7.36 instead of the compatibility of connection and metric it is enough to require that the estimate

$$\|\mathrm{tr}\Gamma_{m'}(A(t,m'), A(t,m'))\| < c_2$$

for $m' \in V_m(r)$ holds on the balls $V_m(r)$ in the charts of a uniform Riemannian atlas where c_2 is a certain constant independent of ball and chart. In Theorem 7.38 the conditions can be weakened in an analogous way.

Now we are in a position to present an example of an equation in Belopolskaya-Daletskii form that is used below.

Example 7.40. Consider an Itô vector field $(a(t,m), \mathbf{A}(t,m))$ as in Example 7.4 where $a(t,m)$ and $B(t,m)$ are smooth. For this field the Itô equation in Belopolskaya-Daletskii form

$$\mathrm{d}\xi(t) = \exp_{\xi(t)}(a(t,\xi(t))\mathrm{d}t + \mathbf{A}(t,\xi(t)))\mathrm{d}w(t) \tag{7.23}$$

is well-defined where $w(t)$ is a Wiener process in $\mathbb{R}^k$ (see Example 7.4) and exp is the exponential mapping of a connection on M. Introduce a Riemannian metric on M (for example, the first fundamental form generated by the inner product in $\mathbb{R}^k$). If M is a compact manifold, the connection and metric are compatible in the sense of Definition 7.35. The existence of a strong solution for all $t \in [0,\infty)$ is proved as above.

We conclude this section with a description of a class of stochastic differential equations on infinite-dimensional Hilbert manifolds that will be used below. This class is a particular case of the equations on infinite-dimensional manifolds considered in [23] (see also [35, 66], where equations in Stratonovich form are considered).

Let M be a Hilbert manifold, H be a connection on M and $G(\cdot,\cdot)$ be a strong Riemannian metric on M (the term "strong" means that $G(\cdot,\cdot)$ determines the topology of the model space in tangent spaces to M; the description of Riemannian metrics and connections on infinite-dimensional manifolds can be found, e.g., in [172], see also Chapter 10 for the particular case of groups of diffeomorphisms). Notice that in this case Definition 7.9 of a uniform Riemannian atlas and Definition 7.35 of compatible metrics and connections remain valid.

Let $\alpha(t,m)$ be a vector field and $A(t,m)$ be a field of linear operators $A(t,m) : \mathbb{R}^n \to T_mM$ where $m \in M$, $t \in [0,l]$ and $\mathbb{R}^n$ is the Euclidean space in which a Wiener process $w(t)$ takes values. As in the finite-dimensional case the pair (α, A) is called an *Itô vector field* and for this field (and for the exponential mapping exp of the connection H) equation (7.18) is well-defined. For convenience of reference we formulate an existence of solution theorem for (7.18) in this case as a separate statement.

Theorem 7.41 *Let the above mentioned objects G, H, $\alpha(t,m)$ and $A(t,m)$ be given on a Hilbert manifold M and let the conditions of Theorem 7.36 be satisfied for them. Then for every $m_0 \in M$ there exists a strong and strongly unique solution $\xi(t)$ of equation (7.18) with initial condition $\xi(0) = m_0$, well-defined for all $t \in [0,l]$.*

Note that Theorem 7.41 is valid for every finite-dimensional Euclidean space $\mathbb{R}^n$. As for Theorem 7.36, the proof of Theorem 7.41 is reduced to Theorem 7.20. Recall (see Remark 7.21) that the proof of Theorem 7.20 is analogous to that of Theorem 2.2 in [23] (proved in the infinite-dimensional case). The modification mentioned in Remark 7.21 is also valid here.

7.4 Completeness of Stochastic Flows

7.4.1 Setting up the problem and a necessary condition for completeness

In this Section we follow [114] and [116].

Let M be a finite-dimensional non-compact manifold. Consider a stochastic flow $\xi(s)$ on M generated by a stochastic differential equation in Itô or in Stratonovich form with smooth coefficients. Since the coefficients are smooth, we can pass from the Stratonovich to the Itô equation and vice versa. By $\mathcal{A}$ we denote the generator of this flow.

Consider the one-point compactification $M\bigcup\{\infty\}$ of M where the system of open neighborhoods of ∞ consists of complements to all compact sets of M. Denote by $\xi(s) : M \to M\bigcup\{\infty\}$ the stochastic flow. For any point $m \in M$ and time t the orbit $\xi_{t,m}(s)$ of this flow is the unique solution of the above-mentioned equation with initial conditions $\xi_{t,m}(t) = m$. As the coefficients of the equation are smooth, this is a strong solution and so a Markov diffusion process given on some random time interval. The point ∞ is the "cemetery" where the solution (defined on a random time interval) arrives after the explosion.

We refer the reader to [174] for more information on the behavior of a diffusion process at infinity.

Recall that the generator $\mathcal{A}$ is a second order vector (see Definition 2.74). In local coordinates one can find the matrix of its pure second order term, which is symmetric and positive semi-definite.

For a stochastic flow the generator plays the same role as the derivative in the direction of a vector field in the right-hand side of an ordinary differential equation. The main result on completeness for stochastic flows here is analogous to Theorem 3.3 where the derivative in the direction of the vector field X^+ is replaced with the corresponding generator. However in the stochastic case there is an additional difficulty that for a flow with inverse time direction the generator does not coincide with the generator for the flow itself. This is why we obtain a necessary and sufficient condition for completeness only for flows with the additional assumption that the flow must be continuous at infinity (see the exact Definition 7.44 below).

We denote the probability space, where the flow is defined, by $(\Omega, \mathcal{F}, \mathrm{P})$ and assume that it is complete. We also deal with separable realizations of all processes.

Let T be a positive real number.

Definition 7.42. The flow $\xi(s)$ is *complete on* $[0, T]$ if $\xi_{t,m}(s)$ a.s. takes values in M for any pair (t, m) (with $0 \le t \le T$) and for all $s \in [t, T]$. The flow $\xi(s)$ is *complete* if it is complete on any interval $[0, T] \subset R$ (cf. Definition 6.30).

We start with a sufficient condition for completeness of a stochastic flow analogous to conditions for completeness of ODE flows with one-sided estimates. This is a simple version of a rather general sufficient condition [66, Theorem IX. 6A]. We use the notion of a proper function introduced in Definition 3.2.

Theorem 7.43 *Let there exist a smooth proper function φ on M such that $\mathcal{A}(t,m)\varphi < C$ for some $C > 0$ at all $t \in [0, +\infty)$ and $m \in M$. Then the flow $\xi(t,s)$ is complete.*

Proof. Consider the collection of subsets $W_k = \varphi^{-1}([0,k))$ of M where k ranges over the positive integers. Since φ is proper, these sets are relatively compact and $\bigcup\limits_k W_k = M$. Besides, by construction $W_i \subset W_{i+1}$, $i = 1, 2, \ldots$.

Let $t \in [0, +\infty)$ and $m \in M$ and consider the orbit $\xi_{t,m}(s)$. Denote by τ_k the first time $\xi_{t,m}(s)$ hits the boundary of W_k. Transform $\varphi(\xi_{t,m}(s \wedge \tau_k))$ by the Itô formula. Since W_k is relatively compact, the Itô integral on the interval $[t, s \wedge \tau_k)$ is a martingale and so its expectation is equal to 0. Then

$$E\varphi(\xi_{t,m}(s \wedge \tau_k)) = \varphi(m) + \int\limits_t^{s\wedge\tau_k} (\mathcal{A}\varphi)(\theta, \xi_{t,m}(\theta))\mathrm{d}\theta < \varphi(m) + Cs,$$

since $\mathcal{A}(t,m)\varphi < C$ and $s \geq s \wedge \tau_k$.

Consider the set $\Omega_s^k = \{\omega \in \Omega \mid s < \tau_k\}$. Obviously $k(1 - \mathsf{P}(\Omega_s^k)) < E\varphi(\xi_{t,m}(s \wedge \tau_k))$, since for $\omega \notin \Omega_s^k$ we get $\xi_{t,m}(s \wedge \tau_k, \omega) = \xi_{t,m}(\tau_k, \omega)$, i.e., $\varphi(\xi_{t,m}(s \wedge \tau_k, \omega)) = k$. Thus,

$$1 - \mathsf{P}(\Omega_s^k) < \frac{\varphi(m) + Cs}{k}. \tag{7.24}$$

Hence $\lim\limits_{k\to\infty}(1 - \mathsf{P}(\Omega_s^k)) = 0$. However by construction, $\lim\limits_{k\to\infty} \Omega_s^k = \bigcup\limits_{i=1}^{\infty} \Omega_s^i = \Omega$, i.e., for any given $s \geq t$ the value $\xi_{t,m}(s)$ exists in M with probability 1. □

7.4.2 A necessary and sufficient condition for completeness of flows continuous at infinity

In this Section the maximal assumption on the stochastic flow is that its generator $\mathcal{A}(t,x)$ is smooth and strictly elliptic (i.e., in a local coordinate system its pure second order term is described by a positive definite, i.e., invertible, matrix). This assumption allows us to apply the machinery from [12]. Notice that using this machinery we can reduce the condition that the stochastic equation is C^∞-smooth to the assumption that it is Hölder continuous. However, in some statements in this Section we use weaker assumptions, indicated in the hypotheses.

Let $\xi(s)$ be a (not necessarily complete) stochastic flow.

Definition 7.44. We say that the flow $\xi(s)$ is *continuous at infinity* if for any $0 \le t \le T$ and any compact $K \subset M$ the equality

$$\lim_{m\to+\infty} \mathrm{P}(\xi_{t,m}(T)) \in K) = 0 \tag{7.25}$$

holds.

One can easily see that continuity at infinity according to Definition 7.44 means that for any $t \in [0,+\infty)$ and for all $s \in [t,+\infty)$ the correspondence $(m,s) \mapsto \xi_{t,m}(s)$ is continuous (in probability) at the point $(s,\{\infty\}) \in [t,\infty) \times (M \bigcup\{\infty\})$. See [203, 206] for details.

As an example we mention that a flow whose diffusion semigroup has the so-called C_0 property is continuous at infinity. We refer the reader to [206] for relations between continuity of a flow on $M \bigcup\{\infty\}$ and the C_0 property of the corresponding diffusion semigroup (see also, e.g., [12] and [67] for details on the C_0-property).

Our next task is to construct a special proper function associated to a complete stochastic flow $\xi(s)$.

Consider an expanding sequence of compact sets M_i such that $M_i \subset M_{i+1}$ for all i and $\bigcup_i M_i = M$. Let (T_i) be an increasing sequence of real numbers tending to $+\infty$.

For $(t,m) \in [0,T_i]\times M_i$ the distribution function $\mu_{t,m,s}$ of random elements $\xi_{t,m}(s)$, $s \in [t,T_i]$, on M form a weakly compact set of measures. Indeed, take an arbitrary sequence of random elements $\xi_{t_k,m_k}(s_k)$ with corresponding measures μ_{t_k,m_k,s_k}. Since $[0,T_i]\times M_i \times [0,T_i]$ is compact, it is possible to select a subsequence $(t_{k_q}, m_{k_q}, s_{k_q})$ of the sequence (t_k,m_k,s_k) which converges (to (t_0,m_0,s_0), say). It is a well-known fact that the function $Ef(\xi_{t,m}(s))$ is continuous jointly in t,m,s for any bounded continuous function $f: M \to \mathbb{R}$. Then we obtain that $E(f(\xi_{t_{k_q},m_{k_q}}(s_{k_q}))) \to E(f(\xi_{t_0,m_0}(s_0)))$, i.e., from any sequence of measures described above it is possible to select a weakly converging subsequence.

Take a monotonically decreasing sequence of positive numbers $\varepsilon_i \to 0$ such that the series $\sum_{i=1}^{\infty} \sqrt{\varepsilon_i}$ converges. From Prokhorov's theorem it follows that for the measures corresponding to $\xi_{t,m}(s)$, $s \in [t,T_i]$, $(t,m) \in [0,T_i] \times M_i$ mentioned above, there exists a compact $\Xi_i \subset M$ such that for all $\mu_{t,m,s}$ the inequality $\mu_{t,m,s}(M \backslash \Xi_i) < \varepsilon_i$ holds. Construct an expanding system of compact sets $\Theta_i \supset \bigcup_{k=0}^{i} \Xi_k$ for any i, being closures of open domains in M with smooth boundary and such that $\Theta_i \subset \Theta_{i+1}$ for any i and $\bigcup_i \Theta_i = M$. By construction, for $s \in [0,T_i]$, $(t,m) \in [0,T_i] \times M_i$ the relation $\mu_{t,m,s}(M\backslash\Theta_i) < \varepsilon_i$ holds. In particular, $\mu_{t,m,s}(\Theta_{i+1}\backslash\Theta_i) < \varepsilon_i$.

Choose neighborhoods $\mathfrak{U}_i \subset \tilde{\mathfrak{U}}_i$ of the set Θ_i that are proper subset of Θ_{i+1} and consider a smooth function ψ_i that equals 0 on $\mathfrak{U}_i$, equals 1 on $\Theta_{i+1}\backslash\tilde{\mathfrak{U}}_i$ and takes values between 0 and 1 on $\tilde{\mathfrak{U}}_i\backslash\mathfrak{U}_i$. Construct the function

θ on M setting its value on $\Theta_{i+1}\backslash\Theta_i$ equal to $\psi_i \frac{1}{\sqrt{\varepsilon_i}} + (1-\psi_i)\frac{1}{\sqrt{\varepsilon_{i-1}}}$. Notice that on $\Theta_{i+1}\backslash\Theta_i$ the values of θ are not greater than $\frac{1}{\sqrt{\varepsilon_i}}$.

Immediately from the above construction we obtain the following:

Lemma 7.45 *For a complete flow $\xi(s)$ the function θ, constructed above, is smooth, positive and proper.*

Theorem 7.46 *If the flow $\xi(t)$ is complete, for every (t,m) and every $T>t$ the inequality $E\theta(\xi_{t,m}(s))<\infty$ holds for each $s\in[t,T]$.*

Proof. Take i such that $[0,T]\subset[0,T_i]$, $t\in[0,T_i]$ and $m\in M_i$. Then $\mu_{t,m,s}(M\backslash\Theta_i)<\varepsilon_i$ and so $\mu_{t,m,s}(\Theta_i)>(1-\varepsilon_i)$. By construction, the values of the continuous function θ on compact Θ_i are bounded by the constant $\frac{1}{\sqrt{\varepsilon_{i-1}}}$. Then also by construction

$$E\theta(\xi_{t,m}(s)) \le \frac{1}{\sqrt{\varepsilon_{i-1}}} + \sum_{k=i}^{\infty}\varepsilon_k\frac{1}{\sqrt{\varepsilon_k}} = \frac{1}{\sqrt{\varepsilon_{i-1}}} + \sum_{k=i}^{\infty}\sqrt{\varepsilon_k} < C < +\infty \quad (7.26)$$

for some positive constant C since by definition the series $\sum\limits_{k=i+1}^{\infty}\sqrt{\varepsilon_k}$ converges. □

Corollary 7.47 *The function $E\theta(\xi_{t,m}(s))$ is integrable in $s\in[t,T]$.*

Proof. From the construction in Theorem 7.46 it follows that for given t,m, estimate (7.26) is valid with the same C for all $s\in[t,T]$. □

Theorem 7.48 *For every stochastic flow $\xi(s)$ on a manifold M that is complete on an interval $[0,T]$, there exists a proper positive function θ on M such that for all $t\in[0,T]$, $m\in M$ and $s\in[t,T]$ the inequality $E\theta(\xi_{t,m}(s))<\infty$ holds.*

Theorem 7.48 follows from Lemma 7.45 and Theorem 7.46.

Let $T>0$ and consider the direct product $M^T=[0,T]\times M$. Denote by $\pi^T : M^T\to M$ the natural projection: $\pi^T(t,m)=m$.

Theorem 7.49 *Let the generator $\mathcal{A}$ of the complete flow be smooth and strictly elliptic. Then the function $u(t,m)=E\theta(\xi_{t,m}(T))$ on M^T is C^1-smooth in $t\in[0,T]$, C^2-smooth in $m\in M$ and satisfies the equation*

$$\left(\frac{\partial}{\partial t}+\mathcal{A}\right)u=0 \quad (7.27)$$

Proof. Since M is locally compact and satisfies the second countability axiom (and hence is paracompact, see Section 1.1) we can choose a countable locally finite open covering $\{V_i\}_{i=1}^{\infty}$ of M such that all V_i have compact closures. Consider a smooth partition of unity $\{\varphi_i\}_{i=1}^{\infty}$ adapted to this covering. Then at any point $m\in M$ the equality $\theta(m)=\sum_{i=1}^{\infty}\varphi_i(m)\theta(m)$ holds.

Introduce the functions $v_i(m) = \varphi_i(m)\theta(m)$ and the functions $u_i(t,m) = Ev_i(\xi_{t,m}(T))$ and $\theta_k(t,m) = \sum_{i=0}^{k} u_i(t,m)$. Notice that all $v_i(m)$ are smooth and bounded. Then all $v_i(m)$ satisfy the conditions of [81, Theorem VIII.4.1] and so all $u_i(t,m)$ are C^1-smooth in t, C^2-smooth in m and satisfy the relation $\frac{\partial}{\partial t}u_i + \mathcal{A}u_i = 0$. Hence all functions $\theta_k(t,m)$, being finite sums of functions $u_i(t,m)$, are also C^1-smooth in t, C^2-smooth in m and satisfy $\frac{\partial}{\partial t}\theta_k + \mathcal{A}\theta_k = 0$.

In addition it is evident that $\theta(t,m)$ is the limit of $\theta_k(t,m)$ as $k \to \infty$ and the functions $\theta_k(t,m)$ form an increasing locally bounded sequence. Then, since $\mathcal{A}$ is strictly elliptic, the assertion of the Theorem follows from standard Schauder estimates. (For autonomous $\mathcal{A}$, see [12, Lemma 1.8].) □

Theorem 7.50 *If a complete flow $\xi(s)$ is continuous at infinity, the function $u(t,m) = E\theta(\xi_{t,m}(T))$ on M^T is proper.*

Proof. Let $\xi(s)$ be continuous at infinity. To prove the properness of $u(t,m)$ it is sufficient to show that $u(t,m) \to \infty$ as $\theta(m) \to \infty$, i.e., that for any $C > 0$ there exists $\varXi > 0$ such that $\theta(m) > \varXi$ yields $u(t,m) > C$ for any $t \in [0,T]$. Since θ is proper, $K = \theta^{-1}([0,2C])$ is compact. From formula (7.25) in the Definition 7.44 of continuity at infinity it follows that for any $t \in [0,T]$ there exists a $\varXi$ such that $\mathrm{P}(\xi_{t,m}(T) \notin K) > \frac{1}{2}$ for $\theta(m) > \varXi$. Then $u(t,m) = E\theta(\xi_{t,m}(T)) > 2C \cdot \frac{1}{2} = C$. Since t belongs to the compact set $[0,T]$ and $u(t,m)$ is continuous in t, this completes the proof. □

On the manifold M^T consider the flow $\eta(s) = (s, \xi(s))$. Obviously for $(t,m) \in M^T$ the trajectory of $\eta_{(t,m)}(s)$ satisfies the relation $\pi^T(\eta_{(t,m)}(s)) = \xi_{t,m}(s)$. It is clear that $\eta(s)$ is the flow with infinitesimal generator $\mathcal{A}^T$ determined by the formula:

$$\mathcal{A}^T_{(t,m)} = \mathcal{A}(t,m) + \frac{\partial}{\partial t}. \tag{7.28}$$

$\mathcal{A}^T$ is a direct analog of the differentiation in the direction of X^+ described in Theorem 3.3.

Theorem 7.51 *A flow $\xi(s)$ on M, continuous at infinity and having a smooth strictly elliptic generator, is complete on $[0,T]$ if and only if there exists a positive proper function $u^T : M^T \to R$ that is C^1-smooth in $t \in [0,T]$, C^2-smooth in $m \in M$ and such that $\mathcal{A}^T u^T < C$ for some constant $C > 0$ at all points $(t,m) \in M^T$.*

Proof. Let there exist a smooth proper positive function $u^T(t,m)$ on M^T such that $\mathcal{A}^T u^T < C$ at all points of M^T. Then from Theorem 7.43 it follows that $\eta(s)$ is complete. Thus $\xi(s)$ is also complete.

Let $\xi(s)$ be complete. Consider the function $\theta(m)$ on M introduced above and the function $u^T(t,m) = E\theta(\xi_{t,m}(T))$ on M^T. Since $\xi(s)$ is continuous at infinity, $u^T(t,m)$ is proper by Theorem 7.50. By Theorem 7.49 it is also C^1 in t, C^2 in m and satisfies the relation $(\frac{\partial}{\partial t} + \mathcal{A})u^T = \mathcal{A}^T u^T = 0$. □

Corollary 7.52 *A flow $\xi(s)$ on M as in Theorem 7.51 is complete if and only if for all $T > 0$ there exists a positive proper function $u^T : M^T \to R$ on M^T that is C^1-smooth in $t \in [0, T]$, C^2-smooth in $m \in M$ and such that $\mathcal{A}^T u(t, m) < C$ for some constant $C > 0$ at all points $(t, m) \in M^T$.*

7.4.3 Remarks on L^1-complete stochastic flows

For a stochastic flow $\xi(s)$ in the Euclidean space $\mathbb{R}^n$ there is a property stronger than ordinary completeness, where $\xi(s)$ is complete and in addition each of its orbits $\xi_{t,x}(s)$ at every $s > t$ belongs to the functional space $L^1((\Omega, \mathcal{F}, \mathsf{P}), \mathbb{R}^n)$, i.e., $E\|\xi_{t,x}(s)\| < \infty$.

In the general case of flows on manifolds it is natural to replace the norm by a proper function, i.e., to suppose that there exists a positive proper function ψ on M such that $E\psi(\xi_{t,x}(s)) < \infty$ for all $s > t$ (note that both the norm in $\mathbb{R}^n$ and the distance with respect to a complete Riemannian metric are proper functions). Moreover, by Theorem 7.48 such a function exists for every complete flow.

In order not to lose some useful properties of $E\|\xi_{t,x}(s)\|$ (which are not possessed by the function from Theorem 7.48 without additional assumptions), in [120, 121] we introduced the notion of an L^1-complete stochastic flow as follows:

Definition 7.53. A flow $\xi(s)$ on a finite-dimensional manifold M is called *L^1-complete on* $[0, T]$ if the following conditions are fulfilled:

(i) $\xi(s)$ is complete on [0,T];
(ii) there exists a smooth proper positive function $v : M \to \mathbb{R}$ such that $Ev(\xi_{t,m}(T)) < \infty$ for all $m \in M$, $t \in [0, T]$;
(iii) for each $K > 0$ there exists a compact $C_{K,T} \subset M$, depending on K and T, such that the inequality $Ev(\xi_{t,m}(T)) < K$ yields $m \in C_{K,T}$;
(iv) the function $f(t, m) = Ev(\xi_{t,m}(T))$ is C^1-smooth in t and C^2-smooth in m.

A flow is *L^1-complete* if it is L^1-complete on every interval $[0, T] \subset [0, \infty)$.

As above, introduce $M^T = [0, T] \times M$ and the process $\eta_{(t,m)}(s) = (s, \xi_{t,m}(s))$ on M^T. By $\mathcal{A}^T$ we denote the generator of the flow $\eta(s)$. Let $u : M^T \to \mathbb{R}$ be a proper function. Consider the sequence of compact subsets $W_k = u^{-1}([0, k])$ of M^T. Specify a point $(t, m) \in M^T$ and for k such that $(t, m) \in W_k$, denote by τ_k the time $\eta_{(t,m)}(s)$ first hits the boundary of W_k.

In [120, 121] the following necessary and sufficient conditions for L^1-completeness are obtained.

Theorem 7.54 ([120, Theorem 3.1]) *The flow $\xi(s)$ on M is L^1-complete on $[0, T]$, $T > 0$, if and only if there exists a smooth proper positive function $u(t, m)$ on M^T such that for all $(t, m) \in M^T$ the equality $\mathcal{A}^T_{(t,m)} u = C$ holds* *where C is some constant, and for all $(t,m) \in M^T$ the random variables $u(\eta_{(t,m)}(T \wedge \tau_k))$ are uniformly integrable.*

Theorem 7.55 *The flow $\xi(s)$ on M is L^1-complete on $[0,T]$, $T > 0$, if and only if there exists a smooth proper positive function u on M^T such that at every $(t,m) \in M^T$ the following conditions are satisfied:*

1) *$\mathcal{A}^T u \le C$ where C is a positive constant;*
2) *$Eu(\eta_{(t,m)}(T)) < \infty$ and $|Eu(\eta_{(t,m)}(T)) - u(t,m)| < C_1$ where C_1 is a positive constant;*
3) *the function $Eu(\eta_{(t,m)}(T))$ is C^1-smooth in t and C^2-smooth in m.*

The assertion of Theorem 7.55 follows from [121, Theorem 3.6].

The constructions from Section 7.4.2 allow us to obtain the following sufficient condition of L^1-completeness.

Theorem 7.56 *A complete flow, continuous at infinity and having a smooth strictly elliptic generator, is L^1-complete.*

Indeed, from Lemma 7.45 and Theorems 7.46, 7.49 and 7.50 it follows that the function θ constructed in Section 7.4.2, under the hypotheses of the Theorem, meets all requirements set up for the function v by Definition 7.53. In particular, Condition (iii) of Definition 7.53 is fulfilled since by Theorem 7.50 the function $u(t,m) = E\theta(\xi_{t,m}(T))$ is proper on M^T.

7.5 A Condition for Weak Compactness of Measures Corresponding to Solutions of Stochastic Differential Equations

Lemma 7.57 *Consider on M a sequence of smooth Itô equations $(\hat{a}_q(t,m), A_q(t,m))$, $t \in [0,T] \subset \mathbb{R}$, with generators $\mathcal{A}_q(t,m)$, respectively, such that:*

(a) *over each compact set $\mathbf{K} \subset M$ the images $(\hat{a}_q([0,T],\mathbf{K}), A_q([0,T],\mathbf{K}))$ for all q belong to a compact set in $I(M)$;*
(b) *there exists a C^2-smooth proper function $\psi : M \to \mathbb{R}$ such that for all q the inequality*
$$|\mathcal{A}_q \psi| < C \tag{7.29}$$
holds for some constant $C > 0$ independent of q, t and m.

Then

(i) *for every $m_0 \in M$ there exist strong solutions $\xi_q(t)$ with initial condition $\xi_q(0) = m_0$ of all Itô equations $(\hat{a}_q(t,m), A_q(t,m))$ that are strongly unique and well-defined on the entire interval $[0,T]$;*
(ii) *the set of measures $\{\mu_q\}$ corresponding to $\xi_q(\cdot)$ on the Banach manifold $\tilde{\Omega} = C^0([0,T],M)$ of continuous curves in M, equipped with the σ-algebra $\tilde{\mathcal{F}}$ generated by cylinder sets, is weakly compact.*

Proof. Recall that a solution that exists up to the first time the boundary of some chart including m_0 is hit is said to be a *local* solution. The existence and strong uniqueness of strong local solutions $\xi_q(\cdot)$ for all q follows from the fact that all $(\hat{a}_q(t,m), A_q(t,m))$ are smooth and hence are bounded on every relatively compact chart. The global existence follows from Condition (b) by Theorem 7.43.

For every integer p consider the set $W_p = \psi^{-1}([0,p])$. Since ψ is proper, these sets are compact and $\bigcup_p W_p = M$. Besides, by construction $W_p \subset W_{p+1}$ for all $p = 1, 2, \ldots$. Thus m_0 belongs to W_p for sufficiently large p.

For a curve $x(\cdot) \in C^0([0,T], M)$ denote by $\theta_p^{x(\cdot)}$ the time the boundary of W_p is first hit. Introduce the subset in $\tilde{\Omega}$ of the form $\Omega_p = \{x(\cdot) \mid T < \theta_p^{x(\cdot)}\}$ (i.e., every $x(t)$ from Ω_p lies in W_p for all t from $t = 0$ to $t = T$). Taking into account condition (b) it follows from the proof of Theorem 7.43 (see (7.24)) that $\mu_q(\Omega_p) > 1 - \frac{\psi(m_0)+CT}{p}$ for all q. Thus for every $\varepsilon > 0$ there exists a p large enough such that

$$\mu_q(\Omega_p) > 1 - \frac{\varepsilon}{2} \tag{7.30}$$

for all q.

Employing the routine machinery of unit decomposition, one can easily construct a sequence of smooth Itô equations $(\tilde{a}_q(t,m), \tilde{A}_q(t,m))$ such that $(\tilde{a}_q(t,m), \tilde{A}_q(t,m))$ coincides with $(\hat{a}_q(t,m), A_q(t,m))$ for $m \in W_p$, for all q, and $(\tilde{a}_q(t,m), \tilde{A}_q(t,m))$ equals zero outside some relatively compact neighborhood V_p of W_p for all q.

Now choose an arbitrary complete Riemannian metric $g(\cdot,\cdot)$ on M and by Nash's Theorem 1.46 embed M isometrically into a Euclidean space $\mathbb{R}^K$ with K large enough. Introduce an Itô vector field $(\breve{a}_q(t,m), \breve{A}_q(t,m))$ on V_p by setting $\breve{A}_q(t,m) = \tilde{A}_q(t,m)$ and in each chart $\mathcal{U}$ in V_p by setting $\breve{a}^k(t,m) = \tilde{a}^k(t,m) + \Gamma^k_{ij}\tilde{\alpha}^{ij}$ where Γ^k_{ij} are Christoffel symbols of the second kind of the Levi-Civitá connection of the metric $g(\cdot,\cdot)$ in $\mathcal{U}$, $(\tilde{\alpha}^{ij}) = (\tilde{A}^i_l)(\tilde{A}^j_l)^*$.

Let $\mathsf{N}(M)$ be the normal bundle of M in $\mathbb{R}^K$ with fibers N_m, $m \in M$. Denote by Θ a relatively compact tubular neighborhood of M over V_p in $\mathbb{R}^K$ (it exists since V_p is relatively compact) and by $r : \Theta \to V_p$ the smooth retraction of Θ onto V_p along the fibers of $\mathsf{N}(M)$.

Recall that Θ has the structure of a direct product

$$\Theta = V_p \times \mathcal{O}, \tag{7.31}$$

where $\mathcal{O}$ is an open ball in $\mathbb{R}^{K-n}$ that at any point $m \in V_p$ can be identified with the normal space N_m.

At any point $(m,x) \in \Theta$ the presentation (7.31) yields the presentation of a tangent space to $\mathbb{R}^K$ of the form $T_{(m,x)}\mathbb{R}^K = T_m M \times T_x\mathcal{O}$. Introduce a new Riemannian metric $g_1(\cdot,\cdot)$ on Θ by transferring the Riemannian inner product from $T_m M$ into the factor $T_m M$ in the above product by determining the inner product in the factor $T_x\mathcal{O}$ as the restriction of the Euclidean inner product in $\mathbb{R}^K$ and by setting the factors in $T_mM \times T_x\mathcal{O}$ to be orthogonal to each other. Let $\mathfrak{U}$ be a chart on V_p and consider the chart $\mathcal{U} = \mathfrak{U} \times \mathcal{O}$ in Θ. In this chart the matrix of the $(0,2)$-metric tensor $g_1(\cdot,\cdot)$ will be denoted by (g^1_{ij}) and the matrix of the corresponding $(2,0)$-metric tensor by (g_1^{ij}).

Calculate the Christoffel symbols $\bar{\Gamma}^l_{ij}$ of the Levi-Civitá connection of $g_1(\cdot,\cdot)$ in $\mathcal{U}$ by the usual formula $\bar{\Gamma}^l_{ij} = \frac{1}{2}g_1^{lk}(\frac{\partial}{\partial q^i}g^1_{jk} + \frac{\partial}{\partial q^j}g^1_{ik} - \frac{\partial}{\partial q^k}g^1_{ij})$. One can easily calculate that:

a) if $\frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j}, \frac{\partial}{\partial q^k}, \frac{\partial}{\partial q^l} \in T_mM$, then $\bar{\Gamma}^l_{ij} = \Gamma^l_{ij}$ where Γ^l_{ij} are the Christoffel symbols of the Levi-Civitá connection of $g(\cdot,\cdot)$ on M in the chart $\mathfrak{U}$;
b) if $\frac{\partial}{\partial q^k} \in T_mM$ and $\frac{\partial}{\partial q^l} \in T_x\mathcal{O}$ or vice versa, then $\bar{\Gamma}^l_{ij} = 0$ for all $\frac{\partial}{\partial q^i}$ and $\frac{\partial}{\partial q^j}$ since $g_1^{kl} = 0$;
c) if $\frac{\partial}{\partial q^k}, \frac{\partial}{\partial q^l} \in T_mM$ and $\frac{\partial}{\partial q^i} \in T_x\mathcal{O}$, $\frac{\partial}{\partial q^j} \in T_mM$, then $g^1_{jk} = g_{jk}$ does not depend on $\frac{\partial}{\partial q^i}$ and so $\frac{\partial}{\partial q^i}g^1_{jk} = 0$. It is also obvious that $g^1_{ik} = 0$ and $g^1_{ij} = 0$. Hence $\bar{\Gamma}^l_{ij} = 0$. Applying analogous arguments we obtain that $\bar{\Gamma}^l_{ij} = 0$ for $\frac{\partial}{\partial q^i} \in T_mM$, $\frac{\partial}{\partial q^j} \in T_x\mathcal{O}$ and $\bar{\Gamma}^l_{ij} = 0$ for $\frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j} \in T_x\mathcal{O}$;
d) if $\frac{\partial}{\partial q^k}, \frac{\partial}{\partial q^l} \in T_x\mathcal{O}$ then for all $\frac{\partial}{\partial q^i}$ and $\frac{\partial}{\partial q^j}$ we obtain $\bar{\Gamma}^l_{ij} = 0$.

Let O be a neighborhood of V_p in Θ such that $\overline{O} \subset \Theta$ where $\overline{O}$ is the closure of O. Let $\phi(y) : \mathbb{R}^K \to \mathbb{R}$ be a smooth function satisfying the relations $0 \le \phi \le 1$, $\phi(y) = 1$ for $y \in \overline{O}$ and $\phi(y) = 0$ for $y \notin \Theta$. Using the presentation of the chart $\mathcal{U}$ on Θ as the above-mentioned direct product, introduce a new object on $\mathcal{U}$ by the formula

$$\widehat{\Gamma}^k_{i,j}(m,x) = (\phi(m,x)\bar{\Gamma}^k_{ij}(m), 0), \quad (m,x) \in \Theta. \tag{7.32}$$

Consider Θ as a chart with local coordinates inherited from the global coordinate system in $\mathbb{R}^K$. This chart will be called *global*. Find the values of the Christoffel symbols $\bar{\Gamma}^k_{ij}$ in the global chart and define the values of $\widehat{\Gamma}^k_{ij}$ on the complement $\mathbb{R}^k \backslash \Theta$ as $\widehat{\Gamma}^k_{ij}(m,x) = 0, \quad (m,x) \notin \Theta$. Thus the values of $\widehat{\Gamma}^k_{ij}$ are given on all of $\mathbb{R}^k$ and since the functions $\widehat{\Gamma}^k_{ij}$ are smooth and have non-zero values only on a compact set, their values are uniformly bounded. By construction, both in the chart $\mathcal{U}$ and in the global chart, the symbols $\widehat{\Gamma}^k_{ij}$ on O coincide with the corresponding $\bar{\Gamma}^k_{ij}$.

Define the vector fields $\widehat{a}_q$, the fields of linear operators $\widehat{A}_q : \mathbb{R} \times \mathbb{R}^N \to T\mathbb{R}^K$ and the $(2,0)$-tensor fields $\mathfrak{a}_q$ on $\mathbb{R}^K$ by defining them in the chart $\mathcal{U}$ by the formulae

$$\begin{aligned} \widehat{a}_q(t,m,x) &= (\phi(m,x)\breve{a}_q(t,m), 0); \\ \widehat{A}_q(t,m,x) &= \begin{pmatrix} \phi(m,x)\breve{A}_q(t,m) \\ 0 \end{pmatrix}; \\ \mathfrak{a}_q(t,m,x) &= \begin{pmatrix} \phi(m,x)(\tilde{\alpha}^{ij}) & 0 \\ 0 & 0 \end{pmatrix}, \end{aligned} \tag{7.33}$$

and by extending them to all of $\mathbb{R}^K$ by setting them to be equal to zero elsewhere.

Notice that from condition (a) and from the construction it follows that all fields $\widehat{a}_q(t,m,x)$, $\widehat{A}_q(t,m,x)$ and $\mathfrak{a}_q(t,m,x)$ are uniformly bounded on $[0,T]\times\mathbb{R}^K$.

Denote the matrix of $\mathfrak{a}_q(t,m,x)$ in the chart $\mathcal{U}$ by $(\mathfrak{a}^{ij})$ and for the object that in $\mathcal{U}$ is determined as $\widehat{\Gamma}^k_{ij}\mathfrak{a}^{ij}$, introduce the invariant notation $\mathrm{tr}(\widehat{\boldsymbol{\Gamma}}(\widehat{A},\widehat{A}))$.

Consider the following problems in $\mathbb{R}^K$:

$$
\begin{aligned}
\mathrm{d}\widehat{\xi}_q(t) = \widehat{a}_q\left(t,\widehat{\xi}_q(t)\right)\mathrm{d}t - \frac{1}{2}\mathrm{tr}\left(\widehat{\boldsymbol{\Gamma}}_{\widehat{\xi}_q(t)}\left(\widehat{A}_q\left(t,\widehat{\xi}_q(t)\right),\widehat{A}_q\left(t,\widehat{\xi}_q(t)\right)\right)\right)\mathrm{d}t \\
+\widehat{A}_q\left(t,\widehat{\xi}_q(t)\right)\mathrm{d}w(t), \qquad (7.34)\\
\widehat{\xi}_q(0) = m_0\in M.
\end{aligned}
$$

Since all the coefficients in (7.34) are smooth and bounded, the equations have unique strong solutions $\widehat{\xi}_q(t)$ well-defined on the interval $[0,T]$.

After transition to the chart $\mathcal{U}$, taking into account the form of the Christoffel symbols (see above), one can easily see that in the neighborhood $O\cap\mathcal{U}$ equations (7.34) are transformed into the system

$$
\begin{cases}
d\breve{\xi}_q(t) = \breve{a}_q\left(t,\breve{\xi}_q(t)\right)\mathrm{d}t - \frac{1}{2}\mathrm{tr}\left(\bar{\boldsymbol{\Gamma}}_{\breve{\xi}_q}\left(\breve{A}_q,\breve{A}_q\right)\right)\mathrm{d}t + \breve{A}_q\left(t,\breve{\xi}_q(t)\right)\mathrm{d}w(t),\\
d\bar{\xi}_q(t) = 0
\end{cases}
\tag{7.35}
$$

with initial conditions $\breve{\xi}_q(0)=m_0$ and $\bar{\xi}_q(0)=0$. Hence the solutions of (7.35) (and so of (7.34)) a.s. belong to M for all $t\in[0,T]$ and coincide with the solutions of the first parts of (7.35). In particular the corresponding measures $\widehat{\mu}_q$ on the path space take the value 1 on the curves lying in M. But since the coefficients of (7.34) are uniformly bounded in $\mathbb{R}^K$, from the Corollary to [83, Lemma III.2.2] it follows that the set of corresponding measures $\{\widehat{\mu}_q\}$ on $\tilde{\Omega}$ is weakly compact. Then by Prokhorov's theorem (see, e.g., [82]) for every $\varepsilon>0$ there exists a compact set $\Xi\subset C^0([0,T],M)$ such that for all q the inequality $\widehat{\mu}_q(\Xi)>1-\frac{\varepsilon}{2}$ holds.

Note that by construction, on W_p the right-hand sides of the first parts of (7.35) coincide with $(\hat{a}_q, A_q)$. Then (see, e.g., [23, theorem III.3.3]) the processes $\xi_q(t)$ and $\widehat{\xi}_q(t)$ a.s. coincide before leaving W_p for all q. From this and from (7.30) it follows that for the compact set $\Omega_p\cap\Xi$ the inequality $\mu_q(\Omega_p\cap\Xi)>1-\varepsilon$ holds for all q.

Thus, for every $\varepsilon>0$ there exists a compact subset of $\tilde{\Omega}$ whose measure μ_q for all q is greater than $1-\varepsilon$. By Prokhorov's theorem this means that the set $\{\mu_q\}$ is weakly compact. □

7.6 Stochastic Development and Parallel Translation

7.6.1 The Eells-Elworthy and Itô developments

Let $\pi : OM \to M$ be the manifold of orthonormal frames on a Riemannian manifold M, H be the Levi-Civitá connection on OM and V be the vertical distribution on OM. Recall (see Section 2.7) that the bundles V and H over OM are trivial: V is trivialized by fundamental vector fields and H by basic vector fields $\mathsf{E}(x)$ where the vector $\mathsf{E}_b(x) \in \mathsf{H}_b$ for $b \in OM$ and $x \in \mathbb{R}^n$ is defined by the equality $\mathsf{E}_b(x) = T\pi^{-1}(bx)_{|\mathsf{H}_b}$ (the frame b is considered here as a linear operator $b : \mathbb{R}^n \to T_{\pi b}M$, see the proof of Theorem 2.67 and Definition 2.68). Thus the tangent bundle $TOM = \mathsf{H} \oplus \mathsf{V}$ is also trivial.

Definition 7.58. The Riemannian metric on OM, generated by the above-mentioned trivialization of the tangent bundle TOM, is said to be *induced*.

Remark 7.59. It is easy to see that the restriction of every induced metric to the connection space H_b in T_bOM coincides with the pull-back of the Riemannian inner product in $T_{\pi b}M$ under the mapping $T\pi$. The restriction of an induced metric to V is determined by a certain inner product in the algebra $\mathfrak{o}(n)$. Thus, a Riemannian metric on M and an inner product in $\mathfrak{o}(n)$ uniquely define an induced metric on OM.

Consider a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and a non-decreasing family $\mathcal{B}_t$ of complete σ-subalgebras of the σ-algebra $\mathcal{F}$ such that a Wiener process $w(t)$, taking values in some Euclidean space $\mathbb{R}^k$, is adapted to it. Let $m_0 \in M$. Let a stochastic process $\alpha(t)$, $t \in [0, l]$ with values in $T_{m_0}M$ and a stochastic process $A(t)$, $t \in [0, l]$, with values in the space of linear mappings $L(R^k, T_{m_0}M)$ be given on $(\Omega, \mathcal{F}, \mathsf{P})$ and let those processes be non-anticipative with respect to $\mathcal{B}_t$. Finally, let $\alpha(t)$ and $A(t)$ a.s. have continuous sample paths and a.s. for $t \in [0, l] \subset \mathbb{R}$ the integral $\int_0^t \alpha(\tau)\mathrm{d}\tau$ and Stratonovich integral $\int_0^t A(\tau) \circ \mathrm{d}w(\tau)$ be well-defined.

Take an orthonormal frame b_0 in $T_{m_0}M$ and consider the processes $b_0^{-1}\alpha(t)$ and $b_0^{-1}A(t)$ in the Euclidean space $\mathbb{R}^n$ (here n is the dimension of M) and in $L(\mathbb{R}^k, \mathbb{R}^n)$, respectively. Construct a basic Itô vector field with random coefficients on OM that at $b \in OM$ has the form $(\mathsf{E}_b(b_0^{-1}\alpha(t)), \mathsf{E}_b(b_0^{-1}A(t)))$.

Consider the Stratonovich stochastic differential equation on OM (cf. Examples 7.5 and 7.6) of the form

$$\mathrm{d}\eta(t) = \mathsf{E}_{\eta(t)}(b_0^{-1}\alpha(t))\mathrm{d}t + \mathsf{E}_{\eta(t)}(b_0^{-1}A(t) \circ \mathrm{d}w(t)). \tag{7.36}$$

Since $\mathsf{E}_b : \mathbb{R}^n \to TOM$ is smooth in b (see Section 2.7), this equation has a strong and strongly unique solution $\eta_{\,0,b_0}(t)$ in $T_{m_0}M$ with initial condition $\eta_{\,0,b_0}(0) = b_0$, well-defined (generally speaking) on some random time interval.

Consider the process $z(t) = \int_0^t \alpha(\tau)\mathrm{d}\tau + \int_0^t A(\tau) \circ \mathrm{d}w(\tau)$ in $T_{m_0}M$.

Definition 7.60. The process $\pi\eta_{\,0,b_0}(t)$ on M is called the *Eells-Elworthy development* of the process $z(t)$ and is denoted by $R_{EE}\,z(t)$ (see, e.g., [66]). The process $\eta_{\,0,b_0}(t)$ is called the *horizontal lift* of $R_{EE}\,z(t)$ to OM with initial condition b_0.

The Eells-Elworthy development is a stochastic generalization of the operation that is the inverse to Cartan's development (see Remark 3.45).

Lemma 7.61 *$R_{EE}\,z(t)$ does not depend on the initial frame b_0.*

Proof. Let $\bar{b} \in O_{m_0}(M)$. Since the fiber $O_{m_0}M$ is isomorphic to the orthogonal group $O(n)$, there exists an operator $b' \in O(n)$ such that $\bar{b} = b_0 \circ b'$. Since the connection H is invariant with respect to the right action of $O(n)$ on OM, from the definition of the mapping E and from the uniqueness of the solution of (7.36) it follows that $\bar{\eta}(t) = \eta_{\,0,b_0}(t) \circ b'$ is the unique strong solution of the equation

$$\mathrm{d}\bar{\eta}(t) = \mathsf{E}_{\bar{\eta}(t)}\left(\bar{b}^{-1}\alpha(t)\mathrm{d}t\right) + \mathsf{E}_{\bar{\eta}(t)}\left(\bar{b}^{-1}A(t) \circ \mathrm{d}w(t)\right)$$

that starts at $\bar{b}$. The projections $\pi\bar{\eta}(t)$ and $\pi\eta_{\,0,b_0}(t)$ coincide. □

Note that $\bar{\eta}(t)$ from the proof of Lemma 7.61 is the horizontal lift of $R_{EE}z(t)$ with initial value $\bar{b}$.

For a development based on equations in Itô form on OM, we need additional constructions.

It is well-known (see, e.g., [26]) that the integral curves of autonomous basic and fundamental vector fields, and only these curves, are respectively the horizontal and vertical geodesics of the Levi-Civitá connection of every induced metric on OM. (But integral curves of constant linear combinations of basic and fundamental vector fields are not geodesics.) Recall (see Section 2.7) that integral curves of basic vector fields, and only these curves, are horizontal lifts of geodesics of the Levi-Civitá connection on M.

Denote by e the exponential mapping of the Levi-Civitá connection of some induced metric on OM.

Lemma 7.62

(i) *For all induced metrics the restrictions $\mathsf{e}_{|\mathsf{H}}$ coincide.*
(ii) *For every $Y \in \mathsf{H}$ the equality $\pi\mathsf{e}(Y) = \exp(T\pi Y)$ holds where $\exp$ is the exponential mapping of the Levi-Civitá connection on M.*
(iii) *For all induced metrics in every specified chart on OM, the restrictions of the local connectors $\boldsymbol{\Gamma}^{\mathsf{e}}(X, X)$ to H coincide as operators of $X \in \mathsf{H}$.*

Statements (i) and (ii) follow from the above-mentioned properties of integral curves of basic vector fields. In order to prove (iii) note in addition that in a chart the operator $-\boldsymbol{\Gamma}^{\mathsf{e}}(X, X)$, $X \in \mathsf{H}$, is the second derivative of a horizontal geodesic with initial velocity X (i.e. it is an integral curve of a basic vector field) that is independent of the choice of induced metric.

Let processes $\alpha(t)$ and $A(t)$ be given as above on the probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Let the Itô integral $\int_0^t A(\tau)\mathrm{d}w(\tau)$ for $t \in [0,l]$ be well-defined (this condition replaces the above condition that the Stratonovich integral exists). As above, construct the basic Itô vector field with random coefficients on OM that at $b \in OM$ takes the form $(\mathsf{E}_b(b_0^{-1}\alpha(t)), \mathsf{E}_b(b_0^{-1}A(t)))$. Consider the following Itô equations in Belopolskaya-Daletskii form on OM:

$$\mathrm{d}\xi(t) = \mathsf{e}_{\xi(t)}\left(\mathsf{E}_b(b_0^{-1}\alpha(t))\mathrm{d}t + \mathsf{E}_b(b_0^{-1}A(t)\mathrm{d}w(t))\right). \tag{7.37}$$

As in the case of equation (7.36), from the fact that E_b is smooth in b it follows that there exists a strong and strongly unique solution $\xi_{\,0,\,b_0}(t)$ of equation (7.37) with initial condition $\xi_{\,0,\,b_0}(0) = b_0$ that is given on some random time interval. In $T_{\pi b_0}M$ consider the Itô process $y(t) = \int_0^t \alpha(\tau)\mathrm{d}\tau + \int_0^t A(\tau)\mathrm{d}w(\tau)$.

Definition 7.63. The process $\pi\xi_{\,0,\,b_0}(t)$ on M is called the *Itô development* of the process $y(t)$ in $T_{\pi b_0}M$ and is denoted by $R_I\, y(t)$. The solution $\xi_{\,0,\,b_0}(t)$ of equation (7.37) is called the *horizontal lift* of the process $R_I\, y(t)$ to OM with initial value b_0.

By construction and by general formula (7.19) the description of equation (7.63) in a chart on OM contains a local connector $\boldsymbol{\Gamma}^{\mathsf{e}}(X, X)$, restricted to H. We emphasize that it follows from Lemma 7.62 that $R_I(y(t))$ is independent of the choice of induced metric on OM.

Lemma 7.64 *$R_I\, y(t)$ does not depend on the choice of the initial value b_0 of the horizontal lift.*

The proof is analogous to that of Lemma 7.61. Here another horizontal lift of $R_I\, y(t)$ also can be obtained by the action of an orthonormal operator to the process $\xi_{\,0,\,b_0}(t)$.

Lemma 7.65 *The process $R_I\, y(t)$ on M is a solution of the Itô equation in Belopolskaya-Daletskii form*

$$\mathrm{d}(R_I\, y(t)) = \exp_{R_I\, y(t)}(T\pi(\mathsf{E}_b(b_0^{-1}\alpha(t))\mathrm{d}t + \mathsf{E}_b(b_0^{-1}A(t)\mathrm{d}w(t))).$$

This statement follows immediately from Theorem 7.34, Lemma 7.62(ii) and the construction of the process $R_I\, y(t)$.

Remark 7.66. The process $R_I z(t)$ can be constructed in another way that is analogous to the classical construction of the Itô integral with varied upper limit. Indeed, for a subdivision $q = (0 = t_0 < t_1 < ... < t_q = l)$ determine the process $\hat{\xi}^q(t)$, starting at $\hat{b}$, as follows. Start integral curves of the vector field $E(\hat{b}^{-1}(\alpha(0)t_1 + A(0)w(t_1)))$ from $\hat{b}$ up to t_1, then start integral curves of $E(\hat{b}^{-1}(\alpha(t_1)(t_2 - t_1) + A(t_1)(w(t_2) - w(t_1))))$ from the points $\xi^q(t_1)$ up to t_2, etc. One can easily see that under the above conditions the process $\hat{\xi}^q(t_1)$ is well-defined on the entire interval $[0,l]$. The process $\hat{\pi}\xi^q(t)$ M a.s.

has piecewise geodesic sample paths. As $\operatorname{diam} q \to 0$ the processes $\pi\hat{\xi}^q(t)$ converge in probability to $\pi\hat{\xi}(t)$ uniformly in t. In particular it is possible to select a subsequence converging to $\pi\widehat{\xi}(t)$ a.s. uniformly.

Remark 7.67. Like R_{EE}, the operator R_I is a stochastic analog of the operation inverse to the classical Cartan development (see Remark 3.45). Note in addition that according to the theory of stochastic processes the Itô processes in $T_{m_0}M$ are analogs of smooth curves, to which the classical Cartan development is applied. From Remark 7.66 it follows that R_I is an extension of the inverse of Cartan's development from the set of piecewise smooth curves to a.s. all sample trajectories of processes $z(t)$ that are continuous but a.s. not smooth [77].

One can apply analogous methods to "develop" processes from Euclidean space $\mathbb{R}^n$ in which a Wiener process $w(t)$ takes values. In this way it is possible to construct developments with random initial data.

Consider a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and a non-decreasing family $\mathcal{B}_t$, $t \in [0, l]$, of complete σ-subalgebras of the σ-algebra $\mathcal{F}$. Let on that probability space the following objects be given: a Wiener process $w(t)$ with values in a Euclidean space $\mathbb{R}^n$ adapted to $\mathcal{B}_t$, a process $a(t)$ in $\mathbb{R}^n$ with a.s. continuous sample paths and a process $A(t)$ in $L(\mathbb{R}^n, \mathbb{R}^n)$ also with a.s. continuous sample paths, both $a(t)$ and $A(t)$ non-anticipative with respect to $\mathcal{B}_t$ and such that the Itô process $y(t) = \int_0^t a(\tau)\mathrm{d}\tau + \int_0^t A(\tau)\mathrm{d}w(\tau)$ is well-defined. Consider the following Itô equation in Belopolskaya-Daletskii form on OM:

$$\mathrm{d}\xi(t) = \mathsf{e}_{\xi(t)}(\mathsf{E}_b(a(t))\mathrm{d}t + \mathsf{E}_b(A(t)\mathrm{d}w(t))). \tag{7.38}$$

Let $x_0 : \Omega \to M$ be a random element independent of $\mathcal{B}_t$, and β_0 be a Borel measurable cross-section of OM (recall that on a non-parallelizable manifold M the bundle OM may not have continuous cross-sections, but Borel measurable cross-sections do exist).

Since the mapping E is smooth, as above a solution $\hat{\xi}_{0,\beta_0(x_0)}(t)$ of equation (7.38) with initial condition $\hat{\xi}_{0,\beta_0(x_0)}(0) = \beta_0(x_0)$ exists on some random time interval.

Definition 7.68. The process $R_I(\beta_0(x_0))y(t) = \pi\hat{\xi}_{0,\beta_0(x_0)}(t)$ is called the *Itô development* of the process $y(t)$ in $\mathbb{R}^n$ generated by $\beta_0(x_0)$. The process $\hat{\xi}_{0,\beta_0(x_0)}(t)$ is called the *horizontal lift* of the process $R_I(\beta_0(x_0))y(t)$ to OM with initial value $\beta_0(x_0)$.

Note that here the development depends on the initial value of the horizontal lift.

An analogous construction is also valid for the Eells-Elworthy development.

7.6.2 Wiener processes on Riemannian manifolds. Stochastic completeness

Let M be a Riemannian manifold and H be the Levi-Civitá connection on OM. Consider a Wiener process $w(t)$ in $\mathbb{R}^n$ and the "basic" stochastic process $E_b\big(w(t)\big)$ in H_b, $b \in OM$. Observe that the field of processes $E\big(w(t)\big)$ is smooth on OM, i.e., obtained from w by means of the smooth map $E\colon OM \times \mathbb{R}^n \to H$.

As above, for every $m \in M$, a frame $b \in O_m(M)$ can be regarded as a linear operator $b\colon \mathbb{R}^n \to T_{\pi b}M$ (see Section 2.7).

Definition 7.69. The process $T\pi E_b(w(t)) = b\,w(t)$ is called a *realization of the Wiener process* w in $T_{\pi b}M$ or simply a *Wiener process* in $T_{\pi b}M$.

A realization of w in T_mM gives rise to the standard Wiener process in T_mM, i.e., a measure on the space of continuous curves in T_mM (cf. Section 6.2.1).

Theorem 7.70 *The standard Wiener process in T_mM is independent of the choice of $w(t)$ on $\mathbb{R}^n$ and $b \in O_mM$.*

Proof. Since b is an orthogonal operator, $b\,w(t)$ is a Wiener process in the Euclidean space T_mM with the inner product given by the Riemannian metric. Thus, the measure determined by $b\,w(\cdot)$ on the space of curves in T_mM is the Wiener measure with respect to this inner product. Let $b_1, b_2 \in O_m(M)$. It is clear that b_1 and b_2 differ by an orthogonal operator on T_mM. The theorem follows, since the Wiener measure is invariant with respect to the group of orthogonal transformations. $\square$

Thus, once a Riemannian metric on M is specified, we have a well-defined standard Wiener process in every tangent space to M. Furthermore, this field of Wiener processes is smooth, i.e., obtained from the standard Wiener process in $\mathbb{R}^n$ by means of a smooth linear transformation, namely, by means of $T\pi E$. We denote the Wiener process on T_mM by w_m or just by w when no confusion may arise. The realization of w in T_mM obtained by the use of b is denoted by $b\,w$.

Having taken a Wiener process $w(t)$ in a tangent space $T_{m_0}M$, we can apply to it either the Eells-Elworthy development R_{EE} or the Itô development R_I.

Theorem 7.71 $R_{EE}w(t) = R_Iw(t)$.

Proof. For $w(t)$ equation (7.36) takes the form

$$\mathrm{d}\eta(t) = \mathsf{E}_{\eta(t)} \circ \mathrm{d}w(t), \tag{7.39}$$

and equation (7.37), the form

$$d\xi(t) = \mathsf{e}_{\xi(t)}\mathsf{E}_{\xi(t)}dw(t). \tag{7.40}$$

We describe these two equations in the local coordinates of some chart on OM. By formula (6.25) we transform equation (7.39) into the Itô equation

$$d\eta(t) = \frac{1}{2}\mathrm{tr}\,\mathsf{E}'\left(\mathsf{E}_{\eta(t)}\right)dt + \mathsf{E}_{\eta(t)}dw(t),$$

and by formula (7.19) equation (7.40) takes the form

$$d\xi(t) = -\frac{1}{2}\mathrm{tr}\,\boldsymbol{\Gamma}^{e}_{\xi(t)}(I,I)dt + \mathsf{E}_{\xi(t)}dw(t),$$

where I is the unit operator. The assertion of the theorem follows from the fact that $\frac{1}{2}\mathrm{tr}\,\mathsf{E}'(\mathsf{E}_b) = -\frac{1}{2}\mathrm{tr}\,\boldsymbol{\Gamma}^e_b(I,I)$ at every $b \in OM$, i.e. the equations coincide. □

Note that both equation (7.39) and equation (7.40) are determined by the Itô vector field $(0, \mathsf{E}_b)$.

Definition 7.72. The development $R_I w_{m_0}(t) = R_{EE}w(t)$ of a Wiener process w_{m_0} in $T_{m_0}M$ is called a *Wiener process on* M beginning at $m_0 \in M$.

The definition of a Wiener process on M as the development of a Wiener process in a tangent space is due to Eells and Elworthy (see the monograph [66] and the bibliography therein.)

Theorem 7.73 *The generator of a Wiener process on a manifold M is $\frac{1}{2}\nabla^2$ where ∇^2 is the Laplace-Beltrami operator (see Definition 2.58) that in local coordinates of some chart takes the form $-g^{ij}\Gamma^k_{ij}\frac{\partial}{\partial q^k} + g^{ij}\frac{\partial^2}{\partial q^i \partial q^j}$ where Γ^k_{ij} are the Christoffel symbols of the Levi-Civitá connection.*

Theorem 7.73 is proved by application of Lemma 7.62(ii) and Theorem 7.34 to equation (7.37) and then by direct calculation in local coordinates by the use of formula (7.10).

Recall the Definition 7.19 of a martingale with respect to a connection.

Theorem 7.74 *A Wiener process $w(t)$ on a Riemannian manifold M is a martingale with respect to the Levi-Civitá connection.*

Proof. By Theorem 7.73 the generator of a Wiener process is $\frac{1}{2}\nabla^2$ that in local coordinates has the presentation $\frac{1}{2}(-g^{ij}\Gamma^k_{ij}\frac{\partial}{\partial q^k} + g^{ij}\frac{\partial}{\partial q^i \partial q^j})$. Apply to this generator the mapping $\mathcal{H}$ of the Levi-Civitá connection by formula (2.45). Then $\mathcal{H}(\frac{1}{2}(-g^{ij}\Gamma^k_{ij}\frac{\partial}{\partial q^k} + g^{ij}\frac{\partial}{\partial q^i \partial q^j})) = -\frac{1}{2}g^{ij}\Gamma^k_{ij}\frac{\partial}{\partial q^k} + \frac{1}{2}g^{ij}\Gamma^k_{ij}\frac{\partial}{\partial q^k} = 0.$ □

Definition 7.75. A Riemannian manifold M is called *stochastically complete* if, for each $m_0 \in M$, every Wiener process beginning at m_0 a.s. extends to $[0, \infty)$.

On a stochastically complete manifold a Wiener process beginning at m_0 gives rise to a measure on the space of continuous curves in M which, in turn, begin at m_0. It is not hard to see that this measure is actually independent of the choice of Wiener process. The coordinate process on the space of such curves is just the standard Wiener process on M beginning at m_0. On the other hand, the measure on the space of curves is not uniquely defined if with a non-zero probability the development $R_I w$ goes to infinity in finite time. In other words, the measure depends on the behavior of $R_I w$ at infinity, i.e., on the geometry of M. (See [66, 138].)

Note that there exist stochastically complete manifolds that are not complete in the usual sense: if we exclude a single point from $\mathbb{R}^n$ it will become incomplete but for every Wiener process $w_x(t) = x + w(t)$ (starting from x at $t = 0$) the sample paths will still a.s. exist for $t \in [0, +\infty)$ since the probability of hitting the excluded point equals zero. Note nevertheless that the flow of a Wiener processes constructed in such a way in $\mathbb{R}^n$ by removing a point will not be strongly complete (see Definition 6.30).

Note also that ordinary completeness is insufficient for the stochastic completeness of a Riemannian manifold.

Applying Theorem 7.43 one can obtain some sufficient conditions for stochastic completeness of a manifold. Denote by $\mathcal{A}^w$ the generator of the flow on OM given by equation (7.40) (or (7.39)). It is not hard to see that in a chart on OM this operator is presented in the form

$$\mathcal{A}^w_b = -\frac{1}{2}\mathrm{tr}\varGamma^{\mathsf{e}}_b(I, I) + \frac{1}{2}a^{ij}\frac{\partial^2}{\partial q^i \partial q^j}, \tag{7.41}$$

where (a^{ij}) is the matrix of the operator $\mathsf{E}_b\mathsf{E}^*_b$.

Theorem 7.76 *If on OM there exists a proper function φ such that the values of the generator $\mathcal{A}^w$ on φ are uniformly bounded, the Riemannian manifold M is stochastically complete.*

This statement follows immediately from Theorem 7.43.

Corollary 7.77 *A compact Riemannian manifold is stochastically complete.*

Indeed, since the group $O(n)$ is compact, from the compactness of M it follows that OM is compact, i.e., M is stochastically complete.

To avoid contradicting some statements in [66] concerning stochastic completeness, we emphasize that we are dealing with the usual Riemannian manifolds, i.e., the manifolds with positive definite Riemannian inner product in $T_m M$ at every point $m \in M$. In this case the group preserving the inner product is the orthogonal group and is compact. If a manifold is semi-Riemannian, the corresponding group that preserves the semi-Euclidean inner product is not compact. Thus, the principal bundle with this group over a compact manifold is not compact either.

The next statement is proved by applying Theorem 7.13 to equation (7.36) on OM.

Theorem 7.78 *Assume that on OM there exists a Riemannian metric possessing a uniform Riemannian atlas and such that the Itô vector field* $(0, \mathsf{E}_b)$ *is uniformly bounded on OM in the norm of the space C^1 generated by this metric. Then the Riemannian manifold M is stochastically complete.*

Corollary 7.79 *A Lie group with left (right) invariant metric is stochastically complete.*

Indeed, the Levi-Civitá connection and basic vector fields $\mathsf{E}_b(x)$ are left (right, respectively) invariant. A uniform Riemannian atlas on the group can be constructed by left (right, respectively) translations of a chart in a neighborhood of the unit element to the points of group. Define in the algebra $\mathfrak{o}(n)$ an inner product and translate it by left (right, respectively) shifts into all points of $O(n)$. For the constructed Riemannian metric on OM a uniform Riemannian atlas can be also constructed by translations of some chart in a neighborhood of the unit. Since the fibers of OM are isomorphic to $O(n)$, we obtain an induced metric on OM (see Definition 7.58 and Remark 7.59) and by construction for this metric and the field $\mathsf{E}_b(x)$ the hypothesis of Theorem 7.78 is fulfilled.

Since the fibers of OM are compact, the proper functions on OM can be constructed from proper functions on M: if φ is a proper function on M, $\hat{\varphi} = \varphi \circ \pi : OM \to R$ is a proper function on OM. It is not hard to see that if on M there is a proper function φ such that $\nabla^2\varphi$ is uniformly bounded (here ∇^2 is the Laplace-Beltrami operator, i.e., by Theorem 7.73 the generator of the Wiener process), for $\hat{\varphi}$ the hypothesis of Theorem 7.76 is fulfilled.

Theorem 7.80 (Elworthy [66]) *If the Ricci curvature of a complete Riemannian manifold M is bounded from below, M is stochastically complete.*

Proof. It is shown in Yau's paper [234] that under the hypothesis of this Theorem there exists a proper function α (constructed from the Riemannian distance) on M, for which $\nabla^2\alpha < C$ for some $C > 0$. A more complete proof can be found in [66]. □

Theorem 7.81 (Grigoryan [138]) *Let M be a complete Riemannian manifold and let $V(r)$ denote the volume of the metric ball with radius r with center $m_0 \in M$. If the condition $\int^\infty \frac{r\mathrm{d}r}{V(r)} = \infty$ is satisfied, M is stochastically complete.*

Notice that if the hypothesis of Theorem 7.80 is fulfilled, $V(r) < e^{Cr}$ where $C > 0$ is a constant and so the hypothesis of Grigoryan's theorem is satisfied. Moreover, the hypothesis of the latter theorem is fulfilled if $V(r) < \mathrm{e}^{Cr^2}$, $V(r) < \mathrm{e}^{Cr^2 \ln r}$, etc. If for some positive function f, regular in some sense, the estimate $\int^\infty \frac{r\mathrm{d}r}{f(r)} < \infty$ is fulfilled, there exists a complete Riemannian manifold for which $V(r) < Cf(r)$ but it is not stochastically complete. A discussion of these questions can be found in Grigoryan's paper [138].

We also refer the reader to another paper of Grigoryan, [139], where additional deep results concerning stochastic completeness are obtained.

Some criteria of stochastic completeness based on the use of atlases which are in some sense uniform, and on the existence of a function with special properties on M, can be found in [66].

7.6.3 Parallel translation along a stochastic process. Itô processes on manifolds

Definition 7.82. A stochastic process on M is called an *Itô process* if it is an Itô development of an Itô process in some tangent space.

Using the transformation of equations of Itô type into equations of Stratonovich type and the expression of Itô processes in tangent spaces via Stratonovich integrals it is possible to present Itô processes on manifolds as the Eells-Elworthy development of a process with Stratonovich integral as in Definition 7.60. For simplicity of presentation here we restrict ourselves to the use of Itô developments.

It is possible to define the Riemannian parallel translation along Itô processes on M analogously to the standard (i.e., non-stochastic) construction. Let $\eta(t)$ be an Itô process given for $t \in [0, l]$ and let $\hat{\xi}(t)$ be a horizontal lift of $\eta(t)$ to OM with the initial condition $\hat{\xi}(0) = \hat{b}$. Let $v \in T_{\eta(0)}M$ be a (random) vector.

Definition 7.83. The *parallel translation* of the vector v along $\eta(\cdot)$ at the point $\eta(t)$ is the vector $(\hat{\xi}(t) \circ (\hat{b}^{-1}v)) \in T_{\eta(t)}M$.

It is obvious that parallel translation preserves the Riemannian norms and does not depend on the choice of horizontal lift $\hat{\xi}$. From the construction of the process $R_I y(t)$ where $y(t) = \int_0^t \alpha(\tau)\mathrm{d}\tau + \int_0^t A(\tau)\mathrm{d}w(\tau)$ in $T_{m_0}M$ it is clear, in particular, that the vector $T\pi E_{\xi(t)}(\hat{b}^{-1}\alpha(t))$ and the operator $T\pi E_{\hat{\xi}(t)}(\hat{b}^{-1}A(t))$ from Lemma 7.65 are parallel along $\eta(t) = R_I y(t)$, respectively, to the vector $\alpha(t)$ and to the operator $A(t)$.

Remark 7.84. We mention the papers [58, 66, 151, 152] where the parallel translation is constructed along stochastic processes of various sorts. It is not hard to see that Itô processes in the sense of Definition 7.82 are local quasi-martingales (i.e., special semi-martingales, see [176]) with continuous sample paths, along which the parallel translation is constructed in [151], and the parallel translations in the sense of [151] and in the sense of Definition 7.83 coincide. From this it follows, in particular, that the parallel translation along an Itô process is the limit in probability of trajectory-wise parallel translations (in the ordinary "piecewise smooth" sense) along processes whose sample paths are piecewise geodesic approximations of the trajectories of the process $\eta(t)$ (see [151]). It is clear that one can always select a sub-sequence of these approximations such that the trajectory-wise parallel translations converge almost surely. Hence the parallel translation along $\eta(t)$ as in Definition 7.83 is an extension of the classical parallel translation from the set of piecewise smooth curves to a.s. all sample paths of $\eta(t)$ that are continuous but a.s. not smooth (cf. Remarks 7.66 and 7.67).

7.7 The Integral Approach to Stochastic Differential Equations on Manifolds

In this Section we present an integral description of Itô stochastic differential equations, analogous to the classical case in linear spaces as presented in Section 6.2. The construction of the required integral operators is a stochastic modification of the integral operators with Riemannian parallel translation from Section 3.2 based on constructions from Section 7.6. Then, making use of parallel translation along a stochastic process, we modify the notion of a stochastic differential equation of Itô type and describe a broader class of equations for which the notion of a non-degenerate (in particular, unit) diffusion coefficient is well-defined.

7.7.1 General constructions

Let $y(t) = \int_0^t b(\tau)\mathrm{d}\tau + \int_0^t B(\tau)\mathrm{d}w(\tau)$ be an Itô process in the tangent space $T_{m_0}M$ to the Riemannian manifold M as in Section 7.6. Introduce for the corresponding Itô process $R_I y(t)$ on M the new notation $R_I y(t) = \mathcal{S}(b(\tau), B(\tau))(t)$. One can easily see that the operator $\mathcal{S}$ introduced in this way is a stochastic analog of the operator $\mathcal{S}$ with parallel translation from Section 3.2.

Let $\xi(t)$ be an Itô process that a.s. exists for $t \in [0,l]$ and $(\alpha(t,m), A(t,m))$ be an Itô vector field on M, $t \in [0,l]$. As in Section 2.2 (see Theorem 2.32), denote by $\Gamma_{t,s}$ the operator of parallel translation along $\xi(\cdot)$ from the random point $\xi(s)$ to the random point $\xi(t)$. For the sake of simplicity, if $t = 0$, i.e., if the translation is performed at the point $\xi(0)$ rather than $\Gamma_{0,s}$, we shall often use the notation Γ analogous to that from Section 3.2.2. Thus $\Gamma\alpha(t,\xi(t))$ is the random vector in $T_{\xi(0)}M$ obtained by parallel translation of the random vector $\alpha(t,\xi(t))$ along $\xi(\cdot)$ at the point $\xi(0)$. Analogously, $\Gamma A(t,\xi(t))$ is the random operator sending $\mathbb{R}^n$ to $T_{\xi(0)}M$ that is obtained by parallel translation of $A(t,\xi(t))$ along $\xi(\cdot)$ at $\xi(0)$.

Let an Itô vector field $(\alpha(t,m), A(t,m))$ be Borel measurable. Consider the processes $\Gamma\alpha(t,\xi(t))$ and $\Gamma A(t,\xi(t))$. Using the properties of horizontal lift, i.e., of a strong solution of equation (7.37), it is not hard to show that these processes are non-anticipative with respect to the family $\mathcal{B}_t$ that is used in the definition of the Itô process $\xi(t)$. Consider equation (7.37) on OM, in which α is replaced by $\Gamma\alpha$ and A by ΓA:

$$\mathrm{d}\xi(t) = \mathsf{e}_{\xi(t)}\left(\mathsf{E}_b\left(b_0^{-1}\Gamma\alpha(t)\right)\mathrm{d}t + \mathsf{E}_b\left(b_0^{-1}\Gamma A(t)\mathrm{d}w(t)\right)\right). \tag{7.42}$$

For further developments it is important to present conditions under which (7.42) has global solutions. This is the case, if, say, on OM there exists a proper function φ such that the values of the generator of the Itô equation on OM, corresponding to (7.42), are uniformly bounded.

In [107] the existence of global solutions of (7.42) is proved for $\alpha(t)$ and $A(t)$ (and hence for $\Gamma\alpha(t)$ and $\Gamma A(t)$ since the parallel translation preserves the norms) uniformly bounded under the additional assumption that M is a uniformly complete Riemannian manifold.

Definition 7.85. A Riemannian manifold M is said to be *uniformly complete* if the following two conditions are satisfied:

(1) there exists an induced metric on OM which possesses a uniform Riemannian atlas;
(2) on the balls $V_b(r)$ of the atlas, the norm of the operator $X \mapsto \Gamma^e_{b'}(X, X)$, where $X \in H_{b'}$ and $b' \in V_b(r)$, as a norm of a quadratic operator, is bounded by a constant $C > 0$ independent of the chart and the ball.

Evidently compact Riemannian manifolds and Lie groups are examples of uniformly complete Riemannian manifolds but the latter class of manifolds is much broader than these two examples.

For uniformly complete manifolds and uniformly bounded $\Gamma\alpha(t)$ and $\Gamma A(t)$ the solvability of (7.42) follows from Theorem 7.38.

Let, for any given initial data, equation (7.42) have a strong and strongly unique solution that is well-defined for all $t \in [0, l]$. It is clear that the projection of this solution to M is the Itô process $\mathcal{S}(\Gamma\alpha(\tau, \xi(t)), \Gamma A(\tau, \xi(\tau)))(t)$.

Definition 7.86. The Itô process $\mathcal{S}(\Gamma\alpha(\tau, \xi(\tau)), \Gamma A(\tau, \xi(\tau)))(t)$ on M is called the *line Itô integral with Riemannian parallel translation* of the field (α, A) along $\xi(t)$.

$\mathcal{S}(\Gamma\alpha(\tau, \xi), \Gamma A(\tau, \xi))(t)$ is a direct analog of the ordinary line integral used in the theory of Itô stochastic differential equations in Euclidean spaces. If M is a Euclidean space, Γ is the identical mapping and

$$\mathcal{S}(\Gamma\alpha(\tau, \xi), \Gamma A(\tau, \xi))(t) = \int_0^t \alpha(\tau, \xi)\mathrm{d}\tau + \int_0^t A(\tau, \xi)\mathrm{d}w(\tau).$$

Like its classical analog, the integral $\mathcal{S}(\Gamma\alpha(\tau, \xi), \Gamma A(\tau, \xi))(t)$ is naturally connected with the Itô equations.

Consider an Itô vector field $(\alpha(t, m), A(t, m))$ on M and the corresponding equation in Belopolskaya-Daletskii form

$$\mathrm{d}\eta(t) = \exp_{\eta(t)}(\alpha(t, \eta(t))\mathrm{d}t + A(t, \eta(t))\mathrm{d}w(t)),\ \eta(0) = m_0, \tag{7.43}$$

where exp is the exponential mapping of the Levi-Civitá connection. It turns out that its solution is an Itô process that satisfies the equation

$$\eta(t) = \mathcal{S}(\Gamma\alpha(\tau, \eta(\tau)), \Gamma A(\tau, \eta(\tau)))(t). \tag{7.44}$$

Indeed, construct a horizontal lift of $\eta(t)$ in the following way. Introduce an Itô vector field $(\bar{\alpha}(t,b), \bar{A}(t,b))$ on OM by the formulae $\bar{\alpha}(t,b) = T\pi^{-1}\alpha(t,\pi b)_{|H_b}$ and $\bar{A}(t,b) = T\pi^{-1}A(t,\pi b)_{|H_b}$. For every $\hat{b} \in O_m(M)$ let there exist a unique strong global solution $\hat{\xi}(t)$ of the equation

$$\mathrm{d}\hat{\xi}(t) = \mathsf{e}_{\hat{\xi}(t)}\left(\bar{\alpha}\left(t, \hat{\xi}(t)\right)\mathrm{d}t + \bar{A}\left(t, \hat{\xi}(t)\right)\mathrm{d}w(t)\right), \qquad \hat{\xi}(0) = \hat{b}.$$

Consider the processes $\alpha(t) = \hat{b}(\hat{\xi}(t)^{-1}\,\alpha(t, \pi\hat{\xi}(t)))$ in $T_{m_0}M$ and $A(t) = \hat{b}(\hat{\xi}(t)^{-1}\,A(t, \pi\hat{\xi}(t)))$ in $L(\mathbb{R}^n, T_{m_0}M)$. By construction we obtain that $\hat{\xi}(t)$ is a solution of equation (7.37) for $\alpha(t)$ and $A(t)$. Thus $\hat{\xi}(t)$ is a horizontal lift of $\eta(t)$. So, $\alpha(t) = \Gamma\alpha(t, \eta(t)), A(t) = \Gamma A(t, \eta(t))$ and for $\eta(t)$ relation (7.44) is satisfied.

Equation (7.44) is an analog of the integral form of the Itô equation in Euclidean space (see Section 6.2.3). For (7.44) the usual notion of strong and weak solutions are introduced as in Section 6.2.3.

Theorem 7.87 *Let equation* (7.44) *have a weak solution* $\eta(t)$*. Then* $\eta(t)$ *is a weak solution of equation* (7.43)*.*

Proof. Let $\hat{\xi}(t)$ be the horizontal lift of $\eta(t)$ with initial condition $\hat{\xi}(0) = \hat{b} \in O_{m_0}(M)$. From (7.44) and from the construction of the operator $\mathcal{S}$ it follows that for every $t \in [0,l]$ a.s. $\alpha(t, \eta(t)) = T\pi E_{\hat{\xi}(t)}(\hat{b}^{-1}\Gamma\alpha(t, \eta(t)))$ and $A(t, \eta(t)) = T\pi E_{\hat{\xi}(t)}(\hat{b}^{-1}\Gamma A(t, \eta(t)))$. Hence by Lemma 7.65 equation (7.43) is satisfied for $\eta(t)$ a.s. for all $t \in [0,l]$. All other conditions of Definition 7.82 are fulfilled by the hypothesis since $\eta(t)$ is a weak solution of (7.44). □

Corollary 7.88 *If* $\eta(t)$ *is a strong solution of* (7.44)*,* $\eta(t)$ *is a strong solution of* (7.43)*.*

Proof. As in Theorem 7.87 it is proven that $\eta(t)$ satisfies (7.43). Here all requirements of Definition 6.23 are fulfilled since $\eta(t)$ is a strong solution of (7.44). □

As in the case of ordinary differential equations (see Section 3.2) the use of integral operators with parallel translation allows one to reduce some questions to the investigation of stochastic differential equations in a single tangent space.

Let $(\alpha(t,m), A(t,m))$ be an Itô vector field on M, $t \in [0,l]$. In $T_{m_0}M$ consider the stochastic differential equation

$$z(t) = \int_0^t \Gamma\alpha(\tau, R_I z(\tau))\mathrm{d}\tau + \int_0^t \Gamma A(\tau, R_I z(\tau))\mathrm{d}w(\tau) \tag{7.45}$$

Theorem 7.89 *An Itô process $z(t)$ in $T_{m_0}M$ is a strong (weak) solution of* (7.45) *if and only if $\eta(t) = R_I z(t)$ is a strong (weak, respectively) solution of* (7.43).

Proof. Assume that a Wiener process $w(t)$ with values in $\mathbb{R}^n$ and an Itô process $z(t) = \int_0^t \alpha(\tau)\mathrm{d}\tau + \int_0^t A(\tau)\mathrm{d}w(\tau)$ with values in $T_{m_0}M$ are given on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and that they satisfy (7.45). Then by Lemma 7.65, by the construction of the mapping E and by the Definition 7.82 of parallel translation we obtain that for all $t \in [0, l]$ equality (7.43) a.s. holds for $R_I z(t)$. Conversely, let $z(t)$ and $w(t)$ be such that the development $\eta(t) = R_I z(t)$ for all t a.s. satisfies (7.43). By the construction of the development there is a horizontal lift $\hat{\xi}(t)$ of $\eta(t)$, $\hat{\xi}(0) = \hat{b} \in O_{m_0}M$ and along $\eta(\cdot)$ the parallel translation is well-defined. From Lemma 7.65 and equation (7.43) it follows that for all $t \in [0, l]$ a.s. $T\pi E_{\hat{\xi}(t)}(\hat{b}^{-1}\alpha(t)) = \alpha(t, \eta(t))$ and $T\pi E_{\hat{\xi}(t)}(\hat{b}^{-1}A(t)) = A(t, \eta(t))$. Applying parallel translation along $\eta(\cdot)$ at the point m_0 to the latter equality, we obtain that for all $t \in [0, l]$ a.s. $\alpha(t) = \Gamma\alpha(t, \eta(t))$ and $A(t) = \Gamma A(t, \eta(t))$. From this it follows that $z(t)$ and $w(t)$ satisfy (7.45). It is easy to see that $z(t)$ and $w(t)$ are non-anticipative with respect to the common family of σ-subalgebras: $\mathcal{P}_t^w$ in the case of a strong solution and $\mathcal{B}_t$ (to which $z(t)$ is adapted) in the case of a weak one. □

As an example of the application of equation (7.45) to the investigation of equation (7.43) we present a statement on the existence of a weak solution of (7.43).

Theorem 7.90 *Let an Itô vector field $(\alpha(t, m), A(t, m))$, $t \in [0, l]$, on a uniformly complete Riemannian manifold M (see Definition 7.85) be jointly continuous in t and m and uniformly bounded in the norm generated by the Riemannian metric. Then for every initial condition $m_0 \in M$ equation* (7.43) *has a weak solution.*

Proof. We use the martingale approach to the construction of solutions [83, 84, 162]. In the case under consideration we need some preliminary constructions that take into account the specific features of the equations. Since (α, A) is uniformly bounded on $[0, l] \times M$, one can easily construct a sequence of smooth approximations (α_i, A_i) that converge uniformly on $[0, l] \times M$ to (α, A). Note that all (α_i, A_i) are uniformly bounded by a common constant since (α, A) is uniformly bounded. Let η_i be a strong solution of the equation

$$\mathrm{d}\eta_i(t) = \exp_{\eta_i(t)}(\alpha_i(t, \eta_i(t))\mathrm{d}t + A_i(t, \eta_i(t))\mathrm{d}w(t)), \quad \eta_i(0) = m_0,$$

that exists by Theorem 7.36. From the above statements it follows that the processes $z_i(t) = \int_0^t \Gamma\alpha_i(\tau, \eta_i(\tau))\mathrm{d}\tau + \int_0^t \Gamma A_i(\tau, \eta_i(\tau))\mathrm{d}w(\tau)$ are strong solutions of the equations

$$z_i(t) = \int_0^t \Gamma\alpha_i(\tau, R_I z_i(\tau))\mathrm{d}\tau + \int_0^t \Gamma A_i(\tau, R_I z_i(\tau))\mathrm{d}w(\tau).$$

Let $\tilde{\Omega} = C^0([0,l], T_{m_0}M)$ be the Banach space of continuous mappings from the interval $[0,l]$ to $T_{m_0}M$ (i.e., continuous curves in $T_{m_0}M$), $\tilde{\mathcal{F}}$ be the σ-algebra in $\tilde{\Omega}$ generated by cylinder sets and $\mathcal{B}_t$ be the σ-subalgebra generated by cylinder sets with bases on $[0,t]$, $t \in [0,l]$ (cf. Sections 6.1.1 and 6.2.1). Recall that all σ-algebras are assumed to be complete (contain all sets of measure zero). Denote by μ_i the probability measure on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ generated by the process z_i. Consider the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu_i)$ where the elementary events are continuous curves $x(\cdot) \in C^0([0,l], T_{m_0}M)$ and the realization of z_i as the coordinate process on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu_i)$: $z_i(t, x(\cdot)) = x(t)$. Note that the coordinate process is not anticipative with respect to $\mathcal{B}_t$. Taking into account Remarks 7.67 and 7.84 we obtain that for a continuous curve $x(\cdot) \in \tilde{\Omega}$ the processes $R_I x(t)$ and $Y(R_I x(t))$ for $Y \in T_{R_I x(t)}M$ are μ_i-a.s. well-defined since R_I and Γ are extensions of the inverse of Cartan's development and parallel translation, respectively, from the set of smooth curves to μ_i-a.s. all continuous curves (sample paths of z_i and $R_I z_i$). This is true for every i, i.e., for every j and for all measures μ_i the processes $\Gamma\alpha_j(t, R_I x(t))$ and $\Gamma A_j(t, R_I x(t))$ are μ_i-a.s. well-defined. From the uniform convergence of (α_i, A_i) to (α, A) and from the properties of parallel translation it follows that $\Gamma\alpha_j(t, R_I x(t))$ and $\Gamma A_j(t, R_I x(t))$ converge as $j \to \infty$, μ_i-a.s. uniformly in t for all i, to $\Gamma\alpha(t, R_I x(t))$ and $\Gamma A(t, R_I x(t))$, respectively. From the fact that the fields (α_i, A_i) are uniformly bounded by a common constant one can easily deduce that the set of measures μ_i is weakly compact.

Let μ be a limit measure. Consider the coordinate process $z(t, x(\cdot)) = x(t)$ on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$. By construction $z(t)$ is not anticipative with respect to $\mathcal{B}_t$. Using Prokhorov's Theorem 6.6 one can easily show that the processes $\Gamma\alpha_j(t, R_I x(t))$ and $\Gamma A_j(t, R_I x(t))$ are μ-a.s. well-defined and μ-a.s. converge uniformly in t to $\Gamma\alpha(t, R_I x(t))$ and $\Gamma A(t, R_I x(t))$, respectively. The concluding arguments are exactly the same as in the classical existence of weak solution theorem for equations with continuous coefficients [83]. Using the above-mentioned convergencies a Wiener process $\tilde{w}(t)$, adapted to $\mathcal{B}_t$, is constructed on $(\tilde{\Omega}, F, \mu)$ so that $z(t)$ and $\tilde{w}(t)$ satisfy (7.45) for all t almost surely. By Theorem 7.89, $R_t z(t)$ is a weak solution of (7.43). □

Remark 7.91.

(i) In the constructions and applications of operators with parallel translation in this section we have not used the fact that the torsion of the Levi-Civitá connection equals zero. Thus all constructions and applications remain true if we use an arbitrary Riemannian connection on M under the condition analogous for that connection to, say, the condition of unform completeness. Note that for some special choice of connection on a Lie group the above-mentioned constructions yield the well-known multiplicative integral.

(ii) Let $(\hat{\alpha}, A)$ be an Itô equation (cross-section of an Itô bundle). Recall that its solution is described by equation (7.13). Using different Riemannian metrics and connections on M, one can describe this equation in the form (7.43) with corresponding (strictly speaking, different) Itô vector fields that canonically correspond to $(\hat{\alpha}, A)$ with respect to the chosen connections. Integral operators with parallel translation and the presentation of the equation $(\hat{\alpha}, A)$ in integral form (7.44) or (7.45) also depend on the choice of metric and connection. We emphasize that a solution of (7.43) or of (7.44) does not depend on the choice of metric and connection since this solution is a solution of (7.13).

7.7.2 Stochastic differential equations in terms of Wiener processes in tangent spaces

Let a be a vector field and A be a (1,1)-tensor field on M, i.e., A_m is a linear operator $T_mM \to T_mM$ for every $m \in M$. Note that the fields may be time-dependent and in this case we shall denote them by $a_{t,m}$ and $A_{t,m}$, respectively. Using these fields we construct a modification of equation (7.18) in the following way. Assume that in every tangent space T_mM a realization $w_m(t)$ of a Wiener process is given. Then the equation

$$\mathrm{d}\xi(t) = \exp_{\xi(t)} \left(a_{t,\xi(t)} \mathrm{d}t + A_{t,\xi(t)} \mathrm{d}w(t) \right) \tag{7.46}$$

is well-defined where exp is the exponential mapping of the Levi-Civitá connection. Indeed, let b_m be a field of orthonormal frames that determines the realizations $w_m(t) = b_m\, w(t)$ of a Wiener process in tangent spaces. Then $(a_{t,m}, A_{t,m}\, b_m)$ is an Itô vector field (unlike the pair (a, A)).

We have changed the notation for equations of type (7.46) to avoid confusion with equations of type (7.18).

For equations of type (7.46) it is necessary to modify the notion of solution.

Definition 7.92. We say that equation (7.46) has a *strong solution* $\xi(t)$ if, for any Wiener process $w(t)$ in $\mathbb{R}^n$, there is a process $\xi(t)$ in M defined on the same probability space as $w(t)$ and non-anticipative with respect to $\mathcal{P}_t^w$, and there is a realization $b_{\xi(t)}\, w(t)$ of the Wiener process at $\xi(t)$ such that the processes $w_{\xi(t)}(t) = b_{\xi(t)}\, w(t)$ and $\xi(t)$ satisfy (7.46) for every t.

Definition 7.93. Equation (7.46) has a *weak solution* if there are:

(1) a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ with a non-decreasing family $\mathcal{B}_t$ of complete σ-subalgebras of $\mathcal{F}$;
(2) a stochastic process $\xi(t)$ on M, non-anticipative with respect to $\mathcal{B}_t$;
(3) a Wiener process $w(t)$ in $\mathbb{R}^n$ relative to $\mathcal{B}_t$;
(4) realizations $w_\xi(t) = b_{\xi(t)}\, w(t)$ of $w(t)$ in $T_{\xi(t)}M$,

such that $w_{\xi(t)}(t)$ and $\xi(t)$ a.s. satisfy (7.46) for every t, as in Definition 7.28.

Using (7.19) and Definitions 7.92 and 7.93 it is easy to show that in a chart equation (7.46) takes the form

$$\mathrm{d}\xi(t) = a\big(t, \xi(t)\big)\,\mathrm{d}t - \frac{1}{2}\mathrm{tr}\,\boldsymbol{\Gamma}_{\xi(t)}(A, A)\,\mathrm{d}t + A_{\xi(t)}\,\big(b_{\xi(t)}\mathrm{d}w(t)\big), \tag{7.47}$$

where $\boldsymbol{\Gamma}_m(\cdot\,\cdot)$ is the local connector and $\mathrm{tr}\,\boldsymbol{\Gamma}_m(A, A) = \mathrm{tr}\,\boldsymbol{\Gamma}_m(A_m b_m, A_m b_m)$ where $b_m \in O_m(M)$ is a cross-section of OM, i.e., a field of orthonormal frames $b_m : \mathbb{R}^n \to T_m M$, $m \in \mathfrak{U}$. Observe that $\mathrm{tr}\,\boldsymbol{\Gamma}_m(A, A)$ is independent of b (consistently with the notation), because the trace is invariant under the action of the orthogonal group. Using this fact and the results of Section 7.3, it is not hard to show that (7.47) is covariant with respect to changes of coordinates. Below $b_{\xi(t)}$ will usually arise as the horizontal lift of $\xi(t)$.

By the definition of a Wiener process on a Riemannian manifold (see Section 7.6.2), we have:

Theorem 7.94 *A Wiener process $\tilde{w}(t)$ on M is a strong solution of the equation $\mathrm{d}\tilde{w}(t) = \exp_{\tilde{w}(t)}(\mathrm{d}w)$. In local coordinates, this equation reads $\mathrm{d}\tilde{w}(t) = -\frac{1}{2}\mathrm{tr}\,\Gamma_{\tilde{w}(t)}(I, I)\,\mathrm{d}t + b_{w(t)}\mathrm{d}w(t)$.*

Assume that M is uniformly complete. To construct the integral operators needed to study (7.46), we have to alter the construction of Section 7.7.1.

Let the probability space $(\Omega, \mathcal{F}, P)$, the family $\mathcal{B}_t$, the manifold M, and the functions $a(t)$ in $T_{m_0}M$ and $A(t)$ in $L(T_{m_0}M)$ be as in Section 7.7.1. Specify a realization $b\,w$ of the Wiener process w in $T_{m_0}M$. It is clear that the operator $\mathcal{S}$ from Section 7.7.1 is applicable to the pair (a, Ab).

Let $a(t, m)$ and $A(t, m)$ be a vector field and a $(1,1)$-tensor field on M, respectively, and let $\eta(t)$, $t \in [0, l]$, be an Itô process on M. Consider the vector and tensor fields $\Gamma a\big(t, \eta(t)\big)$ and $\Gamma A\big(t, \eta(t)\big)$ obtained by the parallel translation of $a\big(t, \eta(t)\big)$ and, respectively, $A\big(t, \eta(t)\big)$ along $\eta(\cdot)$ to $\eta(0)$. The operator $\mathcal{S}$ can be applied to the pair $\big(\Gamma a\big(t, \eta(t)\big), \Gamma A\big(t, \eta(t)\big)\big)$, provided the fields $a(t, m)$ and $A(t, m)$ are bounded and Borel measurable jointly in t and m.

Therefore, we can define the Itô integral and the line integral with parallel translation in terms of the field of Wiener processes. To distinguish these integrals from those of Section 7.7.1, we denote them by $\mathcal{S}\big(a(\tau)\,\mathrm{d}\tau + A(\tau)\,\mathrm{d}w(\tau)\big)(t)$ and $\mathcal{S}\big(\Gamma a\big(\tau, \eta(\tau)\big)\,\mathrm{d}\tau + \Gamma A\big(\tau, \eta(\tau)\big)\,\mathrm{d}w(\tau)\big)(t)$, respectively.

Then (7.44) is to be replaced by the following equation

$$\xi(t) = \mathcal{S}\big(\Gamma a\big(\tau, \xi(\tau)\big)\,\mathrm{d}\tau + \Gamma A\big(\tau, \xi(\tau)\big)\mathrm{d}w(\tau)\big)(t). \tag{7.48}$$

Let $b_0 w$ be the initial realization of the Wiener process in $T_{m_0}M$. Observe that the parallel translation of b_0 along a solution $\xi(\cdot)$ of (7.48) gives rise to a realization of the Wiener process at $\xi(t)$. (See Definitions 7.92 and 7.93.)

The equation

$$z(t) = \int_0^t \Gamma a\big(\tau, R_I z(\tau)\big)\,\mathrm{d}\tau + \int_0^t \Gamma A\big(\tau, R_I z(\tau)\big)\,\mathrm{d}w_{m_0}(\tau) \tag{7.49}$$

is an analog of (7.45). Similar results to Theorems 7.36, 7.87 and 7.90 hold for equations (7.46), (7.48) and (7.49).

7.7.3 Equations with unit diffusion coefficients

Generalizing the classical notion, it is natural to call (7.46) an equation with a non-degenerate diffusion coefficient if the operator $A_{(t,m)} \colon T_mM \to T_mM$ is non-degenerate for all $m \in M$ and $t \in [0,l]$ (cf. Definition 6.18).

Among equations with non-degenerate diffusion coefficients we are especially interested in those with $A = I$, i.e., with the diffusion coefficient equal to the identity operator. Then the equation can be written down in the form:

$$\mathrm{d}\xi(t) = \exp_{\xi(t)}\left(a_{t,\xi(t)}\,\mathrm{d}t + \mathrm{d}w(t)\right). \tag{7.50}$$

Note that in this form the equations with a smooth field of linear operators $A(m) : \mathbb{R}^k \to T_mM$ such that $A(m)A^*(m) = I$ can also be represented.

Solutions of (7.50) will play a crucial role in Chapter 15 and in Section 14.4. Note that for (7.50) the local expression (7.47) turns into the equality

$$\mathrm{d}\xi(t) = a\big(t,\xi(t)\big)\,\mathrm{d}t - \frac{1}{2}\mathrm{tr}\,\Gamma_{\xi(t)}(I,I)\,\mathrm{d}t + b_{\xi(t)}\mathrm{d}w(t).$$

Theorem 7.95 *Assume that the Riemannian manifold M is stochastically complete and the vector field $a(t,m)$ is Borel measurable jointly in $(t,m) \in [0,l] \times M$ and uniformly bounded. Then there exists a weakly unique weak solution $\xi(t)$ of (7.50) for any initial condition $\xi(0) = m_0$ that is well-defined on $[0,l]$.*

Proof. Here, we are using a method based on a change of probability measure [83, 84, 162]. Consider the standard Wiener process $\tilde{w}$ on $T_{m_0}M$, i.e., the coordinate process $\tilde{w}\big(t, x(\cdot)\big) = x(t)$ on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \nu)$, where $\tilde{\Omega} = C^0\big([0,l], T_{m_0}M\big)$, $\tilde{\mathcal{F}}$ is the σ-algebra generated by cylinder sets, and ν is the Wiener measure. Recall that the elementary event in $\tilde{\Omega}$ is a continuous curve $x(\cdot) \in C^0\big([0,l], T_{m_0}M\big)$. Observe that $\tilde{w}(t)$ is non-anticipative with respect to the family of σ-subalgebras $\mathcal{B}_t$ generated by the cylinder sets with bases over $[0,t]$, $t \in [0,l]$. (See Sections 6.2.1 and 7.7.)

Since M is stochastically complete, the development $R_I\tilde{w}(t)$ is well-defined. Taking into account Remarks 7.67 and 7.84, we see that $R_Ix(t)$ and $\Gamma a\big(t, R_Ix(t)\big)$ exist for ν-almost all $x \in \tilde{\Omega}$. Furthermore, it follows from the properties of parallel translation, of R_I and of $a(t,m)$ that the stochastic process $\Gamma a\big(t, R_I\tilde{w}(t)\big)$ is uniformly bounded and non-anticipative with respect to $\mathcal{B}_t$. Consider the measure μ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ with density ρ with respect to ν given by

$$\rho\,(x(\cdot)) = \exp\left(\int_0^l \langle \Gamma a\,(t, R_I x(t))\,, \mathrm{d}\tilde{w}(t)\rangle - \frac{1}{2}\int_0^l \Gamma a\,(t, R_I x(t))^2\ \mathrm{d}t\right). \tag{7.51}$$

It is known (see [83, 162]) that under the hypotheses of the theorem

$$\int_{\tilde{\Omega}} \rho\,\mathrm{d}\nu = 1, \tag{7.52}$$

i.e., μ is a probability measure, and, furthermore, $w\big(t, x(\cdot)\big) = x(t) - \int_0^t \Gamma a\big(\tau, R_I x(\tau)\big)\,\mathrm{d}\tau$ is a Wiener process on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$ relative to $\mathcal{B}_t$. It is not hard to show that $\rho > 0$ everywhere, i.e., ν is absolutely continuous with respect to μ and has density ρ^{-1}. In other words, the probability measures μ and ν are equivalent. The coordinate process $z\big(t, x(\cdot)\big) = x(t)$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$ is non-anticipative with respect to $\mathcal{B}_t$ and, moreover, $\mathcal{B}_t = \mathcal{P}_t^z$. Since the measures μ and ν are equivalent, $R_I(z(t, x(\cdot)) = R_I x(t)$ exists μ-a.s. Thus, $z(t)$ and $w(t)$ are related via the equation

$$\mathrm{d}z(t) = \Gamma a\,(t, R_I z(t))\ \mathrm{d}t + \mathrm{d}w \tag{7.53}$$

on $T_{m_0}M$. In other words, $z(t)$ is a weak solution of (7.53). It is shown in [83, 162] that every solution of (7.53) gives rise to a probability measure on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ with density ρ. This means that a solution of (7.53) is weakly unique. By definition, the process $R_I(t)$ exists on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$. This process is, in fact, a weak solution of (7.50). Hence the solution is weakly unique. $\square$

We point out that $w(t)$, defined as in the proof of Theorem 7.95, is a Wiener process relative to the family $\mathcal{P}_t^z$ generated by the weak solution $z(t)$.

Theorem 7.96 *Assume that M is stochastically complete, $a(t, m)$ is Borel measurable jointly in $(t, m) \in [0, l] \times M$, the inequality $\int_0^l \|a\,(t, m(t))\|^2\ \mathrm{d}t < \infty$ holds for any continuous curve $m(\cdot)\colon [0, l] \to M$, $m(0) = m_0 \in M$, and the density ρ defined by (7.51) satisfies (7.52). Then there exist a weakly unique weak solution ξ of (7.50) with the initial condition $\xi(0) = m_0$ that is well-defined on $[0, l]$.*

This result is a simple generalization of Theorem 7.95. The only refinement needed in the proof is as follows. Even though the hypothesis of Theorem 7.96 does not guarantee that $\Gamma a\big(t, R_I x(t)\big)$ is uniformly bounded, ν-almost all $x(\cdot) \in \tilde{\Omega}$ with $x(0) = 0 \in T_{m_0}M$ satisfy the inequality $\int_0^l \|\Gamma a\,(t, R_I x(t))\|^2\ \mathrm{d}t < \infty$. Arguing in the same way as in the proof of Theorem 7.95, we see that this inequality together with (7.52) yield the existence and weak uniqueness of a solution of (7.53) as in [83, 162]. The rest of the proof of Theorem 7.95 remains unchanged.

Corollary 7.97 *Assume that $\int_0^l \|a\,(t, m(t))\|^2\ \mathrm{d}t < \infty$ for any continuous curve $m(\cdot)\colon [0, l] \to M$. Then, under the hypotheses of Theorem 7.96, the assertion of the theorem holds for any initial condition $\xi(0) = m \in M$.*

From the proofs of Theorems 7.95 and 7.96 it is not hard to see that for processes with unit diffusion coefficients their developments (as well as line integrals with parallel translation) are well-defined on stochastically complete Riemannian manifolds, i.e., on a broader class than the uniformly complete Riemannian manifolds and manifolds satisfying the hypothesis of Theorem 7.76. In addition, for such processes it is possible to weaken the requirement of boundedness and yet still obtain the existence of an Itô development on every *a priori* given non-random time interval.

Consider a stochastic process $\beta(t)$ on $T_{m_0}M$ non-anticipative with respect to $\mathcal{B}_t$ and such that

$$P\left(\int_0^\infty \|\beta(\tau)\|^2 \, \mathrm{d}\tau < \infty\right) = 1. \tag{7.54}$$

Define an Itô process $z(t)$ on $T_{m_0}M$ by the formula

$$z(t) = \int_0^t \beta(\tau) \, \mathrm{d}\tau + w(t). \tag{7.55}$$

Theorem 7.98 *Let M be stochastically complete. Then for any $l > 0$ the development $R_I z(t)$ exists on $[0, l]$ and is weakly unique.*

Proof. Let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \nu)$ and $\mathcal{B}_t$ be as in the proof of Theorem 7.95. Denote by μ_z the probability measure on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ which corresponds to z. Then it follows from (7.54) that μ_z is absolutely continuous with respect to ν (see [175, Chapter 7]). Since M is stochastically complete, the development of a standard Wiener process exists. In other words, the development can be a.s. extended from the space of smooth curves to $\tilde{\Omega}$. Since $\mu_z \ll \nu$, the same holds true when μ_z is replaced by ν.

One can easily see that $z(t)$ coincides with the coordinate process $z\big(t, x(\cdot)\big) = x(t)$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu_z)$. The theorem follows. □

In the Euclidean space $\mathbb{R}^n$ consider a process $\beta(t)$ that satisfies (7.54) and is non-anticipative with respect to $\mathcal{B}_t$ and a Wiener process $w(t)$ adapted to $\mathcal{B}_t$. Construct the Itô process $\vartheta(t) = \int_0^t \beta(\tau)\mathrm{d}\tau + w(t)$. Let $x_0 : \Omega \to M$ be a random element independent of $\mathcal{B}_t$, and let β_0 be a Borel measurable cross-section of OM.

Consider the development $R_I(\beta_0(x_0))\vartheta(t)$ from Definition 7.68.

Theorem 7.99 *If a Riemannian manifold M is stochastically complete, the development $R_I(\beta_0(x_0))\vartheta(t)$ exists, is well-defined for $t \in [0, l]$ and is weakly unique.*

The proof of Theorem 7.99 is analogous to that of Theorem 7.98.

Chapter 8
Mean Derivatives in Linear Spaces

8.1 General Definitions and Results

In this section we briefly describe some preliminary facts about mean derivatives. For details, see [7, 106, 107, 115, 188, 190]. This notion was first introduced by E. Nelson [187, 188, 190] for the needs of so-called stochastic mechanics (see Chapter 15) but it turns out to be useful in some other problems of mathematical physics, economics, and elsewhere.

Consider a stochastic process $\xi(t)$, $t \in [0,T]$, given on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, taking values in a separable Hilbert space and such that $\xi(t)$ is an L^1 random element for all t. For the sake of convenience in this section we work mainly with the Euclidean space $\mathbb{R}^n$. The general case of Hilbert space is quite analogous and we shall describe its features only when necessary.

In Section 6.1.1 for a stochastic process $\xi(t)$ three families of σ-subalgebras of the σ-algebra $\mathcal{F}$ were introduced: "the past" $\mathcal{P}_t^\xi$ of $\xi(t)$, "the future" $\mathcal{F}_t^\xi$ of $\xi(t)$ and "the present" ("now") $\mathcal{N}_t^\xi$ of $\xi(t)$. All the above families we suppose to be complete, i.e., contain all sets of measure zero.

For the sake of convenience we denote by E_t^ξ the conditional expectation $E(\cdot|\mathcal{N}_t^\xi)$ with respect to the "present" $\mathcal{N}_t^\xi$ for $\xi(t)$.

Generally speaking, the sample trajectories of $\xi(\cdot)$ a.s. are not differentiable (see, e.g., Theorem 6.10 for Wiener processes) and so we cannot determine the derivative of $\xi(\cdot)$ in the ordinary way. Following Nelson (see e.g. [187, 188, 190]) we give the following:

Definition 8.1.

(i) The *forward mean derivative* $D\xi(t)$ of the process $\xi(t)$ at time t is the L^1-random variable of the form

$$D\xi(t) = \lim_{\Delta t \to +0} E_t^\xi \left(\frac{\xi(t+\Delta t) - \xi(t)}{\Delta t} \right) \tag{8.1}$$

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_8,

where the limit is assumed to exist in $L^1(\Omega, \mathcal{F}, \mathsf{P})$ and $\Delta t \to +0$ means that $\Delta t \to 0$ and $\Delta t > 0$.

(ii) The *backward mean derivative* $D_*\xi(t)$ of $\xi(t)$ at t is the L^1-random variable

$$D_*\xi(t) = \lim_{\Delta t \to +0} E_t^\xi \left(\frac{\xi(t) - \xi(t - \Delta t)}{\Delta t} \right) \tag{8.2}$$

where (as in (i)) the limit is assumed to exist in $L^1(\Omega, \mathcal{F}, \mathsf{P})$ and $\Delta t \to +0$ means that $\Delta t \to 0$ and $\Delta t > 0$.

If $\xi(t)$ is a Markov process (see Section 6.1.3) then E_t^ξ can be replaced by $E(\cdot|\mathcal{P}_t^\xi)$ in (8.1) and by $E(\cdot|\mathcal{F}_t^\xi)$ in (8.2). In order to distinguish these constructions for non-Markovian processes, we introduce the following definition:

Definition 8.2. The *forward mean derivative relative to the past* ($\mathcal{P}$*-mean derivative*) $D^{\mathcal{P}}\xi(t)$ of a process $\xi(t)$ at time t is the L^1-random element of the form

$$D^{\mathcal{P}}\xi(t) = \lim_{\triangle t \to +0} E\left(\frac{\xi(t + \triangle t) - \xi(t)}{\triangle t} \,\middle|\, \mathcal{P}_t^\xi \right), \tag{8.3}$$

where the limit is assumed to exist in L^1 and $\triangle t \to +0$ means that $\triangle t$ tends to 0 and $\triangle t > 0$.

The *backward mean derivative relative to the future* ($\mathcal{F}$*-mean derivative*) $D_*^{\mathcal{F}}\xi(t)$ of a process $\xi(t)$ at time t is the L^1-random element of the form

$$D_*^{\mathcal{F}}\xi(t) = \lim_{\triangle t \to +0} E\left(\frac{\xi(t) - \xi(t - \triangle t)}{\triangle t} \,\middle|\, \mathcal{F}_t^\xi \right), \tag{8.4}$$

where the limit is assumed to exist in L^1 and $\triangle t \to +0$ means that $\triangle t$ tends to 0 and $\triangle t > 0$.

Remark 8.3. In fact Nelson considered at most the case of Markov processes, and so he gave in different works two equivalent definitions of mean derivatives: with E_t^ξ and with $E(\cdot|\mathcal{P}_t^\xi)$ or $E(\cdot|\mathcal{F}_t^\xi)$, respectively. We mainly consider Itô diffusion type processes which are, generally speaking, non-Markovian, and so these definitions become non-equivalent. Definition 8.1 is compatible with the principle of locality in physics: the derivative depends on the present but neither on the entire past nor on the entire future. Nevertheless the $\mathcal{P}$-mean and $\mathcal{F}$-mean derivatives as in Definition 8.2 also arise in many problems.

It should be noted that in general $D\xi(t) \neq D_*\xi(t)$ (but if $\xi(t)$ a.s. has smooth sample trajectories, these derivatives coincide).

From the properties of conditional expectation it follows that $D\xi(t)$ and $D_*\xi(t)$ are expressed as compositions of $\xi(t)$ and the Borel measurable vector fields, namely the regressions (see Section 6.1.2)

$$Y^0(t,x) = \lim_{\Delta t \to +0} E\left(\frac{\xi(t+\Delta t) - \xi(t)}{\Delta t} \middle| \xi(t) = x \right) \tag{8.5}$$

$$Y^0_*(t,x) = \lim_{\Delta t \to +0} E\left(\frac{\xi(t) - \xi(t-\Delta t)}{\Delta t} \middle| \xi(t) = x \right) \tag{8.6}$$

on $\mathbb{R}^n$, i.e., $D\xi(t) = Y^0(t, \xi(t))$ and $D_*\xi(t) = Y^0_*(t, \xi(t))$.

The mean derivatives of Definition 8.1 are particular cases of the following notions. Let $x(t)$ and $y(t)$ be L^1-stochastic processes given on $(\Omega, \mathcal{F}, \mathsf{P})$. Introduce the y-forward derivative of $x(t)$ by the formula

$$D^y x(t) = \lim_{\Delta t \to +0} E^y_t \left(\frac{x(t+\Delta t) - x(t)}{\Delta t} \right) \tag{8.7}$$

and the y-backward derivative of $x(t)$ by the formula

$$D^y_* x(t) = \lim_{\Delta t \to +0} E^y_t \left(\frac{x(t) - x(t-\Delta t)}{\Delta t} \right) \tag{8.8}$$

where, of course, the limits are assumed to exist in $L^1(\Omega, \mathcal{F}, \mathsf{P})$. If by analogy with (8.3) and (8.4) we replace E^y_t by $E(\cdot \mid \mathcal{P}^y_t)$ in (8.7) and by $E(\cdot \mid \mathcal{F}^y_t)$ in (8.8), we obtain the *y-forward $\mathcal{P}$-derivative* $D^{\mathcal{P}^y} x(t)$ of $x(t)$ and the *y-backward $\mathcal{F}$-derivative* $D^{\mathcal{F}^y}_* x(t)$ of $x(t)$, respectively. As above, if $y(t)$ is Markovian, $D^{\mathcal{P}^y} x(t)$ coincides with $D^y x(t)$ and $D^{\mathcal{F}^y}_* x(t)$ coincides with $D^y_* x(t)$.

Lemma 8.4 *For $s < t$*

$$E\left(x(t) - x(s) \mid \mathcal{P}^y_s\right) = E\left(\int_s^t \left(D^{\mathcal{P}^y} x(\tau) \right) \mathrm{d}\tau \middle| \mathcal{P}^y_s \right), \tag{8.9}$$

$$E\left(x(t) - x(s) \mid \mathcal{F}^y_t\right) = E\left(\int_s^t \left(D^{\mathcal{F}^y}_* x(\tau) \right) \mathrm{d}\tau \middle| \mathcal{F}^y_t \right). \tag{8.10}$$

Proof. Take a partition $q = (s = t_0 < t_1 < \cdots < t_N = t)$ of the interval $[s,t]$ and consider the following integral sum

$$\sum_{i=0}^{N-1} E\left(\frac{x(t_{i+1}) - x(t_i)}{t_{i+1} - t_i} \middle| \mathcal{P}^y_{t_i} (t_{i+1} - t_i) \right) = \sum_{i=0}^{N-1} E\left((x(t_{i+1}) - x(t_i)) \middle| \mathcal{P}^y_{t_i} \right)$$

whose limit as $\operatorname{diam} q \to 0$ is $\int_s^t (D^{\mathcal{P}^y} x(\tau)) \mathrm{d}\tau$. Since $t_i \geq s$, by the properties of conditional expectation $E(E(\cdot \mid \mathcal{P}^y_{t_i}) \mid \mathcal{P}^y_s) = E(\cdot \mid \mathcal{P}^y_s)$. Thus

$$\begin{aligned}
&E\left(\sum_{i=0}^{N-1} E\left(x\left(t_{i+1}\right)-x\left(t_i\right)\middle|\,\mathcal{P}_{t_i}^y\right)\middle|\,\mathcal{P}_s^y\right)\\
&=E\left(\sum_{i=0}^{N-1}\left(x\left(t_{i+1}\right)-x\left(t_i\right)\right)\middle|\,\mathcal{P}_s^y\right)\\
&=E\left(x(t)-x(s)\middle|\,\mathcal{P}_s^y\right).
\end{aligned}$$

This proves (8.9). The proof of (8.10) is similar, replacing the "past" by the "future".

Lemma 8.5

(i) *$x(t)$ is a martingale with respect to $\mathcal{P}_t^y$ if and only if $D^{\mathcal{P}^y}x(t)=0$.*

(ii) *$x(t)$ is a backward martingale with respect to $\mathcal{F}_t^y$ if and only if $D_*^{\mathcal{F}^y}\xi(t)=0$.*

Proof. Let $x(t)$ be a martingale with respect to $\mathcal{P}_t^y$. By the martingale property (see Section 6.1.4) $E(x(t+\Delta t)|\mathcal{P}_t^y)=x(t)$ and so $E(x(t+\Delta t)-x(t)\mid\mathcal{P}_t^y)=0$. Hence $D^{\mathcal{P}^y}x(t)=0$.

Let $D^{\mathcal{P}^y}x(t)=0$. Then by Lemma 8.4 for $t>s$ we have $E(x(t)-x(s)\mid\mathcal{P}_s^y)=0$. Thus $E(x(t)\mid\mathcal{P}_s^y)=E(x(s)\mid\mathcal{P}_s^y)$. But $E(x(s)\mid\mathcal{P}_t^y)=x(s)$ and so $x(t)$ is a martingale with respect to $\mathcal{P}_t^y$.

This proves (i). The proof of (ii) is similar. □

Corollary 8.6

(i) *If $x(t)$ is a martingale with respect to $\mathcal{P}_t^y$, $D^y x(t)=0$.*
(ii) *If $x(t)$ is a backward martingale with respect to $\mathcal{F}_t^y$, $D_*\xi(t)=0$.*

In particular, if a process $\xi(t)$ is a martingale, $D\xi(t)=0$ and if it is a backward martingale, $D_*\xi(t)=0$.

This follows from the fact that $D\xi(t)=E_t^{\xi}(D^{\mathcal{P}}\xi(t))$ and $D_*\xi(t)=E_t^{\xi}(D_*^{\mathcal{F}}\xi(t))$.

Of course, when we use the word "martingale" without indicating the filtration, we mean that it is with respect to it own "past" (and in the case of a backward martingale, with respect to its own "future").

Consider an Itô diffusion type process $\xi(t)$ (see Definition 6.16)

$$\xi(t)=\xi_0+\int_0^t a(s)\mathrm{d}s+\int_0^t A(s)\mathrm{d}w(s) \tag{8.11}$$

and a diffusion process (see Definition 6.17)

$$\eta(t)=\eta_0+\int_0^t a(s,\eta(s))\mathrm{d}s+\int_0^t A(s,\eta(s))\mathrm{d}w(s). \tag{8.12}$$

It should be noted that $\xi(t)$ can neither be a diffusion nor a Markov process.

Theorem 8.7 *For a process $\xi(t)$ of type (8.11) $D\xi(t)$ exists and is equal to $E^\xi_t(a(t))$. For a diffusion process $\eta(t)$ of type (8.12) $D\eta(t) = a(t, \eta(t))$.*

Proof. Evidently

$$\begin{aligned} &D\left(\xi_0 + \int_0^t a(s)\mathrm{d}s + \int_0^t A(s)\mathrm{d}w(s)\right) \\ &= D^\xi\left(\int_0^t a(s)\mathrm{d}s\right) + D^\xi\left(\int_0^t A(s)\mathrm{d}w(s)\right). \end{aligned}$$

Since $\int_0^t A(s)\mathrm{d}w(s)$ is a martingale with respect to $\mathcal{P}^\xi_t$ (see Theorem 6.11), from Corollary 8.6 it follows that $D^\xi(\int_0^t A(s)\mathrm{d}w(s)) = 0$. Then $D^\xi(\int_0^t a(s)\mathrm{d}s) = E^\xi_t(a(t))$ and $D\eta(t) = a(t,\eta(t))$ since $a(t,\eta(t))$ is measurable with respect to $\mathcal{N}^\eta_t$. □

Theorem 8.8 *For diffusion type process (8.11) the derivative $D^{\mathcal{P}}\xi(t)$ exists and takes the form $D^{\mathcal{P}}\xi(t) = a(t)$.*

Proof. Note that $\xi(t+\triangle t)-\xi(t) = \int_t^{t+\triangle t} a(s)\mathrm{d}s+\int_t^{t+\triangle t} A(s)\mathrm{d}w(s)$. Since the Itô integral is a martingale with respect to $\mathcal{P}^\xi_t$, $E(\int_t^{t+\triangle t} A(s)\mathrm{d}w(s) \mid \mathcal{P}^\xi_t) = 0$ and so we get

$$E(\xi(t+\triangle t) - \xi(t) \mid \mathcal{P}^\xi_t) = E\left(\int\limits_t^{t+\triangle t} a(t)\mathrm{d}t \,\middle|\mathcal{P}^\xi_t\right) = \int\limits_t^{t+\triangle t} E\left(a(t)\,\middle|\mathcal{P}^\xi_t\right)\mathrm{d}t.$$

Applying formula (8.3) we obtain that $D^{\mathcal{P}}\xi(t) = E(a(t) \mid \mathcal{P}^\xi_t)$. Since, by definition of a diffusion type process, $a(t)$ is measurable with respect to $\mathcal{P}^\xi_t$, $E(a(t) \mid \mathcal{P}^\xi_t) = a(t)$. □

Theorem 8.9 *For a backward Itô process*

$$\xi(t) = \xi_0 + \int_0^t a(s)\mathrm{d}s + \int_0^t A(s)\mathrm{d}_*w(s)$$

given on an interval $[0,T]$ and such that $A(t)$ is measurable with respect to $\mathcal{N}^\xi_t$ for all t, the derivative $D_*\xi(t)$ at $t \in (0,T]$ exists and equals $E^\xi_t(a(t)) + A(t)D^\xi_* w(t)$.

Proof. Indeed, as in Theorem 8.7, $D^\xi_*(\int_0^t a(s)\mathrm{d}s) = E^\xi_t(a(t))$. Approximate the backward increment of $\int_0^t A(s)\mathrm{d}_*w(s)$ by a summand of the backward integral sum (6.4) of the form $A(t)(w(t) - w(t-\triangle t))$. Since $A(t)$ is measurable with respect to $\mathcal{N}^\xi_t$, $\lim\limits_{\triangle t\to+0} E^\xi_t(A(t)\frac{(w(t)-w(t-\triangle t))}{\triangle t}) = A(t)D^\xi_* w(t)$. The Theorem follows. □

By Theorem 8.7 the forward mean derivative gives information about the drift of an Itô process. Following [6, 7] we introduce a new mean derivative D_2, called the quadratic mean derivative, that is responsible for the diffusion term of a process. This is a slight modification of an idea of Nelson from [188, 190].

Definition 8.10. For an L^1-stochastic process $\xi(t)$, $t \in [0, T]$, its *quadratic mean derivative* $D_2\xi(t)$ is defined by the formula

$$D_2\xi(t) = \lim_{\triangle t \to +0} E_t^\xi \left(\frac{(\xi(t+\triangle t) - \xi(t)) \otimes (\xi(t+\triangle t) - \xi(t))}{\triangle t} \right), \tag{8.13}$$

where $\otimes$ denotes the tensor product and the limit is assumed to exist in $L^1(\Omega, \mathcal{F}, \mathsf{P})$.

Note that here the tensor product of two vectors in $\mathbb{R}^n$ is the $n \times n$ matrix formed by the products of every component of the first vector with every component of the second one. Note also that for column vectors $X, Y \in \mathbb{R}^n$ their tensor product $X \otimes Y$ equals the matrix product XY^* of the column vector X and the row vector Y^* (the transpose of column Y).

As in the case of forward mean derivatives, if $\xi(t)$ is not Markovian, the quadratic mean derivative with respect to the past differs from that in Definition 8.10. To distinguish these constructions we introduce:

Definition 8.11. The *quadratic mean derivative relative to the past* (for short, *quadratic $\mathcal{P}$-mean derivative*) $D_2^{\mathcal{P}}\xi(t)$ of $\xi(t)$ at t is the L^1-random element of the form

$$D_2^{\mathcal{P}}\xi(t) = \lim_{\triangle t \to +0} E\left(\frac{(\xi(t+\triangle t) - \xi(t)) \otimes (\xi(t+\triangle t) - \xi(t))}{\triangle t} \,\Big|\, \mathcal{P}_t^\xi \right), \tag{8.14}$$

where the limit is assumed to exist in L^1, $\triangle t \to +0$ means that $\triangle t$ tends to 0 and $\triangle t > 0$ and $\otimes$ denotes the tensor product in $\mathbb{R}^n$.

Denote by $\mathrm{S}_+(n)$ the set of symmetric positive definite $n \times n$ matrices and by $\bar{\mathrm{S}}_+(n)$ the set of symmetric positive semi-definite matrices (the closure of $\mathrm{S}_+(n)$ in the space of all symmetric matrices $S(n)$).

We emphasize that the tensor product in (8.13) is a symmetric positive semi-definite matrix so that $D_2\xi(t)$ is a function with values in $\bar{\mathrm{S}}_+(n)$.

From the properties of conditional expectation (see, e.g., [194]) it follows that there exists a Borel mapping $\alpha(t, x)$ from $[0, T] \times \mathbb{R}^n$ to $\bar{\mathrm{S}}_+(n)$ such that $D_2\xi(t) = \alpha(t, \xi(t))$. As above, following [194], we call $\alpha(t, x)$ the *regression*.

Theorem 8.12 *Let $\xi(t)$ be a diffusion type process of the form* (8.11). *Then $D_2\xi(t) = E_t^\xi[\alpha(t)]$ where $\alpha(t) = A(t)A^*(t)$, $A^*(t)$ is the transposed matrix $A(t)$ and $A(t)A^*(t)$ is the matrix product. If $\xi(t)$ is a diffusion process, $D_2\xi(t) = \alpha(t, \xi(t))$ where α is the diffusion coefficient. In particular, $D_2\xi(t) = 0$ if and only if $\xi(t)$ a.s. has C^1-smooth sample paths.*

Proof. By direct calculation it follows that the components of $(\xi(t+\triangle t)-\xi(t))\otimes(\xi(t+\triangle t)-\xi(t))$ are elements of the matrix $(\xi(t+\triangle t)-\xi(t))(\xi(t+\triangle t)-\xi(t))^*$ where we use the matrix multiplication of the column-vector $(\xi(t+\triangle t)-\xi(t))$ and the row-vector $(\xi(t+\triangle t)-\xi(t))^*$ (i.e., the transpose of $(\xi(t+\triangle t)-\xi(t)))$. The product is a symmetric positive semi-definite matrix. Since $\xi(t+\triangle t)-\xi(t)=\int_t^{t+\triangle t}a(s)\mathrm{d}s+\int_t^{t+\triangle t}A(s)\mathrm{d}w(s)$, taking into account the properties of the Lebesgue and Itô integrals one can see that $(\xi(t+\triangle t)-\xi(t))(\xi(t+\triangle t)-\xi(t))^*$ is approximated by $a(t)a(t)^*(\Delta t)^2+(a(t)\Delta t)(A(t)\Delta w(t))^*+(A(t)\Delta w(t))(a(t)\Delta t)^*+A(t)A(t)^*\Delta t$. Thus we see that only $A(t)A(t)^*\Delta t$ is infinitesimal of the same order as Δt while the other summands are infinitesimals of order higher than Δt. Applying formula (8.13), we obtain the assertion of Theorem since $AA^*=\alpha$ (see above). If $\xi(t)$ a.s. has C^1-sample paths, $A=0$ a.s. and so $D_2\xi(t)=0$. On the other hand, if $D_2\xi(t)=0$, this means that in the expression for $(\xi(t+\triangle t)-\xi(t))\otimes(\xi(t+\triangle t)-\xi(t))$ there is no term with the same infinitesimal order as $\triangle t$. Hence, a.s. $A=0$. □

Theorem 8.13 *For the above-mentioned Itô diffusion type process $\xi(t)$ the derivative $D_2^{\mathcal{P}}\xi(t)$ exists and takes the form $D_2^{\mathcal{P}}\xi(t)=A(t)A^*(t)$ where $A^*(t)$ is the transpose of the matrix $A(t)$.*

The proof of Theorem 8.13 is a simple modification of that for Theorem 8.12 based on the fact that $A(t)$ is measurable with respect to $\mathcal{P}_t^{\xi}$.

Below we will often deal with the particular case of the process (8.11) for $F=\mathbb{R}^n$ with $A=\sigma I$, where $\sigma>0$ is a real constant and I is the identity operator; i.e., with diffusion type processes in $\mathbb{R}^n$ of the form

$$\xi(t)=\xi_0+\int_0^t a(s)\mathrm{d}s+\sigma w(t). \tag{8.15}$$

Note that Theorem 8.7 is valid for processes of the form (8.15).

For a process of type (8.15) we can obtain from the Itô formulae (6.10) and (6.13) and from formulae (8.9) and (8.10) two important relations.

Lemma 8.14 *Let $f(t,x)$ be a smooth mapping. For a process of type (8.15) for every $t>s$ the relations*

$$\begin{aligned}&E_s^{\xi}(f(t,\xi(t))-f(s,\xi(s)))\\&=E_s^{\xi}\left(\int_s^t\frac{\partial f}{\partial\tau}\mathrm{d}\tau+\int_s^t f'(Y^0(\tau,\xi(\tau)))\mathrm{d}\tau+\frac{\sigma^2}{2}\int_s^t\nabla^2 f(\tau,\xi(\tau))\mathrm{d}\tau\right)\end{aligned}\tag{8.16}$$

and

$$\begin{aligned}&E_t^{\xi}(f(t,\xi(t))-f(s,\xi(s)))\\&=E_t^{\xi}\left(\int_s^t\frac{\partial f}{\partial\tau}\mathrm{d}\tau+\int_s^t f'(Y_*^0(\tau,\xi(\tau)))\mathrm{d}\tau-\frac{\sigma^2}{2}\int_s^t\nabla^2 f(\tau,\xi(\tau))\mathrm{d}\tau\right)\end{aligned}\tag{8.17}$$

hold.

Proof. Note that by Theorem 8.7 and by formula (8.5) defining the regression Y^0, the equality $E_t^\xi(a(t)) = Y^0(t,\xi(t))$ holds. Also, $\mathrm{tr} f''(\cdot,\cdot) = \nabla^2 f$ for a smooth mapping $f(t,x)$ where ∇^2 is the Laplacian. Then for $t > s$ and for a smooth mapping $f(t,x)$ one can easily derive from (6.10) and (8.9) that

$$
\begin{aligned}
&E_s^\xi(f(t,\xi(t)) - f(s,\xi(s)))\\
&= E_s^\xi\left(\int_s^t D^\xi f(\tau,\xi(\tau))\mathrm{d}\tau\right)\\
&= E_s^\xi\left(\int_s^t \frac{\partial f}{\partial \tau}\mathrm{d}\tau + \int_s^t f'(Y^0(\tau,\xi(\tau)))\mathrm{d}\tau + \frac{\sigma^2}{2}\int_s^t \nabla^2 f(\tau,\xi(\tau))\mathrm{d}\tau\right)
\end{aligned}
$$

since $E_s^\xi(\int_s^t f'(\mathrm{d}w(\tau)) = 0$.

For the process (8.15) the regression Y_*^0 introduced by formula (8.6) takes the form $Y_*^0 = Y^0 + Y^1$ where Y^1 is the regression for $D_*^\xi w(t)$, i.e., $Y^1(t,\xi(t)) = D_*^\xi w(t)$. Obviously, the backward Itô formula (6.13) is applicable to $\xi(t)$ as well as (6.10). Then for $t > s$ and for a smooth mapping $f(t,x)$ one can easily derive from (6.13), (8.10) and Theorem 8.9 that

$$
\begin{aligned}
&E_t^\xi(f(t,\xi(t)) - f(s,\xi(s)))\\
&= E_t^\xi\left(\int_s^t D^\xi f(\tau,\xi(\tau))\mathrm{d}\tau\right)\\
&= E_t^\xi\left(\int_s^t \frac{\partial f}{\partial \tau}\mathrm{d}\tau + \int_s^t f'(Y^0(\tau,\xi(\tau)))\mathrm{d}\tau - \frac{\sigma^2}{2}\int_s^t \nabla^2 f(\tau,\xi(\tau))\mathrm{d}\tau\right.\\
&\qquad \left. + \int_s^t f'(D_*^\xi w(\tau))\mathrm{d}\tau\right)\\
&= E_t^\xi\left(\int_s^t \frac{\partial f}{\partial \tau}\mathrm{d}\tau + \int_s^t f'(Y^0(\tau,\xi(\tau)))\mathrm{d}\tau - \frac{\sigma^2}{2}\int_s^t \nabla^2 f(\tau,\xi(\tau))\,\mathrm{d}\tau\right.\\
&\qquad \left. + \int_s^t f'(Y^1(\tau,\xi(\tau)))\mathrm{d}\tau\right)\\
&= E_t^\xi\left(\int_s^t \frac{\partial f}{\partial \tau}\mathrm{d}\tau + \int_s^t f'(Y_*^0(\tau,\xi(\tau)))\mathrm{d}\tau - \frac{\sigma^2}{2}\int_s^t \nabla^2 f(\tau,\xi(\tau))\mathrm{d}\tau\right).
\end{aligned}
$$

□

Definition 8.15. The derivative $D_S = \frac{1}{2}(D + D_*)$ is called the *symmetric mean derivative*. The derivative $D_A = \frac{1}{2}(D - D_*)$ is called the *antisymmetric mean derivative*.

Consider the vectors $v^\xi(t,x) = \frac{1}{2}(Y^0(t,x) + Y_*^0(t,x))$ and $u^\xi(t,x) = \frac{1}{2}(Y^0(t,x) - Y_*^0(t,x))$.

Definition 8.16. $v^\xi(t) = v^\xi(t, \xi(t)) = D_S\xi(t)$ is called the *current velocity* of the process $\xi(t)$ and $u^\xi(t) = u^\xi(t, \xi(t)) = D_A\xi(t)$ is called the *osmotic velocity* of the process $\xi(t)$.

In Nelson's works it is shown that in many natural problems the current velocity plays the same role as the ordinary velocity of a deterministic process. The osmotic velocity measures the variation of randomness of a process. The precise meaning of osmotic and current velocities is clarified by the following.

Denote by $\rho^\xi(t,x)$ the density of the process (8.15) with respect to Lebesgue measure λ on $[0,l] \times \mathbb{R}^n$. This means that for any continuous integrable function $f(t,x)$ on $[0,l] \times \mathbb{R}^n$ the following equality holds

$$\int_{[0,l]\times\mathbb{R}^n} f(t,x)\rho^\xi(t,x)\mathrm{d}\lambda = \int_{\Omega\times[0,l]} f(t,\xi(t))\mathrm{d}\mathsf{P}\mathrm{d}t.$$

Lemma 8.17 *For the process (8.15) in $\mathbb{R}^n$ the vector field $u^\xi(t,x)$ is presented in the form*

$$u^\xi(t,x) = \frac{1}{2}\sigma^2 \mathrm{grad}\ \log\ \rho^\xi(t,x) = \sigma^2 \mathrm{grad}\ \log\ \sqrt{\rho^\xi(t,x)}. \tag{8.18}$$

Proof. We shall prove (8.18) using an idea developed in [65] for more complicated processes. For an alternative proof see, e.g., [187, 188, 190] where only Markovian diffusion processes are considered.

Let f be a smooth function on $\mathbb{R}^n$ with compact support. Since $f(\xi(t))$ is $\mathcal{N}_t^\xi$-measurable,

$$E[f(\xi(t))E_t^\xi(w(t) - w(t-\Delta t))] = E[f(\xi(t))(w(t) - w(t-\Delta t))]$$

(see the properties of conditional expectations in Section 6.1.2). Since $f(\xi(t-\Delta t))$ and $w(t) - w(t-\Delta t)$ are independent and $E(w(t) - w(t-\Delta t)) = 0$ (see Theorem 6.8), we obtain the equality $E[f(\xi(t-\Delta t))(w(t) - w(t-\Delta t))] = 0$. Thus $E[f(\xi(t))E_t^\xi(w(t) - w(t-\Delta t))] = E[(f(\xi(t)) - f(\xi(t-\Delta t)))(w(t) - w(t-\Delta t))]$. Using the Itô formula (6.10) and taking into account that $f'(a) = \mathrm{grad} f \cdot a$, we obtain the presentation

$$\begin{aligned} &f(\xi(t)) - f(\xi(t-\Delta t)) \\ &= \int_{t-\Delta t}^{t} (\mathrm{grad} f(\xi(s)) \cdot a(s))\mathrm{d}s + \frac{\sigma^2}{2}\int_{t-\Delta t}^{t} \mathrm{tr} f''(\xi(s))\mathrm{d}s \\ &\qquad + \int_{t-\Delta t}^{t} \sigma \mathrm{grad} f(\xi(s)) \cdot \mathrm{d}w(s). \end{aligned}$$

Hence,

$$
\begin{aligned}
&E\left[f(\xi(t))E_t^\xi(w(t)-w(t-\Delta t))\right]\\
&= E\left[\int_{t-\Delta t}^{t}(\mathrm{grad} f(\xi(s))\cdot a(s))\mathrm{d}s\,\mathrm{d}w(s)+\frac{\sigma^2}{2}\int_{t-\Delta t}^{t}\mathrm{tr} f''(\xi(s))\mathrm{d}s\,\mathrm{d}w\right.\\
&\qquad \left.+\int_{t-\Delta t}^{t}\sigma\mathrm{grad} f(\xi(s))\mathrm{d}s\right] = E\left[\int_{t-\Delta t}^{t}\sigma\mathrm{grad} f(\xi(s))\mathrm{d}s\right],
\end{aligned}
$$

and so

$$
\begin{aligned}
E[f((\xi(t))u^\xi(t,\xi(t))] &= -\frac{1}{2}E\left[\sigma f(\xi(t))\lim_{\Delta t\to+0}\frac{1}{\Delta t}E_t^\xi(w(t)-w(t-\Delta t)\right]\\
&= -\frac{1}{2}E[\sigma^2\mathrm{grad}\, f(\xi(t))]\\
&= -\frac{1}{2}\int_{[0,l]\times\mathbb{R}^n}\sigma^2\mathrm{grad}\, f(x)\rho^\xi(t,x)\mathrm{d}\lambda\\
&= \frac{1}{2}\int_{[0,l]\times\mathbb{R}^n}\sigma^2 f(x)\mathrm{grad}\,\rho^\xi(t,x)\mathrm{d}\lambda\\
&= \frac{1}{2}\int_{[0,l]\times\mathbb{R}^n}\sigma^2 f(x)\left(\frac{\mathrm{grad}\,\rho^\xi(t,x)}{\rho^\xi(t,x)}\right)\rho^\xi(t,x)\mathrm{d}\lambda\\
&= \frac{1}{2}E[\sigma^2 f(\xi(t))\mathrm{grad}\ \log\ \rho^\xi(t,\xi(t))].
\end{aligned}
$$

Since this is valid for every smooth function f with compact support, (8.18) holds. □

From (8.18) it follows that the osmotic velocity does indeed characterize the 'variation of randomness' of the process.

Lemma 8.18 *For the process* (8.15) *in* $\mathbb{R}^n$*, the vector field* $v^\xi(t,x)$ *and the density* $\rho^\xi(t,x)$ *satisfy the continuity equation*

$$
\frac{\partial\rho^\xi(t,x)}{\partial t} = -\mathrm{div}(\rho^\xi v^\xi). \tag{8.19}
$$

Proof. Let $f(t,x)$ be a smooth real-valued function on $[0,l]\times\mathbb{R}^n$ with compact support. Note that for such f and a vector Y the equality $f'(Y)=(\mathrm{grad} f\cdot Y)$ holds where the dot denotes the inner product in $\mathbb{R}^n$. Recall that $E(E_t^\xi(\cdot)) = E(\cdot)$. Then by formula (8.16) for $t>s$ the equality

$$
\begin{aligned}
&E[f(t,\xi(t))-f(s,\xi(s))]\\
&= E\left[\int_s^t\frac{\partial f}{\partial t}\mathrm{d}\tau\int_s^t(\mathrm{grad} f\cdot Y^0(t,\xi(t)))\mathrm{d}\tau+\int_s^t\frac{\sigma^2}{2}\nabla^2 f\mathrm{d}\tau\right]
\end{aligned}
$$

holds. On the other hand, by formula (8.17) we get the equality

$$
\begin{aligned}
&E[f(t,\xi(t)) - f(s,\xi(s))] \\
&= E\left[\int_s^t \frac{\partial f}{\partial t}\mathrm{d}\tau + \int_s^t (\mathrm{grad} f \cdot Y_*^0(t,\xi(t)))\mathrm{d}\tau - \int_s^t \frac{\sigma^2}{2}\nabla^2 f \mathrm{d}\tau\right].
\end{aligned}
$$

Thus

$$
E[f(t,\xi(t)) - f(s,\xi(s))] = E\left[\int_s^t \frac{\partial f}{\partial t}\mathrm{d}\tau + \int_s^t (\mathrm{grad} f \cdot v^{\xi}(t,\xi(t)))\mathrm{d}\tau\right],
$$

where $v^{\xi}(t,x) = \frac{1}{2}(Y^0(t,x) + Y_*^0(t,x))$ (see above). But

$$
\begin{aligned}
&E\left[\int_s^t \frac{\partial f}{\partial t}\mathrm{d}\tau + \int_s^t (\mathrm{grad} f \cdot v^{\xi}(t,\xi(t)))\mathrm{d}\tau\right] \\
&= \int_{[s,t]\times\mathbb{R}^n} \left[\frac{\partial f(\tau,x)}{\partial \tau}\rho(\tau,x) + \big(\mathrm{grad} f \cdot v^{\xi}(\tau,x)\rho(\tau,x)\big)\right]\mathrm{d}\lambda \\
&= \int_{[s,t]\times\mathbb{R}^n} \frac{\partial}{\partial t}\left[f(\tau,x)\rho^{\xi}(\tau,x)\right]\mathrm{d}\lambda - \int_{[s,t]\times\mathbb{R}^n}\left[f(\tau,x)\frac{\partial \rho^{\xi}(\tau,x)}{\partial \tau}\right]\mathrm{d}\lambda \\
&\qquad - \int_{[s,t]\times\mathbb{R}^n}\left[f(\tau,x)\mathrm{div}\left(v^{\xi}(\tau,x)\rho^{\xi}(\tau,x)\right)\right]\mathrm{d}\lambda \\
&= E[f(t,\xi(t)) - f(s,\xi(s))] - \int_{[s,t]\times\mathbb{R}^n} f(\tau,x)\frac{\partial \rho^{\xi}(\tau,x)}{\partial \tau}\mathrm{d}\lambda \\
&\qquad - \int_{[s,t]\times\mathbb{R}^n}\left[f(\tau,x)\mathrm{div}\left(v^{\xi}(\tau,x)\rho^{\xi}(\tau,x)\right)\right]\mathrm{d}\lambda.
\end{aligned}
$$

Hence

$$
\int_{[s,t]\times\mathbb{R}^n}\left[f(\tau,x)\frac{\partial \rho^{\xi}(\tau,x)}{\partial \tau}\right]\mathrm{d}\lambda = -\int_{[s,t]\times\mathbb{R}^n}\left[f(\tau,x)\mathrm{div}\left(v^{\xi}(\tau,x)\rho^{\xi}(\tau,x)\right)\right]\mathrm{d}\lambda.
$$

Since the last equality is valid for any f, the Lemma follows. An alternative proof for a Markovian diffusion $\xi(t)$ can be found, e.g., in [187, 188, 190]. □

Lemmas 8.17 and 8.18 can be generalized for processes with more complicated diffusion terms in the following way.

Consider an autonomous smooth field of non-degenerate linear operators $A(x) : \mathbb{R}^n \to \mathbb{R}^n$, $x \in \mathbb{R}^n$. Suppose that $\xi(t)$ is a diffusion type process whose diffusion integrand is $A(\xi(t))$. Then its diffusion coefficient $A(x)A^*(x)$ is a smooth field of symmetric positive definite matrices $\alpha(x) = (\alpha^{ij}(x))$. Since all such matrices are invertible, the field of inverse matrices (α_{ij}) exists and is smooth and at any x the matrix $(\alpha_{ij})(x)$ is symmetric and positive definite. Thus it defines a new Riemannian metric $\alpha(\cdot,\cdot) = \alpha_{ij}\mathrm{d}x^i \otimes \mathrm{d}x^j$ on $\mathbb{R}^n$. Consider the Riemannian volume form of this Riemannian metric $\Lambda_{\alpha} = \sqrt{\det(\alpha_{ij})}\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n$ (see (1.31)).

Denote by $\rho^\xi(t,x)$ the probability density of $\xi(t)$ with respect to the volume form $\mathrm{d}t \wedge \Lambda_\alpha = \sqrt{\det(\alpha_{ij})}\mathrm{d}t \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n$ on $[0,T] \times \mathbb{R}^n$, i.e., for any continuous bounded function $f : [0,T] \times \mathbb{R}^n \to R$ the relation

$$\int_0^T E(f(t,\xi(t)))\mathrm{d}t = \int_0^T \left(\int_\Omega f(t,\xi(t))\mathrm{dP} \right) \mathrm{d}t = \int_{[0,T]\times\mathbb{R}^n} f(t,x)\rho^\xi(t,x)\mathrm{d}t \wedge \Lambda_\alpha$$

holds. We have the following generalization of formula (8.18):

$$u^\xi(t,x) = \frac{1}{2}\mathrm{Grad}\log\rho^\xi(t,x) = \mathrm{Grad}\log\sqrt{\rho^\xi(t,x)}, \tag{8.20}$$

where Grad denotes the gradient with respect to the Riemannian metric $\alpha(\cdot,\cdot)$. Analogously for $v^\xi(t,x)$ and $\rho^\xi(t,x)$ we have the following generalization of formula (8.19) (equation of continuity)

$$\frac{\partial \rho^\xi(t,x)}{\partial t} = -\mathrm{Div}(v^\xi(t,x)\rho^\xi(t,x)) \tag{8.21}$$

where Div denotes divergence with respect to the Riemannian metric $\alpha(\cdot,\cdot)$. The arguments for the derivation of (8.20) and (8.21) are analogous to those in the proofs of Lemmas 8.17 and 8.18 with a natural modification for using Grad and Div. For an alternative proof for Markovian diffusion processes, see [190].

Let $Z(t,x)$ be a C^2-smooth vector field, and $\xi(t)$ be a stochastic process.

Definition 8.19. The *forward* and *backward mean derivatives* of Z along $\xi(\cdot)$ at t (denoted by $DZ(t,\xi(t))$ and $D_*Z(t,\xi(t))$, respectively) are the L^1-limits of the form

$$DZ(t,\xi(t)) = \lim_{\Delta t\to+0} E_t^\xi \left(\frac{Z(t+\Delta t,\xi(t+\Delta t)) - Z(t,\xi(t))}{\Delta t} \right) \tag{8.22}$$

$$D_*Z(t,\xi(t)) = \lim_{\Delta t\to+0} E_t^\xi \left(\frac{Z(t,\xi(t)) - Z(t-\Delta t,\xi(t-\Delta t))}{\Delta t} \right) \tag{8.23}$$

As in Definition 8.1, if $\xi(t)$ is a Markov process, E_t^ξ can be replaced by $E(\cdot|\mathcal{P}_t^\xi)$ in (8.22) and by $E(\cdot|\mathcal{F}_t^\xi)$ in (8.23), see Remark 8.3.

Of course $DZ(t,\xi(t))$ and $D_*Z(t,\xi(t))$ can be presented as compositions of $\xi(t)$ with certain Borel vector fields on $\mathbb{R}^n$. These vector fields (regressions) will also be denoted by DZ and D_*Z, respectively.

Lemma 8.20 *For the process* (8.15) *in* $\mathbb{R}^n$ *the following formulae hold*

$$DZ = \frac{\partial}{\partial t}Z + (Y^0 \cdot \nabla)Z + \frac{\sigma^2}{2}\nabla^2 Z, \tag{8.24}$$

$$D_*Z = \frac{\partial}{\partial t}Z + (Y_*^0 \cdot \nabla)Z - \frac{\sigma^2}{2}\nabla^2 Z, \tag{8.25}$$

where $\nabla = (\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^n})$, ∇^2 *is the Laplacian, the dot denotes the inner product in* $\mathbb{R}^n$ *and* $Y^0(t,x)$ *and* $Y^0_*(t,x)$ *are as introduced in formulae* (8.5) *and* (8.6).

Proof. The vector field $Z(t,x)$ can be considered as a map $Z : [0,l]\times\mathbb{R}^n \to \mathbb{R}^n$ and so one may apply formulae (8.16) and (8.17). Note that $Z'(Y) = (Y\cdot\nabla)Z$ for a vector Y. Formula (8.24) follows immediately from (8.16) and (8.22) while (8.25) follows from (8.17) and (8.23). □

In fact the forward mean derivative and the quadratic mean derivative together make it possible in principle to recover a stochastic process from its mean derivatives: the forward mean derivative gives information about the drift while the quadratic mean derivative gives information about the diffusion term. It turns out that such recovery is also possible for more complicated relations with mean derivatives that we call *equations with mean derivatives (EMDs)*.

Specify a homogeneous polynomial $f(x,y)$ of order k of two variables, analogous polynomials $g_i(x,y)$, $i = 1,...,k-1$ of order i and two mappings: $F : \mathbb{R}\times\mathbb{R}^n\times ...\times\mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$.

Definition 8.21. A *k-th-order stochastic equation with mean derivatives* (*EMD*) is a system

$$\begin{aligned} f(D,D_*)\xi(t) &= F(t,\xi(t), g_1(D,D_*)\xi(t), ..., g_{k-1}(D,D_*)\xi(t)), \qquad (8.26)\\ D_2\xi(t) &= g(t,\xi(t)). \end{aligned}$$

The definition of a differential inclusion with mean derivatives is analogous. In Section 8.4 below we prove some existence theorems for several types of first order equations and inclusion with mean derivatives. Various second order equations and inclusions with mean derivatives are considered in Chapters 14 and 15 and in Section 16.4. For the construction of their solutions we need to have formulae of mean derivatives for processes from a sufficiently broad class. The next two sections are devoted to the derivation of such formulae.

8.2 Calculation of Mean Derivatives for a Wiener Process and for Diffusion Processes

For a Wiener process $w(t)$ in $\mathbb{R}^n$ $Dw(t) = 0$, $t\in[0,l)$, by Lemma 8.5(i) since $w(t)$ is a martingale.

Lemma 8.22 *For* $t\in(0,l]$ *the equality* $D_*w(t) = \frac{w(t)}{t}$ *holds.*

Proof. In this case, from the definition of the osmotic velocity $u^w(t, w(t))$ it follows that $D_* w(t) = -2u^w(t, w(t))$. Recall that the density $\rho^w(t,x)$ is given in formula (6.1). Thus according to formula (8.18) we have

$$u^w(t,x) = -\frac{1}{2} \cdot \frac{x}{t} \text{ i.e. } D_* w(t) = \frac{w(t)}{t}.$$

$\square$

Obviously the process $\frac{w(t)}{t}$ does not exist at $t = 0$. Nevertheless the following statement holds:

Lemma 8.23 *The integral $\int_0^t \frac{w(s)}{s} \mathrm{d}s$ exists a.s. for all $t \in [0, l]$.*

Proof. By standard calculations, using the density $\rho^w(t,x)$ one can easily obtain the estimate $E \int_0^t \| \frac{w(s)}{s} \| \mathrm{d}s < C \cdot \sqrt{t}$, where E denotes the expectation and the constant $C > 0$ depends only on the dimension n. Then the result follows from the classical Tchebyshev inequality. $\square$

Remark 8.24. To emphasize that mean derivatives essentially depend on the σ-algebras with respect to which they are calculated, we consider the following example. By Lemma 8.23 the process $\eta(t) = -\int_0^t \frac{w(s)}{s} \mathrm{d}s + w(t)$ is well-defined. From Lemma 8.22 one can easily derive that $D_*^w \eta(t) = 0$ and $D^w \eta(t) = -D_* w(t)$. But it is shown in [153] that $\eta(t)$ is a Wiener process with respect to its own 'past' family of σ-algebras and so $D\eta(t) = 0$.

Lemma 8.25

(i) $D^w \frac{w(t)}{t} = -\frac{w(t)}{t^2}$ *for* $t \in (0, l)$.
(ii) $D_*^w \frac{w(t)}{t} = 0$ *for* $t \in (0, l]$.

Proof. It is easy to see that $D^w \frac{w(t)}{t} = (\frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{t}) w(t) + \frac{1}{t} D w(t) = -\frac{w(t)}{t^2}$ and $D_*^w \frac{w(t)}{t} = (\frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{t}) w(t) + \frac{1}{t} D_* w(t) = 0$. $\square$

Lemma 8.26 *Let a Markovian diffusion process $\xi(t)$ be a solution of the Itô equation* (6.16)*. Then:*

(i) $D\xi(t) = a(t, \xi(t))$ *for* $t \in (0, l]$;
(ii) $D_* \xi(t) = a(t, \xi(t)) - \mathrm{tr} A'(A(t, \xi(t))) + A(t, \xi(t)) D_*^\xi w(t)$ *for* $t \in (0, l]$,

where $D_*^\xi w(t)$ is the backward mean derivative introduced in (8.8)*.*

Proof. Assertion (i) is a corollary of Theorem 8.7. To prove (ii), represent $\xi(t)$ by formula (6.26). Then using the fact that the first two summands on the right hand side of (6.26) are processes with a.s. smooth trajectories, as well as the properties of the conditional expectation and formula (8.8), we obtain

$$
\begin{aligned}
D_*\xi(t) &= D_*^{\xi}\left(\int_0^t a(s,\xi(s))\mathrm{d}s - \int_0^t \mathrm{tr}A'(A(s,\xi(s))\right)\mathrm{d}s \\
&\qquad + \int_0^t A(s,\xi(s))\mathrm{d}_* w(s)) \\
&= a(t,\xi(t)) - \mathrm{tr}A'(A(t,\xi(t))) \\
&\qquad + \lim_{\Delta t\to+0} E_t^{\xi}\left(A(t,\xi(t))\left(\frac{w(t)-w(t-\Delta t)}{\Delta t}\right)\right) \\
&= a(t,\xi(t)) - \mathrm{tr}A'(A(t,\xi(t))) \\
&\qquad + A(t,\xi(t))\lim_{\Delta t\to+0} E_t^{\xi}\left(\frac{w(t)-w(t-\Delta t)}{\Delta t}\right) \\
&= a(t,\xi(t)) - \mathrm{tr}A'(A(t,\xi(t)) + A(t,\xi(t))D_*^{\xi}w(t). \square
\end{aligned}
$$

Theorem 8.27 *Let $\xi(t)$ be a solution of the Itô equation*

$$
\begin{aligned}
\xi(t) = \xi_0 &+ \int_0^t a(s,\xi(s))\mathrm{d}s + \int_0^t \mathrm{tr}A'(A(s,\xi(s)))\mathrm{d}s \\
&- \int_0^t A(s,\xi(s))D_*^{\xi}w(s)\mathrm{d}s + \int_0^t A(s,\xi(s))\mathrm{d}w(s).
\end{aligned} \tag{8.27}
$$

Then for $t \in (0,l]$ we have $D_*\xi(t) = a(t,\xi(t))$.

In fact Theorem 8.27 is a corollary of Lemma 8.26 and the proof is absolutely analogous to the proof of the latter.

Lemma 8.28 *For a diffusion type process $\xi(t)$ the relation $ED_*^{\xi}w(t) = 0$ holds.*

Proof. By the definition of D_*^{ξ} and the properties of conditional expectation

$$
\begin{aligned}
ED_*^{\xi}w(t) &= E\lim_{\Delta t\to+0} E_t^{\xi}\frac{w(t)-w(t-\Delta t)}{\Delta t} \\
&= \lim_{\Delta t\to+0} E\frac{w(t)-w(t-\Delta t)}{\Delta t} \\
&= 0.
\end{aligned}
$$
$\square$

Let $\xi(t)$ be a Markovian diffusion process given on a finite interval $t \in [0,T]$.

Definition 8.29. The process $w_*^{\xi}(t) = \int_t^T D_*^{\xi}w(s)\mathrm{d}s + w(t) - w(T)$ is called the *backward Wiener process* with respect to $\xi(t)$.

Lemma 8.30 *The process $w_*^{\xi}(t)$ is a backward martingale with zero mean relative to $\mathcal{F}_t^{\xi}$.*

Proof. Since $\xi(t)$ is Markovian, $D^{\xi}_{*}w^{x}i_{*}(t) = D^{\mathcal{F}^{\xi}}_{*} w^{\xi}_{*}(t)$. Note that $D^{\xi}_{*}w^{\xi}_{*}(t) = -D^{\xi}_{*}w(t) + D^{\xi}_{*}w(t) = 0$. Hence, $D^{\mathcal{F}^{\xi}}_{*} w^{\xi}_{*}(t) = 0$ and the assertion of the Theorem follows from Lemma 8.5(ii) and Lemma 8.28. □

We should emphasize that $w^{\xi}_{*}(t)$ depends on the given process $\xi(t)$; from Lemmas 8.22 and 8.23 it follows that $w^{w}_{*}(t) = \int_t^T \frac{w(s)}{s}\mathrm{d}s + w(t) - w(T)$ and it is well-defined. Below in formula (8.37) we find $D^{\xi}_{*}w(s)$ for an Itô diffusion type process $\xi(t)$ with unit diffusion coefficient and so it is possible to find $w^{\xi}_{*}(t)$ in this case. The same can also be done for more general processes $\xi(t)$. As mentioned above (after Theorem 8.7), it is convenient to use $w^{\xi}_{*}(t)$ for representing $\xi(t)$ via this backward martingale and hence for calculating the backward derivatives (see below).

Let $\xi(t)$ be a diffusion process, i.e., a solution of the Itô equation

$$\xi(t) = \xi_0 + \int\limits_0^t a(\tau, \xi(\tau))\mathrm{d}\tau + \int\limits_0^t A(\tau, \xi(\tau))\mathrm{d}w(\tau) \tag{8.28}$$

for $t \in [0, T]$. Specify a time t. From the above formulae it follows that the process $\eta(t)$ satisfying for $s < t$ the relation

$$\eta(t) - \eta(s) = \int\limits_s^t a_{*}(\tau, \eta(\tau))\mathrm{d}\tau + \int\limits_s^t A(\tau, \eta(\tau))\mathrm{d}_{*}w^{\eta}_{*}(\tau) \tag{8.29}$$

with $a_{*}(t, x) = a(t, x) - \mathrm{tr}A'(A(t, x)) + A(t, x)D^{\eta}_{*}w(t)$ and such that $\eta(t) = \xi(t)$, has the same backward mean derivative at t as $\xi(t)$.

Lemma 8.31 $\int_s^t A(\tau, \eta(\tau))\mathrm{d}_{*}w^{\eta}_{*}(\tau)$ *is a backward martingale with respect to* $\mathcal{F}^{\eta}_{t}$.

Proof. By Lemma 8.30 and Lemma 8.5(ii), $D^{\mathcal{F}^{\eta}}_{*} w^{\eta}_{*}(t) = 0$. Approximate the backward increment of $\int_0^t A(\tau, \eta(\tau))\mathrm{d}_{*}w^{\eta}_{*}(s)$ by a summand of the backward integral sum (6.4), $A(t, \eta(t))(w^{\eta}_{*}(t) - w^{\eta}_{*}(t - \triangle t))$. Since $A(t, \eta(t))$ is measurable with respect to $\mathcal{N}^{\xi}_{t}$ and hence with respect to $\mathcal{F}^{\eta}_{t}$,

$$\begin{aligned}
&D^{\mathcal{F}^{\eta}}_{*} \int_s^t A(\tau, \eta(\tau))\mathrm{d}_{*}w^{\eta}_{*}(\tau)\\
&= \lim_{\triangle t \to +0} E\left(\left(A(t, \eta(t))\ \frac{(w^{\eta}_{*}(t) - w^{\eta}_{*}(t - \triangle t))}{\triangle t}\right) \middle| \mathcal{F}^{\eta}_{t}\right)\\
&= A(t, \eta(t))D^{\mathcal{F}^{\eta}}_{*} w^{\eta}_{*}(t)\\
&= 0.
\end{aligned}$$

Thus the assertion of the Theorem follows from Lemma 8.5(ii). □

Lemma 8.31 is ‘symmetric’ to Theorem 6.11(2).

Definition 8.32. The equality (8.29) is called an *equation in backward differentials* and is denoted as follows

$$\mathrm{d}_*\xi(t) = a(t, \xi(t))\mathrm{d}_*t + A(t, \xi(t))\mathrm{d}_*w_*^{\xi}(t). \tag{8.30}$$

Here d_*t means the increment in the negative time direction. We use this notation only for convenience. From the above arguments it follows that for $s < t$ close enough to t, a solution of (8.30) approximates a solution of (8.29).

Lemma 8.33 *Let $f(t, x)$ be a function that is C^1 in $t \in \mathbb{R}$ and C^2 in $x \in \mathbb{R}^n$. Then for the process $\xi(t)$ satisfying (8.30), the backward Itô formula*

$$\begin{aligned}\mathrm{d}_*f(\xi(t)) = \frac{\partial f}{\partial t}\mathrm{d}_*t &+ f'(a_*(t, \xi(t))\mathrm{d}_*t \\ &+ \frac{1}{2}\mathrm{tr} f''(A(t, \xi(t)), A(t, \xi(t)))\mathrm{d}_*t \\ &+ f'(A(t, \xi(t))\mathrm{d}_*w_*^{\xi}(t)\end{aligned}$$

is valid.

The proof of Lemma 8.33 is analogous to that of formula (10.11). Here we have to use the Taylor expansion at the right point and so the summands with even numbers change sign. Taking into account the properties of integrals of higher order, we obtain the assertion of the Lemma.

Remark 8.34. We refer the reader to Nelson's books [188, 190] where another approach to backward processes and equations is developed.

8.3 Calculation of Mean Derivatives for Itô Processes

This section is devoted to the calculation of mean derivatives for processes of diffusion type of the form (8.15). To do this, we first describe a method for calculating conditional expectations under a change of probability measure.

On a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ consider a new probability measure μ. Let μ be absolutely continuous with respect to P with a certain density θ, let $\mathcal{B}$ be a σ-subalgebra of $\mathcal{F}$ and ψ be a measurable map from $(\Omega, \mathcal{F})$ into $\mathbb{R}^n$ equipped with the Borel σ-algebra. Denote by $E^0(\psi|\mathcal{B})$ the conditional expectation of ψ with respect to $\mathcal{B}$ on the probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and by $E'(\psi|\mathcal{B})$ the same expectation on the probability space $(\Omega, \mathcal{F}, \mu)$. Using $E^0(\psi|\mathcal{B})$ we can calculate $E'(\psi|\mathcal{B})$ as follows (cf., e.g., [175]). For any function λ, measurable with respect to $\mathcal{B}$, we have $E'(\lambda\psi) = E'(\lambda E'(\psi|B)) = E^0(\lambda E'(\psi|B)\theta) = E^0(\lambda E'(\psi|B)E^0(\theta|B))$, and on the other hand $E'(\lambda\psi) = E^0(\lambda\psi\theta) = E^0(\lambda E^0(\psi\theta|B))$. Thus

$$E'(\psi|B) = E^0(\theta|B)^{-1}E^0(\psi\theta|B). \tag{8.31}$$

Consider a process of diffusion-type (8.15) and for the sake of simplicity suppose that $\sigma = 1$. For the space of continuous curves $\tilde{\Omega} = C^0([0,l], \mathbb{R}^n)$ and for the σ-algebra $\tilde{\mathcal{F}}$ of cylinder sets on $\tilde{\Omega}$, consider two probability spaces $(\tilde{\Omega}, \tilde{\mathcal{F}}, \nu)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$ where ν is the Wiener measure and the measure μ corresponds to the process $\xi(t)$. Denote the coordinate process on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ by $\zeta(t)$. Recall that $\zeta(t)$, considered as a process on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \nu)$, is a Wiener process. Let us denote this process by $W(t)$. The process $\zeta(t)$, considered on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$, is $\xi(t)$.

It is well-known that if $\xi(t)$ satisfies the condition

$$P\left(\int_0^l a(s)^2 \mathrm{d}s < \infty\right) = 1, \tag{8.32}$$

then μ is absolutely continuous with respect to ν. Under some additional assumptions (see, e.g., [175]) one can show that the density of μ with respect to ν has the form

$$\theta(l) = \exp\left(-\frac{1}{2}\int_0^l a(s)^2 \mathrm{d}s + \int_0^l (a(s)\cdot \mathrm{d}W(s))\right) \tag{8.33}$$

(the above assumptions mean that $\theta(l)$ is a probability density) and so μ and ν are equivalent. For the remainder of this Section we suppose that (8.32) and the assumptions are satisfied. Clearly

$$\theta(l)^{-1} = \exp\left(\frac{1}{2}\int_0^l a^2 \mathrm{d}s - \int_0^l (a(s)\cdot \mathrm{d}W(s))\right).$$

Determine $\theta(t)$ by analogy with formula (8.33) where l is replaced by t. Then using the Itô formula one can easily show that

$$\theta(l) = 1 + \int_0^l \theta(s)(a(s)\cdot \mathrm{d}W(s)) \tag{8.34}$$

(for details see, e.g., [175]).

Denote by E^0 the (conditional) expectation on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \nu)$ and by E' the (conditional) expectation on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$. Then using formulae (8.31) and (8.34) we can calculate

$$
\begin{aligned}
D\xi(t) &= \lim_{\Delta t \to +0} E_t^{\prime\zeta}\left(\frac{\zeta(t+\Delta t)-\zeta(t)}{\Delta t}\right)\\
&= \lim_{\Delta t \to +0} E_t^{0\zeta}(\theta(l))^{-1} E_t^{0\zeta}\left(\frac{\zeta(t+\Delta t)-\zeta(t)}{\Delta t}\theta(l)\right)\\
&= E_t^{0\zeta}(\theta(l))^{-1} \lim_{\Delta t \to +0} E_t^{0\zeta}\left(\frac{\zeta(t+\Delta t)-\zeta(t)}{\Delta t}\right)
\end{aligned}
$$

$$
+E_t^{0\zeta}(\theta(l))^{-1} \lim_{\Delta t \to +0} E_t^{0\zeta}\left(\left(\frac{\zeta(t+\Delta t)-\zeta(t)}{\Delta t}\right)\left(\int_0^l \theta(t)(a(t)\cdot \mathrm{d}W(t)\right)\right).
$$

Let $f : M \to \mathbb{R}$ be a smooth function with compact support. Then since here $\zeta(t) = W(t)$ and so $f(\zeta(t))(\zeta(t+\Delta t)-\zeta(t))) = \int_t^{t+\Delta t} f(W(t))\mathrm{d}W(s)$, we can apply the usual properties of the multiplication of Itô integrals (see Theorem 6.12) to obtain

$$
\begin{aligned}
&E^0(f(\zeta(t)) \lim_{\Delta t \to +0} E_t^{0\zeta}\left(\frac{\zeta(t+\Delta t)-\zeta(t)}{\Delta t}\right)\\
&= \lim_{\Delta t \to +0} E^0\left(f(\zeta(t))\left(\frac{\zeta(t+\Delta t)-\zeta(t)}{\Delta t}\right)\right)\\
&= \lim_{\Delta t \to +0} E^0\left(\frac{\int_t^{t+\Delta t} f(W(t))\theta(s)a(s)\mathrm{d}s}{\Delta t}\right)\\
&= E^0(f(\zeta(t))\theta(t)a(t)).
\end{aligned}
$$

Thus, since f is an arbitrary function of the above-mentioned type,

$$
\lim_{\Delta t \to +0} E_t^{0\zeta}\left(\frac{\zeta(t+\Delta t)-\zeta(t)}{\Delta t}\int_0^t \theta(s)(a(s)\cdot \mathrm{d}W(s)\right) = E_t^{0\zeta}(\theta(t)a(t))
$$

On the other hand, by Theorem 8.7 $D\xi(t) = E_t^{\xi}(a(t))$, thus

$$
E_t^{0\zeta}(\theta(l))^{-1} E_t^{0\zeta}(\theta(t)a(t)) = E_t^{\xi}(a(t)) \tag{8.35}
$$

(note that formula (8.35) can also easily be obtained by direct calculation).

Then

$$
\begin{aligned}
D_*\xi(t) &= \lim_{\Delta t \to +0} E_t^{\prime\zeta}\left(\frac{\zeta(t)-\zeta(t-\Delta t)}{\Delta t}\right)\\
&= E^{0}{}_t^{\zeta}(\theta(l))^{-1} \lim_{\Delta t \to +0} E^{0}{}_t^{\zeta}\left(\frac{\zeta(t)-\zeta(t-\Delta t)}{\Delta t}\theta(l)\right).
\end{aligned}
$$

Using the same arguments as above we easily obtain

$$
\begin{aligned}
&E^0(f(\zeta(t)) \lim_{\Delta t\to+0} E_t^{0\zeta}\left(\frac{\zeta(t)-\zeta(t-\Delta t)}{\Delta t}\theta(l)\right)\\
&= \lim_{\Delta t\to+0} E^0\left((f(W(t))-f(W(t-\Delta t)))\frac{W(t)-W(t-\Delta t)}{\Delta t}\theta(l)\right)\\
&\quad + \lim_{\Delta t\to+0} E^0\left(f(W(t-\Delta t))\frac{W(t)-W(t-\Delta t)}{\Delta t}\theta(l)\right).
\end{aligned}
$$

As above, the second summand on the right hand side is equal to $E^0(f(\zeta(t))\theta(t)a(t))$. Let us calculate the first summand. Here we apply Theorem 6.12, the Itô formula, and integration by parts. Denoting by the same symbol the conditional expectation and the corresponding regression (see Section 6.1.2), we have:

$$
\begin{aligned}
&\lim_{\Delta t\to+0} E^0\left((f(W(t)-f(W(t-\Delta t)))\frac{W(t)-W(t-\Delta t)}{\Delta t}\theta(l)\right)\\
&= \lim_{\Delta t\to+0} E^0([f''(W(t-\Delta t))\Delta t\\
&\qquad + (\mathrm{grad} f(W(t-\Delta t)))(W(t)-W(t-\Delta t))]\frac{W(t)-W(t-\Delta t)}{\Delta t}\theta(l))\\
&= E^0(\mathrm{grad} f(W(t))\theta(l))\\
&= \int_{[0,l]\times\mathbb{R}^n}[(\mathrm{grad} f(W(t)))\theta(l)]\rho^W \mathrm{d}\lambda\\
&= \int_{[0,l]\times\mathbb{R}^n}\left[f(W(t))\left(-\frac{\mathrm{grad}\rho^W}{\rho^W}\right)\theta(l)\right]\rho^W \mathrm{d}\lambda\\
&\qquad + \int_{[0,l]\times\mathbb{R}^n}\left[f(W(t))\left[-\frac{\mathrm{grad} E_t^W(\theta(l))}{\theta(l)}\right]\theta(l)\right]\rho^W \mathrm{d}\lambda\\
&= E^0\left(f(W(t))\left(-\frac{\mathrm{grad}\rho^W}{\rho^W}\right)\theta(l)\right)\\
&\qquad - E^0\left(f(W(t))\left[\theta(l)^{-1}\mathrm{grad} E_t^W(\theta(l))\right]\theta(l)\right)\\
&= E^0\left(f(\zeta(t))\left(\frac{W(t)}{t}\right)\theta(l)\right) - E^0\left(f(\zeta(t))\left[\theta(l)^{-1}\mathrm{grad} E_t^W(\theta(l))\right]\theta(l)\right),
\end{aligned}
$$

where $-\frac{\mathrm{grad}\rho^W}{\rho^W} = \frac{W(t)}{t}$ by Lemma 8.22.

Lemma 8.35 *The following formulae hold:*

$$D_*\xi(t) = E_t^\xi(a(t)) + \frac{\xi(t)}{t} - E_t^\xi(\kappa(t)), \tag{8.36}$$

$$D_*^\xi w(t) = \frac{\xi(t)}{t} - E_t^\xi(\kappa(t)), \tag{8.37}$$

where $\kappa(t) = \theta(l)^{-1}\mathrm{grad} E_t^W(\theta(l))$.

Proof. From the above formulae it follows that

$$D_*\xi(t) = E_t^W(\theta(l))^{-1}\left\{E_t^W(\theta(t)a(t)) + E_t^W\left[\left(\frac{W(t)}{t}\right)\theta(l)\right]\right.$$
$$\left.+E_t^W\left([\theta(l)^{-1}\mathrm{grad}E_t^W(\theta(l))]\theta(l)\right)\right\}$$

and having applied (8.31), (8.34) and (8.35) we obtain (8.36). (8.37) is a consequence of (8.35) and (8.36). □

Lemma 8.36 *Let $g(t)$ and $h(t)$ be L^1-stochastic processes with continuous sample paths in $\mathbb{R}^n$ defined for $t \in [0, l)$ on the same probability space. Consider the process $E_t^h g(t)$. Suppose $Dh(t)$ and $D_*h(t)$ exist. Then:*

(i) *$D^h g(t)$ exists if and only if $D^h E_t^h g(t)$ exists and $D^h E_t^h g(t) = D^h g(t)$;*
(ii) *$D_*^h g(t)$ exists if and only if $D_*^h E_t^h g(t)$ exists and $D_*^h E_t^h g(t) = D_*^h g(t)$.*

Proof. Fix an arbitrary smooth function $f : \mathbb{R}^n \to \mathbb{R}$ with compact support. Using the equality

$$\begin{aligned}
&(E_{t+\Delta t}^h g(t+\Delta t))f(h(t+\Delta t)) - (E_t^h g(t))f(h(t))\\
&= \{(E_{t+\Delta t}^h g(t+\Delta t) - E_t^h g(t)\}f(h(t))\\
&\quad + E_{t+\Delta t}^h g(t+\Delta t))\{f(h(t+\Delta t)) - f(h(t))\}
\end{aligned}$$

we obtain

$$\begin{aligned}
&ED^h\{(E_t^h g(t))f(h(t))\}\\
&= \lim_{\Delta t\to+0} E\left(E_t^h\left(\frac{E_{t+\Delta t}^h g(t+\Delta t))f(h(t+\Delta t)) - (E_t^h g(t))f(h(t))}{\Delta t}\right)\right)\\
&= E((D^h E_t^h g(t))f(h(t))) + E((E_t^h g(t))D_*^h f(h(t))),
\end{aligned}$$

if the limit exists (cf. [188, 190]). Note that the existence of the second summand on the right-hand side follows from the conditions of the Lemma. Thus the limit exists if and only if $D^h E_t^h g(t)$ exists. On the other hand

$$\begin{aligned}
&E(E_t^h\{(E_{t+\Delta t}^h g(t+\Delta t))f(h(t+\Delta t)) - (E_t^h g(t))f(h(t))\})\\
&= E(g(t+\Delta t)f(h(t+\Delta t)) - g(t)f(h(t)))
\end{aligned}$$

and by analogous arguments we obtain

$$ED^h\{(E_t^h g(t))f(h(t))\} = E((D^h g(t))f(h(t))) + E(g(t)D_*^h f(h(t)))$$

if and only if $D^h g(t)$ exists. Clearly,

$$E((E_t^h g(t))D_*^h f(h(t))) = E(g(t)D_*^h f(h(t))),$$

hence

$$E((D^h E_t^h g(t))f(h(t))) = E((D^h g(t))f(h(t))).$$

This proves (i). The proof of (ii) is analogous and is based on the equality

$$\begin{aligned}&(E_{t+\Delta t}^h g(t+\Delta t))f(h(t+\Delta t)) - (E_t^h g(t))f(h(t))\\ &= \{(E_{t+\Delta t}^h g(t+\Delta t)) - E_t^h g(t)\}f(h(t+\Delta t))\\ &\quad + E_t^h g(t))\{f(h(t+\Delta t)) - f(h(t))\}.\end{aligned}$$

□

Lemma 8.37

(i) $D^\xi\left(\frac{\xi(t)}{t}\right) = E_t^\xi\left(\frac{a(t)}{t}\right) - \frac{\xi(t)}{t^2}$.

(ii) $D_*^\xi\left(\frac{\xi(t)}{t}\right) = E_t^\xi\left(\frac{a(t)}{t} - \frac{\kappa(t)}{t}\right)$.

Lemma 8.37 is a corollary of Theorem 8.7 and Lemma 8.35. In particular, (ii) follows immediately from Theorem 8.7 and the construction of the derivative.

Lemma 8.38 *The following equalities hold:*

$$DD_*\xi(t) = D^\xi a(t) + E_t^\xi\left(\frac{a(t)}{t}\right) - \frac{\xi(t)}{t^2}, \tag{8.38}$$

$$D_*D\xi(t) = D_*^\xi a(t), \tag{8.39}$$

$$D^\xi E_t^\xi(\kappa(t)) = 0. \tag{8.40}$$

Proof. To prove (8.40) we apply Lemma 8.36 and formula (8.31) as follows:

$$\begin{aligned}&D^\xi E_t^\xi(\kappa(t))\\ &= D^\xi E_t^\xi\left[\theta(l)^{-1}\mathrm{grad}E_t^W(\theta(l))\right]\\ &= D^\xi\left[\theta(l)^{-1}\mathrm{grad}E_t^W(\theta(l))\right]\\ &= \lim_{\Delta t\to+0} E_t^\xi\left\{\frac{\theta(l)^{-1}\mathrm{grad}\left(E_{t+\Delta t}^W(\theta(l)\right) - E_t^W(\theta(l)))}{\Delta t}\right\}\\ &= E_t^W(\theta(l))^{-1}\lim_{\Delta t\to+0} E_t^W\left(\left\{\frac{\theta(l)^{-1}\mathrm{grad}\left(E_{t+\Delta t}^W(\theta(l)\right) - E_t^W(\theta(l)))}{\Delta t}\right\}\theta(l)\right)\end{aligned}$$

$$\begin{aligned}&= E_t^W(\theta(l))^{-1}\lim_{\Delta t\to+0} E_t^W\left(\frac{\mathrm{grad}\left(E_{t+\Delta t}^W(\theta(l)\right) - E_t^W(\theta(l)))}{\Delta t}\right)\\ &= E_t^W(\theta(l))^{-1}\mathrm{grad}D^W\left(E_t^W\theta(l)\right)\\ &= E_t^W(\theta(l))^{-1}\mathrm{grad}D^W\theta(l)\\ &= 0.\end{aligned}$$

Formulae (8.38) and (8.39) follow from Lemmas 8.35–8.37, Theorem 8.7 and formula (8.40). □

8.4 First Order Differential Equations and Inclusions with Mean Derivatives

Let $a(t,x)$ and $\alpha(t,x)$ be Borel mappings from $[0,T]\times\mathbb{R}^n$ to $\mathbb{R}^n$ and to $\bar{\mathrm{S}}_+(n)$, respectively. According to Definition 8.21 we call the system of the form

$$\begin{cases} D\xi(t) = a(t,\xi(t)), \\ D_2\xi(t) = \alpha(t,\xi(t)), \end{cases} \tag{8.41}$$

a *first order differential equation with forward mean derivatives.*

It is clear that the first equation of (8.41) determines the drift and the second one determines the diffusion coefficient of the process.

Definition 8.39. We say that (8.41) has a *weak solution* on $[0,T]$ with initial condition $\xi(0)=x_0$ if there exists a probability space $(\Omega,\mathcal{F},\mathsf{P})$ and a process $\xi(t)$ given on $(\Omega,\mathcal{F},\mathsf{P})$ and taking values in $\mathbb{R}^n$ such that for almost all $t\in[0,T]$ equation (8.41) is satisfied P-a.s. by $\xi(t)$

Later we shall need the following technical statement.

Lemma 8.40 *Let $\alpha(t,x)$ be a jointly continuous (measurable, smooth) mapping from $[0,T]\times\mathbb{R}^n$ to $\mathrm{S}_+(n)$. Then there exists a jointly continuous (measurable, smooth, respectively) mapping $A(t,x)$ from $[0,T]\times\mathbb{R}^n$ to $L(\mathbb{R}^n,\mathbb{R}^n)$ such that for all $t\in\mathbb{R}$, $x\in\mathbb{R}^n$ the equality $A(t,x)A^*(t,x)=\alpha(t,x)$ holds.*

Proof. Since the symmetric matrices $\alpha(t,x)$ from $\mathrm{S}_+(n)$ are positive definite, all diagonal minors of $\alpha(t,x)$ are positive (in particular, they are not equal to zero). Then the Gauss decomposition is valid for $\alpha(t,x)$ (see [239, Theorem II.9.3]): $\alpha=\zeta\delta z$, where ζ is a lower-triangular matrix with units on the diagonal, z is an upper-triangular matrix with units on the diagonal and δ is a diagonal matrix. In addition, the elements of matrices ζ, δ and z are rationally expressed via the elements of α, hence if the matrices $\alpha(t,x)$ are continuous (measurable, smooth) jointly in t,x, the matrices ζ, δ and z are also continuous (measurable, smooth, respectively) jointly in variables t,x. From the fact that the elements of α are symmetric matrices one can easily derive that $z=\zeta^*$ (i.e., z equals the transpose of ζ). One can also easily see that the elements of the diagonal matrix δ are positive. Thus the diagonal matrix $\sqrt{\delta}$ is well-defined: its diagonal contains the square roots of the corresponding diagonal elements of δ. Consider the matrix $A(t,x)=\zeta\sqrt{\delta}$. By construction $A(t,x)$ is jointly continuous (measurable, smooth, respectively) in t,x and $A(t,x)A^*(t,x)=\zeta(t,x)\delta(t,x)z(t,x)=\alpha(t,x)$. □

If $\alpha(t,x)$ takes values in $\bar{S}_+(n)$ it is possible to construct continuous $A(t,x)$ under some stronger assumptions.

Lemma 8.41 *If $\alpha(t,x)$ is a C^2-smooth mapping from $[0,T]\times\mathbb{R}^n$ to $\bar{S}_+(n)$, there exists a mapping $A(t,x)$ from $[0,T]\times\mathbb{R}^n$ to the space $L(\mathbb{R}^n,\mathbb{R}^n)$ of $n\times n$ matrices, jointly continuous in t,x, such that for all $t\in\mathbb{R}$ and $x\in\mathbb{R}^n$ the equality $A(t,x)A^*(t,x)=\alpha(t,x)$ holds.*

Lemma 8.41 is derived from [75, Theorem 1].

We can now prove several simple existence of solution theorems for (8.41).

Theorem 8.42 *Let $\alpha(t,x)$ in the system (8.41) be jointly continuous in t,x and positive definite (i.e., for all $t\in[0,T]$ and $x\in\mathbb{R}^n$, $\alpha(t,x)$ belongs to $S_+(n)$). In addition, let the estimate*

$$\|\mathrm{tr}\alpha(t,x)\| < K(1+\|x\|)^2 \tag{8.42}$$

hold for some $K>0$. Let $a(t,x)$ be Borel measurable and satisfy the estimate

$$\|a(t,x)\| < K(1+\|x\|) \tag{8.43}$$

for some $K>0$. Then for every initial condition $\xi(0)=x_0\in\mathbb{R}^n$ equation (8.41) has a weak solution that is well-defined on the entire interval $[0,T]$.

Proof. Since $\alpha(t,x)$ is continuous and positive-definite, by Lemma 8.40 there exists a continuous $A(t,x)$ such that $A(t,x)A^*(t,x)=\alpha(t,x)$. Directly from the definition of trace we obtain in this case that $\mathrm{tr}\,\alpha(t,x)$ equals the sum of the squares of the elements of the matrix $A(t,x)$, i.e., it is the square of the Euclidean norm in the space of $n\times n$ matrices. Since in a finite-dimensional space $S(n)$ of symmetric matrices all norms are equivalent, from condition (8.42) it immediately follows that $\|A(t,x)\|<K(1+\|x\|)$ for some $K>0$. Since $\alpha(t,x)$ is positive-definite, the matrix $A(t,x)$ is invertible for all t,x. Since $a(t,x)$ is Borel measurable and satisfies (8.43), the pair $a(t,x)$ and $A(t,x)$ satisfies [83, Theorem III.3.3] and so there exists a weak solution of the stochastic differential equation

$$\xi(t)=\xi_0+\int_0^t a(s,\xi(s))\mathrm{d}s+\int_0^t A(s,\xi(s))\mathrm{d}w(s), \tag{8.44}$$

that is well-defined on the entire interval $[0,T]$. From Theorems 8.7 and 8.12 it follows that the solution $\xi(t)$ of (8.44) P-a.s. satisfies (8.41). □

Theorem 8.43 *Let $\alpha(t,x)$ be C^2-smooth, positive semi-definite (i.e., for all $t\in[0,T]$ and $x\in\mathbb{R}^n$, $\alpha(t,x)$ belongs to $\bar{S}_+(n)$) and satisfy (8.42). Let $a(t,x)$ be continuous and satisfy (8.43). Then for every initial condition $\xi(0)=x_0\in\mathbb{R}^n$ equation (8.41) has a weak solution well-defined on the entire interval $[0,T]$.*

Proof. By Lemma 8.41 there exists a continuous mapping $A(t,x)$ to $L(\mathbb{R}^n,\mathbb{R}^n)$ such that $\alpha(t,x) = A(t,x)A^*(t,x)$. As in the proof of Theorem 8.42, one can obtain the estimate $\|A(t,x)\| < K(1+\|x\|)$ for some $K > 0$. Since $a(t,x)$ and $A(t,x)$ are continuous and (8.43) holds, equation (8.44) satisfies the conditions of [83, Theorem III.2.4], i.e., it has a weak solution well-defined on the entire interval $[0,T]$. It is obvious that P-a.s. the solution satisfies (8.41). □

Now we turn to differential inclusions with mean derivatives. We refer the reader to Section 4.1 for the definitions and main results in the theory of set-valued mappings that we use here.

Let $\mathbf{a}(t,x)$ and $\boldsymbol{\alpha}(t,x)$ be set-valued mappings from $[0,T]\times\mathbb{R}^n$ to $\mathbb{R}^n$ and to $\bar{\mathrm{S}}_+(n)$, respectively. The system

$$\begin{cases} D\xi(t) \in \mathbf{a}(t,\xi(t)), \\ D_2\xi(t) \in \boldsymbol{\alpha}(t,\xi(t)). \end{cases} \tag{8.45}$$

is called *the first order differential inclusion with forward mean derivatives.*

Definition 8.44. We say that (8.45) has a *weak solution* on $[0,T]$ with initial condition $\xi(0) = x_0$ if there exist a probability space $(\Omega,\mathcal{F},\mathsf{P})$ and a process $\xi(t)$ given on $(\Omega,\mathcal{F},\mathsf{P})$ and taking values in $\mathbb{R}^n$ such that P-a.s. and for almost all t (8.45) is satisfied.

Analogous definitions are also valid for inclusions with backward derivatives and with current velocities.

In this section we will mainly look for weak solutions in the class of diffusion type processes.

In the simplest cases the problem of the existence of weak solutions for (8.45) can be reduced to that for (8.41). We present some examples of such statements.

Everywhere below for the set B in $\mathbb{R}^n$ or in $L(\mathbb{R}^n,\mathbb{R}^n)$ we use the norm defined by the formula $\|B\| = \sup\limits_{y\in B}\|y\|$.

Theorem 8.45 *Assume that $\boldsymbol{\alpha}(t,x)$ takes values in positive definite matrices $S_+(n)$, has closed convex values, is lower semicontinuous and for every $\alpha \in \boldsymbol{\alpha}(t,x)$ the estimate*

$$\|\mathrm{tr}\alpha(t,x)\| < K(1+\|x\|)^2$$

holds for a certain $K > 0$. Let $\mathbf{a}(t,x)$ be a Borel measurable set-valued mapping that satisfies the estimate

$$\|\mathbf{a}(t,x)\| < K(1+\|x\|) \tag{8.46}$$

for some $K > 0$. Then for every initial condition $\xi(0) = \xi_0$ there exists a weak solution of (8.45) that is well-defined on the entire interval $[0,T]$.

Proof. Under the hypothesis, by Michael's Theorem (Theorem 4.7) the set-valued mapping $\boldsymbol{\alpha}(t,x)$ has a continuous single-valued selector $\alpha(t,x)$. The

Borel measurable set-valued mapping $\mathbf{a}(t,x)$ has a Borel measurable single-valued selector $a(t,x)$. Then the system

$$\begin{cases} D\xi(t) = a(t,\xi(t)), \\ D_2\xi(t) = \alpha(t,\xi(t)) \end{cases}$$

satisfies the conditions of Theorem 8.42 and so it has a weak solution that is evidently a solution of (8.45). □

Assume that $\boldsymbol{\alpha}(t,x)$ and $\mathbf{a}(t,x)$ are lower semicontinuous, have closed convex values in $\bar{S}_+$ and satisfy the estimates from the hypothesis of Theorem 8.45. Suppose in addition that it is known that a continuous selector $\alpha(t,x)$ of $\boldsymbol{\alpha}(t,x)$ (that exists by Michael's Theorem) is represented in the form $\alpha(t,x) = A(t,x)A^*(t,x)$ with continuous $A(t,x)$. Then one can easily prove the existence of a weak solution of (8.45) by reducing the problem to Theorem 8.43.

Theorem 8.46 *Let $\mathbf{a}(t,x)$ be an upper semicontinuous set-valued mapping with closed convex values from $[0,T]\times\mathbb{R}^n$ to $\mathbb{R}^n$ and let it satisfy the estimate*

$$\|\mathbf{a}(t,x)\| < K(1+\|x\|) \tag{8.47}$$

for some $K>0$.

Let $\boldsymbol{\alpha}(t,x)$ be an upper semicontinuous set-valued mapping with closed convex values from $[0,T]\times\mathbb{R}^n$ to $\bar{\mathrm{S}}_+(n)$ such that for each $\alpha(t,x)\in\boldsymbol{\alpha}(t,x)$ the estimate

$$\|\mathrm{tr}\,\alpha(t,x)\| < K(1+\|x\|)^2 \tag{8.48}$$

holds for some $K>0$.

Then for any initial condition $\xi(0)=\xi_0\in\mathbb{R}^n$ inclusion (8.45) has a weak solution $\xi(t)$, well-defined on the entire interval $t\in[0,T]$, that is a semimartingale.

Proof. For the norm in $\mathrm{S}(n)$ we take the restriction to $\mathrm{S}(n)$ of the Euclidean norm (i.e., the square root of the sum of the squares of the elements of a matrix) in the space $L(\mathbb{R}^n,\mathbb{R}^n)$ isomorphic to $\mathbb{R}^{n^2}$. Since all norms in the finite-dimensional space $\mathrm{S}(n)$ are equivalent to each other, for this norm (8.48) is also valid, perhaps with another constant, for which we keep the notation K.

Since $\mathbf{a}(t,x)$ is an upper semicontinuous set-valued mapping with closed convex values, for any $\varepsilon>0$ it has an ε-approximation (see Section 4.1, in particular Definition 4.10). We shall use the ε-approximations from Theorem 4.11, i.e., for $\varepsilon_i\to 0$ the ε_i-approximations point-wise converge to a Borel measurable selector of the set-valued mapping.

Choose a positive sequence $\varepsilon_i\to 0$. Denote by $a_i(t,x)$ the continuous ε_i-approximations of $\boldsymbol{a}(t,x)$ in $\mathbb{R}^n$ from Theorem 4.11 and by $a(t,x)$ the Borel measurable selector of $\boldsymbol{a}(t,x)$ to which $a_i(t,x)$ converge point-wise. It is clear that all $a_i(t,x)$ and $a(t,x)$ satisfy (8.47) for some constant that is bigger than the constnant K from the condition of the Theorem. Nevertheless, for simplicity, we shall retain the notation K for this constant.

Like $\mathbf{a}(t,x)$, $\boldsymbol{\alpha}(t,x)$ has in $\mathrm{S}(n)$ an ε-approximation from Theorem 4.11 for any $\varepsilon > 0$ since $\boldsymbol{\alpha}(t,x)$ is an upper semicontinuous set-valued mapping with closed convex values. Note that $\bar{\mathrm{S}}_+(n)$ is a convex set in $\mathrm{S}(n)$ and so by Theorem 4.11 those approximations also take values in $\bar{\mathrm{S}}_+(n)$. For the sequence (ε_i) (see above) denote by $\bar{\alpha}_i(t,x)$ an $\frac{\varepsilon_i}{2}$-approximation of $\boldsymbol{\alpha}(t,x)$. Let $\alpha_i(t,x) = \bar{\alpha}_i(t,x) + \frac{\varepsilon_i}{4} I$ where I is the unit matrix. Immediately from the construction it follows that $\alpha_i(t,x)$, for any i, is a continuous ε_i-approximation of $\boldsymbol{\alpha}(t,x)$ and that at any (t,x) it belongs to $\mathrm{S}_+(n)$, i.e., it is strictly positive definite. In addition, $\alpha_i(t,x)$ satisfies (8.48) where the constant $K > 0$ is bigger than the constant from the hypothesis of the Theorem. Also, by construction, the sequence $\alpha_i(t,x)$ point-wise converges to a Borel measurable selector $\alpha(t,x)$ of $\boldsymbol{\alpha}(t,x)$.

By Lemma 8.40 for any i there exists a continuous $A_i(t,x)$ such that $A_i(t,x)A_i^*(t,x) = \alpha_i(t,x)$. Directly from the definition of trace we obtain that $\operatorname{tr}\alpha_i(t,x)$ is equal to the sum of the squares of the elements of $A_i(t,x)$, i.e., it is the square of the Euclidean norm in $L(\mathbb{R}^n,\mathbb{R}^n)$. Hence from (8.48) it follows that $\|A_i(t,x)\| < K(1+\|x\|)$ for some $K > 0$.

Thus the stochastic differential equation

$$\xi(t) = \xi_0 + \int_0^t a_i(s,\xi(s))\mathrm{d}s + \int_0^t A_i(s,\xi(s))\mathrm{d}w(s) \tag{8.49}$$

satisfies the hypothesis of Theorem 6.26 and so it has a weak solution that is well-defined on the entire interval $[0,T]$. Denote this solution by $\xi_i(t)$.

Below in this section we use the measure space $(\tilde{\Omega},\tilde{\mathcal{F}})$ and the family $\tilde{\mathcal{P}}_t$ of σ-subalgebras of $\tilde{\mathcal{F}}$ introduced in Section 6.1.1. On the measure space $([0,T],\mathcal{B})$, where $\mathcal{B}$ is the Borel σ-algebra, we denote the Lebesgue measure by λ_1.

The process $\xi_i(t)$ determines a measure μ_i on $(\tilde{\Omega},\tilde{\mathcal{F}})$. On the probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\mu_i)$ the process $\xi_i(t)$ is the coordinate process, i.e., $\xi_i(t,x(\cdot)) = x(t)$, $x(\cdot)\in\tilde{\Omega}$. In addition it is clear that Lemma 6.28 is valid for measures μ_i and so the set of measures $\{\mu_i\}$ is weakly compact, i.e., it is possible to select a subsequence weakly convergent to some measure μ. Denote by $\xi(t)$ the coordinate process on the probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\mu)$.

Define the measures ν_i on $(\tilde{\Omega},\tilde{\mathcal{F}})$ by the relations $\mathrm{d}\nu_i = (1+\|x(\cdot)\|)\mathrm{d}\mu_i$. By Lemma 6.29 these measures weakly converge to the measure ν given by the relation $\mathrm{d}\nu = (1+\|x(\cdot)\|)\mathrm{d}\mu$.

Since the sequence $a_i(t,x(\cdot))$ converges to $a(t,x(\cdot))$ point-wise, it converges almost surely with respect to all $\lambda\times\mu_k$ and so the functions $\frac{a_i(t,x(\cdot))}{1+\|x(\cdot)\|}$ converge to $\frac{a(t,x(\cdot))}{1+\|x(\cdot)\|}$ almost surely with respect to all $\lambda\times\nu_k$.

Let $\delta > 0$. By Egorov's theorem (see, e.g., [235]) for every k there exists a subset $\tilde{K}_\delta^k \subset [0;T]\times\tilde{\Omega}$ such that $(\lambda\times\nu_k)(\tilde{K}_\delta^k) > 1-\delta$ and the sequence

$\frac{a_i(t,x(\cdot))}{1+\|x(\cdot)\|}$ converges to $\frac{a(t,x(\cdot))}{1+\|x(\cdot)\|}$ on $\tilde{K}^k_\delta$ uniformly. Let $\tilde{K}_\delta = \bigcup_{k=0}^{\infty} \tilde{K}^k_\delta$. Then the sequence $\frac{a_i(t,x(\cdot))}{1+\|x(\cdot)\|}$ converges on $\tilde{K}_\delta$ to $\frac{a(t,x(\cdot))}{1+\|x(\cdot)\|}$ uniformly and $(\lambda \times \nu_k)(\tilde{K}_\delta) > (\lambda \times \nu_k)([0;T] \times \tilde{\Omega}) - \delta$ for all $k = 0, \ldots, \infty$ where $\nu_\infty = \nu$.

Note that $a(t, x(\cdot))$ is continuous on a set of full measure $\lambda \times \nu$ on $[0;T] \times \tilde{\Omega}$. Indeed, consider a sequence $\delta_k \to 0$ and the corresponding sequence $\tilde{K}_{\delta_k}$ from Egorov's theorem. From the above arguments we see that $a(t, x(\cdot))$ is a uniform limit of continuous functions on every $\tilde{K}_{\delta_k}$. Thus this mapping is continuous on every $\tilde{K}_{\delta_k}$ and so on each finite unit $\bigcup_{k=1}^{n} \tilde{K}_{\delta_i}$. We have $\lim_{n\to\infty} (\lambda \times \nu)(\bigcup_{k=1}^{n} \tilde{K}_{\delta_k}) = (\lambda \times \nu)([0;T] \times \tilde{\Omega})$. Thus $\frac{a(t,x(\cdot))}{1+\|x(\cdot)\|}$ is continuous on a set of full measure $\lambda \times \nu$ on $[0;T] \times \tilde{\Omega}$.

Let $g_t(x(\cdot))$ be an arbitrary continuous bounded function on $\tilde{\Omega}$ that is $\tilde{\mathcal{P}}_t$ measurable. In particular let $|g_t(x(\cdot))| < \Xi$ for all $x(\cdot)$ from $\tilde{\Omega}$.

From the above-mentioned uniform convergence of $\frac{a_i(t,x(\cdot))}{1+\|x(\cdot)\|}$ to $\frac{a(t,x(\cdot))}{1+\|x(\cdot)\|}$ on $\tilde{K}_\delta$ for all k and from the boundedness of g_t we obtain that for i large enough

$$\left\| \int_{\tilde{K}_\delta} \left(\int_t^{t+\Delta t} (a_i(\tau, x(\cdot)) - a(\tau, x(\cdot)))\mathrm{d}\tau \right) g_t(x(\cdot))\mathrm{d}\mu_k \right\|$$
$$= \left\| \int_{\tilde{K}_\delta} \left(\int_t^{t+\Delta t} \frac{a_i(\tau, x(\cdot)) - a(\tau, x(\cdot))}{1 + \|x(\cdot)\|}\mathrm{d}\tau \right) g_t(x(\cdot))\mathrm{d}\nu_k \right\| < \delta$$

uniformly for all k. Since $(\lambda \times \mu_k)(\tilde{K}_\delta) > 1 - \delta$ for all k, $\|\frac{a_i(t,x(\cdot))}{1+\|x(\cdot)\|}\| < K$ by (8.47) for all $i = 0, 1, \ldots, \infty$ (where $i = \infty$ corresponds to a) and since $|g_t(x(\cdot))| < \Xi$, we obtain

$$\left\| \int_{\tilde{\Omega}\setminus\tilde{K}_\delta} \left(\int_t^{t+\Delta t} (a_i(\tau, x(\cdot)) - a(\tau, x(\cdot)))\mathrm{d}\tau \right) g_t(x(\cdot))\mathrm{d}\mu_k \right\|$$
$$= \left\| \int_{\tilde{\Omega}\setminus\tilde{K}_\delta} \left(\int_t^{t+\Delta t} \frac{a_i(\tau, x(\cdot)) - a(\tau, x(\cdot))}{1 + \|x(\cdot)\|}\mathrm{d}\tau \right) g_t(x(\cdot))\mathrm{d}\nu_k \right\| < 2Q\Xi\delta.$$

From the last two formulae it follows that for k large enough

$$\left\| \int_{\tilde{\Omega}} \left(\int_t^{t+\Delta t} (a_k(\tau, x(\cdot)) - a(\tau, x(\cdot)))\mathrm{d}\tau \right) g_t(x(\cdot))\mathrm{d}\mu_k \right\| < \delta(2Q\Xi + 1)$$

From the fact that δ is an arbitrary number it follows that

$$\lim_{k\to\infty}\int_{\tilde{\Omega}}\left(\int_t^{t+\Delta t} a_k(\tau,x(\cdot))\mathrm{d}\tau - \int_t^{t+\Delta t} a(\tau,x(\cdot))\mathrm{d}\tau\right) g_t(x(\cdot))\mathrm{d}\mu_k = 0. \tag{8.50}$$

The function $\frac{a(t,x(\cdot))}{1+\|x(\cdot)\|}$ is continuous $\lambda\times\nu$-a.s. (see above) and bounded on $[0;T]\times\tilde{\Omega}$. Hence by a lemma from [82, Section VI.4] we derive from the weak convergence of the measures ν_k to ν that

$$\begin{aligned}
&\lim_{k\to\infty}\int_{\tilde{\Omega}}\left(\int_t^{t+\Delta t} a(\tau,x(\cdot))\mathrm{d}\tau\right) g_t(x(\cdot))\mathrm{d}\mu_k \\
&= \lim_{k\to\infty}\int_{\tilde{\Omega}}\left(\int_t^{t+\Delta t} \frac{a(\tau,x(\cdot))}{1+\|x(\cdot)\|}\mathrm{d}\tau\right) g_t(x(\cdot))\mathrm{d}\nu_k \\
&= \int_{\tilde{\Omega}}\left(\int_t^{t+\Delta t} \frac{a(\tau,x(\cdot))}{1+\|x(\cdot)\|}\mathrm{d}\tau\right) g_t(x(\cdot))\mathrm{d}\nu \\
&= \int_{\tilde{\Omega}}\left(\int_t^{t+\Delta t} a(\tau,x(\cdot))\mathrm{d}\tau\right) g_t(x(\cdot))\mathrm{d}\mu.
\end{aligned} \tag{8.51}$$

Using the same arguments as above, we obtain

$$\begin{aligned}
&\lim_{i\to\infty}\int\limits_{\tilde{\Omega}} (x(t+\Delta t)-x(t))g_t(x(\cdot))\mathrm{d}\mu_i \\
&= \lim_{i\to\infty}\int\limits_{\tilde{\Omega}} \frac{x(t+\Delta t)-x(t)}{1+\|x(\cdot)\|} g_t(x(\cdot))\mathrm{d}\nu_i \\
&= \int\limits_{\tilde{\Omega}} \frac{x(t+\Delta t)-x(t)}{1+\|x(\cdot)\|} g_t(x(\cdot))\mathrm{d}\nu \\
&= \int\limits_{\tilde{\Omega}} (x(t+\Delta t)-x(t))g_t(x(\cdot))\mathrm{d}\mu.
\end{aligned} \tag{8.52}$$

Recall that

$$\begin{aligned}
&\int_{\tilde{\Omega}}\left(x(t+\Delta t)-x(t)-\int_t^{t+\Delta t} a_i(\tau,x(\cdot))\mathrm{d}\tau\right) g_t(x(\cdot))\mathrm{d}\mu_i \\
&= E\left[\left(\xi_i(t+\Delta t)-\xi_i(t)-\int_t^{t+\Delta t} a_k(\tau,\xi_i(\tau))\mathrm{d}\tau\right) g_t(\xi_i(\cdot))\right] \\
&= 0
\end{aligned} \tag{8.53}$$

since $\xi_i(t)$ is a solution of (8.49) and $g_t(\xi_i(\cdot))$ is independent from $\xi_i(t+\Delta t)-\xi_i(t)-\int\limits_t^{t+\Delta t} a_k(\tau,\xi_i(\tau))\mathrm{d}\tau$.

From (8.50), (8.51), (8.52) and (8.53) we obtain that

$$
\begin{aligned}
0 &= \lim_{k\to\infty}\int_{\tilde{\Omega}}\left(x(t+\Delta t)-x(t)-\int_t^{t+\Delta t} a_k(\tau,x(\cdot))\mathrm{d}\tau\right) g_t(x(\cdot))\mathrm{d}\mu_k\\
&= \lim_{k\to\infty}\int_{\tilde{\Omega}}\left(x(t+\Delta t)-x(t)-\int_t^{t+\Delta t} a(\tau,x(\cdot))\mathrm{d}\tau\right) g_t(x(\cdot))\mathrm{d}\mu_k\\
&= \int_{\tilde{\Omega}}\left[(x(t+\Delta t)-x(t))-\int_t^{t+\Delta t} a(\tau,x(\cdot))\mathrm{d}\tau\right] g_t(x(\cdot))\mathrm{d}\mu.
\end{aligned}
$$

Thus

$$
\int_{\tilde{\Omega}}\left[(x(t+\Delta t)-x(t))-\int_t^{t+\Delta t} a(s,x(\cdot))\mathrm{d}s\right] g_t(x(\cdot))\mathrm{d}\mu = 0. \tag{8.54}
$$

Since (8.54) is valid for every g_t, we have proved the following:

Lemma 8.47 *The process $\xi(t)-\int_0^t a(s,\xi(s))\mathrm{d}s$ is a martingale with respect to $\tilde{\mathcal{P}}_t$.*

Define the measures ρ_i on $(\tilde{\Omega},\tilde{\mathcal{F}})$ by the relations $\mathrm{d}\rho_i=(1+\|x(\cdot)\|^2)\mathrm{d}\mu_i$. By Lemma 6.29 these measures weakly converge to the measure ρ defined by the relation $\mathrm{d}\rho=(1+\|x(\cdot)\|^2)\mathrm{d}\mu$.

Using an elementary modification of the above arguments (in particular, replacing the measures ν_k by ρ_k, a_i by α_i, $1+\|x\|$ and $1+\|x\|^2$, etc.) one can easily show that for every bounded continuous function $g_t:\Omega\to\mathbb{R}$ that is measurable with respect to $\tilde{\mathcal{P}}_t$, the relation

$$
\begin{aligned}
&\lim_{i\to\infty}\int_{\Omega}\Bigg[(x(t+\Delta t)-x(t))(x(t+\Delta t)-x(t))^*\\
&\qquad\qquad -\int_t^{t+\Delta t}\alpha_i(s,x(\cdot))\mathrm{d}s\Bigg] g_t(x(\cdot))\mathrm{d}\mu_i\\
&=\int_{\Omega}\Bigg[(x(t+\Delta t)-x(t))(x(t+\Delta t)-x(t))^*\\
&\qquad\qquad -\int_t^{t+\Delta t}\alpha(s,x(\cdot))\mathrm{d}s\Bigg] g_t(x(\cdot))\mathrm{d}\mu
\end{aligned}
$$

holds. Besides, for every i

$$\int_{\Omega} \Big[(x(t+\Delta t) - x(t))(x(t+\Delta t) - x(t))^* \\ - \int_t^{t+\Delta t} \alpha_i(s, x(\cdot))\mathrm{d}s \Big] g_t(x(\cdot))\mathrm{d}\mu_i = 0$$

and so

$$\int_{\Omega} \Big[(x(t+\Delta t) - x(t))(x(t+\Delta t) - x(t))^* \\ - \int_t^{t+\Delta t} \alpha(s, x(\cdot))\mathrm{d}s \Big] g_t(x(\cdot))\mathrm{d}\mu = 0.$$

From this we obtain:

Lemma 8.48 *For the coordinate process $\xi(t)$ on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$ the process $\xi(t)\xi^*(t) - \int_0^t \alpha(s, \xi(\cdot))\mathrm{d}s$ is a martingale with respect to $\tilde{\mathcal{P}}_t$.*

From Lemmas 8.5 and 8.47 it immediately follows that

$$D^{\xi}\left(\xi(t) - \int_0^t a(\tau, \xi(\tau))\mathrm{d}\tau \right) = 0$$

and so $D\xi(t) = a(t, \xi(t)) \in \mathbf{a}(t, \xi(t))$.

Note that from the Definition 8.10 of the quadratic derivative it follows that $D^{\xi}(\xi(t)\xi^*(t)) = D_2\xi(t)$. Then from Lemma 8.48 we obtain that $D_2\xi(t) = \alpha(t, \xi(t)) \in \boldsymbol{\alpha}(t, \xi(t))$.

Thus $\xi(t)$ satisfies (8.45). From Lemma 8.47 it follows that $\xi(t)$ is a semi-martingale. $\square$

Theorem 8.49 *Suppose that $\boldsymbol{\alpha}(t, x)$ takes values in the space $\bar{\mathrm{S}}_+(n)$ of positive semi-definite symmetric matrices, has closed convex values, is lower semicontinuous and for each $\alpha \in \boldsymbol{\alpha}(t, x)$ the following estimate*

$$\|\mathrm{tr}\, \alpha(t, x)\| < K(1 + \|x\|)^2 \tag{8.55}$$

holds for some $K > 0$. Let also $\mathbf{a}(t, x)$ be a Borel measurable set-valued mapping and satisfy the estimate

$$\|\mathbf{a}(t, x)\| < K(1 + \|x\|) \tag{8.56}$$

for some $K > 0$. Then for any initial condition $\xi(0) = \xi_0$ there exists a weak solution of (8.45) that is well-defined on the entire interval $t \in [0, T]$.

Proof. From Michael's theorem it follows that under the conditions of Theorem 8.49 the set-valued mapping $\boldsymbol{\alpha}(t, x)$ has a single-valued continuous selector $\alpha(t,x)$. Obviously $\alpha(t,x)$ belongs to $\bar{\mathrm{S}}_+(n)$ for all t, x. The Borel measurable set-valued mapping $\mathbf{a}(t,x)$ has a Borel measurable single-valued selector $a(t,x)$.

Let $\varepsilon_i \to 0$ be a positive sequence. Define $\alpha_i(t,x) = \alpha(t,x) + \varepsilon_i I$ where I is the unit $n \times n$ matrix. Clearly the α_i are strictly positive definite and continuous. Then by Lemma 8.40 there exists a continuous $A_i(t,x)$ such that $A_i(t,x)A_i^*(t,x) = \alpha_i(t,x)$. Recall that $\operatorname{tr}\alpha_i(t,x)$ is equal to the sum of the squares of the elements of $A_i(t,x)$, i.e., it is the square of the Euclidean norm in $L(\mathbb{R}^n, \mathbb{R}^n)$. Since in the finite-dimensional linear space $\mathrm{S}(n)$ all norms are equivalent, from (8.55) it immediately follows that $\|A(t,x)\| < K(1+\|x\|)$ for some $K > 0$. As $\alpha_i(t,x)$ is positive definite, the matrix $A_i(t,x)$ is invertible for all t, x. Since $a(t,x)$ is measurable and satisfies (8.55), under the above-mentioned properties of $A_i(t,x)$ by [83, Theorem III.3.3] there exists a weak solution of the stochastic differential equation

$$\xi_i(t) = \xi_0 + \int_0^t a(s, \xi_i(s))\mathrm{d}s + \int_0^t A_i(s, \xi_i(s))\mathrm{d}w(s), \tag{8.57}$$

well-defined on the entire interval $t \in [0,T]$. Denote this solution by $\xi_i(t)$. $\xi_i(t)$ determines a measure μ_i on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ where $(\tilde{\Omega}, \tilde{\mathcal{F}})$ was introduced in the proof of Theorem 8.46.

The rest of the proof is analogous to that of Theorem 7.51. All equations (8.57) satisfy the hypothesis of Lemma 6.27. The set of measures $\{\mu_i\}$ is weakly compact so that there exists a subsequence that weakly converges to some measure μ. Denote by $\xi(t)$ the coordinate process on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$. Construct $A(t, x(\cdot))$ by analogy with Theorem 7.51, i.e., as a weak limit in the corresponding L^2 space of the bounded (and so weakly compact) set A_i. The process $\xi(t)$ satisfies the equality $\xi(t) = \xi_0 + \int_0^t a(s, \xi(s))\mathrm{d}s + \int_0^t A(s, \xi(\cdot))\mathrm{d}w$ where $w(t)$ is some Wiener process. Since by construction α_i converges to α uniformly, one can easily show that $E_t^\xi(AA^*) = \alpha$. By Theorem 8.7 and Theorem 8.12, this means that $\xi(t)$ is the weak solution of (8.45) that we are looking for. $\square$

Equations and inclusions with backward mean derivatives arise in the description of some special stochastic processes of mathematical physics. For example (see, e.g., [113, 106, 115]) a second order equation in backward mean derivatives of the group of Sobolev diffeomorphisms may be derived that describes a process whose expectation is a flow of a viscous incompressible fluid. It should be pointed out that the study of such equations and inclusions is generally much more complicated than that of equations and inclusions with forward mean derivatives. Nevertheless there exists a simple method which uses the inverse time direction to solve equations and inclusions with forward mean derivatives, allowing one to obtain results for the case of backward mean derivatives. We refer the reader to [7] for some statements of this sort.

As mentioned in Section 8.1, the notion of current velocity is analogous to that of ordinary velocity for a non-random process. Thus, from the physical point of view, it is most natural to study equations and inclusions with current velocities.

The system

$$\begin{cases} D_S\xi(t) = v(t,\xi(t)) \\ D_2\xi(t) = \alpha(t,\xi(t)) \end{cases} \tag{8.58}$$

is called *the first order differential equation with current velocities.*

Theorem 8.50 *Let $v : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ be smooth and $\alpha : \mathbb{R}^n \to \mathrm{S}_+(n)$ be smooth and autonomous (so it determines the Riemannian metric $\alpha(\cdot,\cdot)$ on $\mathbb{R}^n$, introduced in Section 8.1). Suppose in addition that v and α satisfy the estimates*

$$\|v(t,x)\| < K(1+\|x\|), \tag{8.59}$$

$$\mathrm{tr}\ \alpha(x) < K(1+\|x\|^2) \tag{8.60}$$

for some $K > 0$. Let ξ_0 be a random element with values in $\mathbb{R}^n$ whose probability density ρ_0 with respect to the volume form Λ_α of $\alpha(\cdot,\cdot)$ on $\mathbb{R}^n$ (see Section 8.1) is smooth and nowhere equal to zero. Then for the initial condition $\xi(0) = \xi_0$ equation (8.58) has a weak solution that is well-defined on the entire interval $t \in [0,T]$.

Proof. Since $v(t,x)$ is smooth and the estimate (8.59) is fulfilled, its flow g_t is well-defined on the entire interval $t \in [0,T]$. By $g_t(x)$ we denote the orbit of the flow (i.e., the solution of the equation $x'(t) = v(t,x)$) with the initial condition $g_0(x) = x$. Since $v(t,x)$ is smooth, its flow is also smooth.

The continuity equation (8.21) can clearly be transformed into the form

$$\frac{\partial \rho}{\partial t} = -\alpha(v, \mathrm{Grad}\ \rho) - \rho\ \mathrm{Div}\ v. \tag{8.61}$$

Suppose that $\rho(t,x)$ is nowhere zero in $[0,T]\times\mathbb{R}^n$. Then we can divide (8.61) by ρ so that it is transformed into the equation

$$\frac{\partial p}{\partial t} = -\alpha(v, \mathrm{Grad}\ p) - \mathrm{Div}\ v \tag{8.62}$$

where $p = \log\rho$. Let $p_0 = \log\rho_0$.

We show that the solution of (8.62) with initial condition $p(0) = p_0$ is described by the formula $p(t,x) = p_0(g_{-t}(x)) - \int_0^t (\mathrm{Div}\ v)(s, g_s(g_{-t}(x))\mathrm{d}s$. Consider the function p_0 as given on the level surface $(0,\mathbb{R}^n)$ of the product $[0,T]\times\mathbb{R}^n$. Consider the vector field $(1, v(t,x))$ on $[0,T]\times\mathbb{R}^n$. The orbits of its flow $\hat{g}_t$, starting at the points of $(0,\mathbb{R}^n)$, have the form $\hat{g}_t(0,x) = (t, g_t(x))$ and, like g_t, the flow $\hat{g}_t$ is smooth. Introduce on $[0,T]\times\mathbb{R}^n$ the Riemannian metric $\hat{\alpha}(\cdot,\cdot)$ by the formula $\hat{\alpha}((X_1,Y_1),(X_2,Y_2)) = X_1X_2 + \alpha(Y_1,Y_2)$. Notice that for any (t,x) the point $\hat{g}_{-t}(t,x)$ belongs to $(0,\mathbb{R}^n)$ where the function p_0 is given. Thus on the one hand $(1,v)p(t,x)$, the derivative of $p(t,x)$ in the direction of $(1,v)$, by construction equals $-\mathrm{Div}\ v(t,x)$, and on the other hand one can easily calculate that $(1, v)p(t, x) = \frac{\partial}{\partial t}p(t, x) + \alpha(v(t, x), \mathrm{Grad}\ p(t, x))$. Thus (8.62) is satisfied.

Note that $\rho = e^p$ is indeed nowhere zero, i.e., our arguments are well-defined.

Since $\rho(t, x)$ is well-defined for all $t \in [0, T]$, it determines a process $\xi(t)$ with this probability density and so with initial density ρ_0. By construction $D_S\xi(t) = v(t, \xi(t))$.

Let $u = \frac{1}{2}\mathrm{Grad}\ p = \mathrm{Grad} \log \sqrt{\rho}$ and $a(t, x) = v(t, x) + u(t, x)$.

From Lemma 8.40 and from the hypothesis of the Theorem it follows that there exists a smooth $A(x)$ such that $A(x)A^*(x) = \alpha(x)$ and the relation $\|A(x)\| < K(1 + \|x\|)$ holds. Then $\xi(t)$ satisfies the stochastic differential equation

$$\xi(t) = \xi_0 + \int_0^t a(s, \xi(s))\mathrm{d}s + \int_0^t A(\xi(s))\mathrm{d}w(s) \tag{8.63}$$

and so by Theorem 8.12 $D_2\xi(t) = \alpha(\xi(t))$. □

Lemma 8.51 *Let $\alpha(x)$, $\rho(t, x)$ and Λ_α be as in Theorem 8.50 and let the vector field v from Theorem 8.50 be autonomous. Then the flow $\hat{g}_t$ of the vector field $(1, v(x))$ on $[0, T] \times \mathbb{R}^n$ preserves the volume form $\rho(t, x)\mathrm{d}t \wedge \Lambda_\alpha$ (i.e., $\hat{g}_t^*(\rho(t, x)\mathrm{d}t \wedge \Lambda_\alpha) = \rho_0(x)\mathrm{d}t \wedge \Lambda_\alpha$ where $\hat{g}_t^*$ is the pull-back) and so for any measurable set $Q \subset \mathbb{R}^n$ and for any $t \in [0, T]$ the relation $\int_Q \rho_0(x)\Lambda_\alpha = \int_{g_t(Q)} \rho(t, x)\Lambda_\alpha$ holds.*

Proof. It is enough to show that $L_{(1,v)}(\rho(t, x)\mathrm{d}t \wedge \Lambda_\alpha) = 0$ where $L_{(1,v)}$ is the Lie derivative along $(1, v)$. Clearly $L_{(1,v)}(\rho(t, x)\mathrm{d}t \wedge \Lambda_\alpha) = (L_{(1,v)}\rho(t, x))\mathrm{d}t \wedge \Lambda_\alpha + \rho(t, x)(L_{(1,v)}\mathrm{d}t \wedge \Lambda_\alpha)$. For a function the Lie derivative coincides with the derivative in the direction of the vector field, hence $L_{(1,v)}\rho(t, x) = \frac{\partial\rho}{\partial t} + \alpha(v, \mathrm{Grad}\,\rho)$ (see the proof of Theorem 8.50) and so $(L_{(1,v)}\rho(t, x))\mathrm{d}t \wedge \Lambda_\alpha = (\frac{\partial\rho}{\partial t} + \alpha(v, \mathrm{Grad}\,\rho))\mathrm{d}t \wedge \Lambda_\alpha$. Since neither the form Λ_α nor the vector field $v(x)$ depend on t, $L_{(1,v)}\mathrm{d}t \wedge \Lambda_\alpha = \mathrm{d}t \wedge (L_v\Lambda_\alpha) = \mathrm{Div}\, v\, (\mathrm{d}t \wedge \Lambda_\alpha)$ as the Lie derivative along v of the volume form Λ_α equals $(\mathrm{Div}\, v)\Lambda_\alpha$ (see Section 1.7). Taking into account (8.61), we obtain $L_{(1,v)}(\rho(t, x)\mathrm{d}t \wedge \Lambda_\alpha) = 0$. □

We refer the reader to [7] where differential inclusions with current velocities are considered and a certain existence of solution result is obtained.

8.5 The Case of $\mathcal{P}$-mean Derivatives

In what follows, for the sake of simplicity of presentation, we deal with processes given on a finite time interval $t \in [0, T] \subset \mathbb{R}$.

As in the previous section, we use the measure space $(\tilde{\Omega}, \tilde{\mathcal{F}})$ and the family $\tilde{\mathcal{P}}_t$ introduced in Section 6.1.1.

Let $a : [0, T] \times \tilde{\Omega} \to \mathbb{R}^n$ and $\alpha : [0, T] \times \tilde{\Omega} \to \bar{S}_+(n)$ be measurable.

Definition 8.52. An *equation with $\mathcal{P}$-mean derivatives* is a system of the form

$$\begin{cases} D^{\mathcal{P}}\xi(t) = a(t,\xi(\cdot)), \\ D_2^{\mathcal{P}}\xi(t) = \alpha(t,\xi(\cdot)). \end{cases} \tag{8.64}$$

Definition 8.53. We say that the equation (8.64) has a *weak solution* $\xi(t)$ if there exists a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and a stochastic process $\xi(t)$, given on $(\Omega, \tilde{\mathcal{F}}, \mathsf{P})$ and taking values in $\mathbb{R}^n$, such that equation (8.64) is fulfilled P-a.s.

For simplicity we deal with deterministic initial conditions only.

Lemma 8.54 *For a continuous (measurable, C^k-smooth, $k \geq 1$) mapping $\alpha : [0,T] \times \tilde{\Omega} \to S_+(n)$ satisfying Condition* 4.12*, there exists a continuous (measurable, C^k-smooth, respectively) mapping $A : [0,T] \times \tilde{\Omega} \to L(\mathbb{R}^n, \mathbb{R}^n)$ that satisfies Condition* 4.12 *and such that $\alpha(t,x(\cdot)) = A(t,x(\cdot))A^*(t,x(\cdot))$ for each $(t,x(\cdot)) \in \mathbb{R} \times \tilde{\Omega}$.*

The proof of Lemma 8.54 is a simple modification of that for Lemma 8.40. The fact that $\zeta(t,x(\cdot))$, $\sqrt{\delta}(t,x(\cdot))$, and hence $A(t,x(\cdot))$, satisfies Condition 4.12 follows from the construction.

Theorem 8.55 *Let $a : [0,T] \times \tilde{\Omega} \to \mathbb{R}^n$ and $\alpha : [0,T] \times \tilde{\Omega} \to S_+(n)$ be jointly continuous in $t, x(\cdot)$ and satisfy Condition* 4.12*. Let also the following estimates hold:*

$$\operatorname{tr} \alpha(t,x(\cdot)) < K_1(1 + \|x(\cdot)\|)^2, \tag{8.65}$$

$$\|a(t,x(\cdot))\| < K_2(1 + \|x(\cdot)\|). \tag{8.66}$$

Then for every initial condition $\xi_0 \in \mathbb{R}^n$ equation (8.64) *has a weak solution that is well-defined on the entire interval* $[0,T]$*.*

Proof. Note that $\alpha(t,x(\cdot))$ satisfies the hypothesis of Lemma 8.54 and so there exists a continuous $A(t,x(\cdot))$ such that $A(t,x(\cdot))A^*(t,x(\cdot)) = \alpha(t,x(\cdot))$ and $A(t,x(\cdot))$ satisfies Condition 4.12. Immediately from the definition of trace in this case it follows that $\operatorname{tr} \alpha(t,x(\cdot))$ equals the sum of the squares of all the elements of the matrix $A(t,x(\cdot))$, i.e., it is the square of the Euclidean norm in $L(\mathbb{R}^n, \mathbb{R}^n)$. Since in the finite dimensional vector space all norms are equivalent, from estimate (8.65) it follows that $\|A(t,x(\cdot))\| < K_3(1 + \|x(\cdot)\|)$ for some $K_3 > 0$. Recall that $a(t,x(\cdot))$ is continuous and satisfies Condition 4.12 and the estimate (8.66). Under all these conditions, by [83, Theorem III.2.4] there exists a weak solution $\xi(t)$ of the diffusion type stochastic differential equation

$$\xi(t) = \xi_0 + \int_0^t a(s,\xi(\cdot))\mathrm{d}s + \int_0^t A(s,\xi(\cdot))\mathrm{d}w(s),$$

that is a diffusion type process, well-defined on the entire interval $[0,T]$. From Theorems 8.8 and 8.13 it follows that $\xi(t)$ a.s. satisfies (8.64). $\square$

A more general existence result where $\alpha(t, x(\cdot))$ may take values in $\bar{S}_+(n)$ is obtained in Theorem 8.59 below.

Consider set-valued mappings $\boldsymbol{a}(t, x)$ and $\boldsymbol{\alpha}(t, x)$ sending $[0, T] \times \tilde{\Omega}$ to $\mathbb{R}^n$ and $\overline{S}_+(n)$, respectively, and satisfying Condition 4.12. The *differential inclusion with forward $\mathcal{P}$-mean derivatives* is a system of the form

$$\begin{cases} D^{\mathcal{P}}\xi(t) \in \boldsymbol{a}(t, \xi(\cdot)), \\ D_2^{\mathcal{P}}\xi(t) \in \boldsymbol{\alpha}(t, \xi(\cdot)). \end{cases} \tag{8.67}$$

Definition 8.56. We say that the inclusion (8.67) has a *weak solution* with initial condition $\xi_0 \in \mathbb{R}^n$ if there exists a probability space and a stochastic process $\xi(t)$ given on it, taking values in $\mathbb{R}^n$, such that $\xi(0) = \xi_0$ and a.s. $\xi(t)$ satisfies inclusion (8.67).

As in the case of equations with $\mathcal{P}$-mean derivatives, we deal with deterministic initial conditions only.

If, say, $\boldsymbol{a}(t, x)$ and $\boldsymbol{\alpha}(t, x)$ are lower semi-continuous and have closed convex values, then by Michael's theorem they have continuous selectors $a(t, x(\cdot))$ and $\alpha(t, x(\cdot))$, respectively. If those selectors satisfy the conditions of Theorem 8.55, the weak solution of (8.64) with coefficients $a(t, x(\cdot))$ and $\alpha(t, x(\cdot))$, that exists by Theorem 8.55, is obviously a weak solution of (8.67).

Theorem 8.57 *Let $\boldsymbol{\alpha}(t, x)$ be an upper semi-continuous set-valued mapping from $[0, T] \times \tilde{\Omega}$ to $\overline{S}_+(n)$ with closed convex values that satisfies Condition* 4.12 *and let for every $\alpha(t, x(\cdot)) \in \boldsymbol{\alpha}(t, x(\cdot))$ the estimate*

$$\operatorname{tr} \alpha(t, x(\cdot)) < K_1(1 + \|x(\cdot)\|)^2 \tag{8.68}$$

hold for some $K_1 > 0$.

Let $\boldsymbol{a}(t, x(\cdot))$ be an upper semi-continuous set-valued mapping from $[0, T] \times \tilde{\Omega}$ to $\mathbb{R}^n$ with closed convex values that satisfies Condition 4.12 *and let the estimate*

$$\|\boldsymbol{a}(t, x(\cdot))\| < K_2(1 + \|x(\cdot)\|) \tag{8.69}$$

hold for some $K_2 > 0$.

Then for any initial condition $\xi(0) = \xi_0 \in \mathbb{R}^n$ inclusion (8.67) has a weak solution.

Proof. Choose a sequence of positive numbers $\varepsilon_k \to 0$. The set-valued mapping $\boldsymbol{a}(t, x(\cdot))$ satisfies the conditions of Lemma 4.14 and so there exists a sequence of continuous single-valued mappings $a_k : [0, T] \times \tilde{\Omega} \to \mathbb{R}^n$ that point-wise converges to a measurable selector $a(t, x(\cdot))$ of $\boldsymbol{a}(t, x(\cdot))$ and every $a_k(t, x(\cdot))$ satisfies both Condition 4.12 and the estimate

$$\|a_k(t, x(\cdot))\| < K_2(1 + \|x(\cdot)\|). \tag{8.70}$$

The mapping $\boldsymbol{\alpha}(t, x(\cdot))$ that takes values in the closed convex set $\bar{S}_+(n)$ in the space of all symmetric $n \times n$ matrices also satisfies the conditions of Lemma 4.14 and so there exists a sequence of continuous single-valued mappings $\tilde{\alpha}_k : [0,T]\times\tilde{\Omega}\to\bar{S}_+(n)$ that point-wise converges to a measurable selector $\alpha(t,x(\cdot))$ of $\boldsymbol{\alpha}(t,x(\cdot))$ and every $\tilde{\alpha}_k(t,x(\cdot))$ satisfies both Condition 4.12 and the estimate

$$\operatorname{tr}\tilde{\alpha}_k(t,x(\cdot)) < K_1(1+\|x(\cdot)\|)^2. \tag{8.71}$$

Define another sequence $\alpha_k(t,x(\cdot)) = \tilde{\alpha}_k(t,x(\cdot)) + \varepsilon_k I$ where I is the unit matrix, which evidently also point-wise converges to $\alpha(t,x(\cdot))$. All mappings $\alpha_k(t,x(\cdot))$ are continuous, satisfy Condition 4.12 and estimate (8.71) – at least for k large enough – and in addition they all take values in the convex open set $S_+(n)$ of positive definite symmetric matrices. Thus by Lemma 8.54 for every $\alpha_k(t,x(\cdot))$ there exist continuous $A_k : [0,T]\times\tilde{\Omega} :\to L(\mathbb{R}^n,\mathbb{R}^n)$ such that $\alpha_k(t,x(\cdot)) = A_k(t,x(\cdot))A_k^*(t,x(\cdot))$ and all $A_k(t,x(\cdot))$ satisfy Condition 4.12.

As in Theorem 8.42, immediately from the definition of trace in this case it follows that $\operatorname{tr}\alpha_k(t,x(\cdot))$ equals the sum of the squares of all elements of the matrix $A_k(t,x(\cdot))$, i.e., it is the square of the Euclidean norm of $A_k(t,x(\cdot))$ in $L(\mathbb{R}^n,\mathbb{R}^n)$. Hence, (8.71) means that $\|A_k(t,x(\cdot))\| < K_1(1+\|x(\cdot)\|)$ and so from the latter estimate and (8.70) it follows that for each k the pair $(a_k(t,x(\cdot)),\ A_k(t,x(\cdot)))$ satisfies the Itô condition (6.22) with $K>0$ the same for all k.

Consider the sequence of diffusion type Itô stochastic differential equations

$$\xi_k(t) = \xi_0 + \int_0^t a_k(s,\xi_k(\cdot))\mathrm{d}s + \int_0^t A_k(s,\xi_k(\cdot))\mathrm{d}w(s). \tag{8.72}$$

Since their coefficients are continuous and satisfy Condition 4.12 and estimate (6.22) with the same K, by Theorem 6.26 they all have weak solutions $\xi_k(t)$, well-defined on the entire interval $[0,T]$, and by Lemma 6.28 the set of measures μ_k generated by $\xi_k(t)$ on $(\tilde{\Omega},\tilde{\mathcal{F}})$ is weakly compact. Hence we can choose a subsequence (we retain the notation μ_k for this subsequence) that weakly converges to a probability measure μ. Denote by $\xi(t)$ the coordinate process on the probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\mu)$. Note that $\mathcal{P}_t$ is the "past" σ-algebra of $\xi(t)$.

The fact that $\xi(t)-\int_0^t a(s,\xi(\cdot))\mathrm{d}s$ is a martingale on $(\tilde{\Omega},\tilde{\mathcal{F}},\mu)$ with respect to $\tilde{\mathcal{P}}_t$ and hence

$$E\left([\xi(t+\Delta t)-\xi(t)] - \int_t^{t+\Delta t} a(s,\xi(\cdot))\mathrm{d}s \,\middle|\, \tilde{\mathcal{P}}_t\right) = 0 \tag{8.73}$$

is proved by analogy with Lemma 8.47. From (8.73) it follows that

$$D^{\mathcal{P}}\xi(t) = a(t,\xi(\cdot)) \subset \boldsymbol{a}(t,\xi(\cdot)). \tag{8.74}$$

Now turn to $A_k(t,x(\cdot))$. Recall that $\alpha_k(t,x(\cdot)) = A_k(t,x(\cdot))A_k^*(t,x(\cdot))$ point-wise converges to $\alpha(t,x(\cdot))$, a measurable selector of $\boldsymbol{\alpha}(t,x(\cdot))$.

The fact that the process $[\,\xi(t) \otimes \xi(t)\,] - \int_0^t \alpha(s, \xi(\cdot))\mathrm{d}s$ is a martingale on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$ with respect to $\tilde{\mathcal{P}}_t$ and hence

$$E\left([(\xi(t+\Delta t) - \xi(t)) \otimes (\xi(t+\Delta t) - \xi(t))] - \int_t^{t+\Delta t} \alpha(s, \xi(\cdot)))\mathrm{d}s \;\middle|\; \tilde{\mathcal{P}}_t \right) = 0,$$

from which it follows that

$$D_2^{\mathcal{P}} \xi(t) = \alpha(t, \xi(\cdot)) \subset \boldsymbol{\alpha}(t, \xi(\cdot)), \tag{8.75}$$

is proved by analogy with Lemma 8.48.

Relations (8.74) and (8.75) imply that $\xi(t)$ is the solution of (8.67) that we are looking for. $\square$

Remark 8.58. From (8.73) it follows that the solution $\xi(t)$ of (8.67) obtained in Theorem 8.57 is a semi-martingale with respect to $\tilde{\mathcal{P}}_t$ since $\xi(t) - \int_0^t a(s, \xi(\cdot))\mathrm{d}s$ is a martingale with respect to $\tilde{\mathcal{P}}_t$.

Theorem 8.59 *Let $a(t, x(\cdot))$ and $\alpha(t, x(\cdot))$ be as in Theorem 8.55 with the exception that α sends $[0, T] \times \tilde{\Omega}$ to $\bar{S}_+(n)$ instead of $S_+(n)$. Then for every initial condition $\xi_0 \in \mathbb{R}^n$ equation (8.64) has a weak solution that is well-defined on the entire interval $[0, T]$.*

Indeed, we can construct a sequence of continuous single-valued mappings $\alpha_k = \alpha + \varepsilon_k I : [0, T] \times \tilde{\Omega} \to S_+(n)$ satisfying Condition 4.12 that converge to α. The proof of Theorem 8.59 then follows the same argument as that of Theorem 8.57.

Chapter 9
Mean Derivatives on Manifolds

9.1 Forward and Backward Mean Derivatives

Let a connection H be given on a manifold M. Let $\xi(t)$ be a stochastic process on M. According to formulae (8.1) and (8.2) we can introduce mean forward and mean backward derivatives of $\xi(t)$, if they exist, in any chart. However, from formula (7.19) it follows that for solutions of (7.18) we would obtain the mean derivatives depending on the local connector of the connection H in the chart and even on A, while for physical reasons the derivatives should be vectors. This is why we modify the definition of mean derivatives as follows.

Consider the Borel fields $Y^0(t,\cdot)_\alpha$ and $Y^0_*(t,\cdot)_\alpha$ on a chart $\mathcal{U}_\alpha$ such that the forward (backward, respectively) mean derivative of $\xi(t)$ at t, calculated in $\mathcal{U}_\alpha$, is presented in the form $Y^0(t,\xi(t))_\alpha$ ($Y^0_*(t,\xi(t))_\alpha$, respectively); see Section 8.1. Of course $Y^0(t,\cdot)_\alpha$ and $Y^0_*(t,\cdot)_\alpha$ do not transform as vectors under changes of coordinates. Now construct the vector field $Y^0(t,\cdot)$ (and $Y^0_*(t,\cdot)$) on M whose vector at any $m \in M$ coincides with $Y^0(t,m)_n$ (with $Y^0_*(t,m)_n$, respectively), where $Y^0(t,m)_n$ ($Y^0_*(t,m)_n$, respectively) is calculated in the normal chart $\mathcal{U}_n(m)$ of H at m. Clearly the fields Y^0 and Y^0_* are Borel measurable cross-sections of the tangent bundle TM.

Definition 9.1. $D^{\mathsf{H}}\xi(t) = Y^0(t,\xi(t))$ and $D^{\mathsf{H}}_*\xi(t) = Y^0_*(t,\xi(t))$ are called the *forward* and *backward*, respectively, *mean derivatives* of $\xi(\cdot)$ at t on M with respect to H; $D_S\xi(t) = v^\xi(t,\xi(t))$ and $D_A\xi(t) = u^\xi(t,\xi(t))$ are called the *current* and *osmotic velocities*, respectively, of $\xi(\cdot)$ where $v^\xi(t,m) = \frac{1}{2}(Y^0(t,m) + Y^0_*(t,m))$ and $u^\xi(t,m) = \frac{1}{2}(Y^0(t,m) - Y^0_*(t,m))$.

The current and osmotic velocities do not depend on the connection (see Theorem 9.12 below) and so we do not indicate the connection in the notation. If H is specified, we shall not indicate it in the notation of mean forward and backward derivatives either.

In the same manner as in Definition 9.1 we modify the definitions of $D^y x(t)$ and $D^y_* x(t)$ (see (8.7) and (8.8)).

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics, DOI 10.1007/978-0-85729-163-9_9,

Remark 9.2. Let $f : M \to M_1$ be a smooth mapping of manifolds. Since the value of a mean derivative depends on the "now" σ-algebra of the process, the tangent mapping Tf sends mean derivatives of a process $\eta(t)$ to mean derivatives of the process $\xi(t) = f(\eta(t))$ only in the following form: $Tf(D\eta(t)) = D^{\eta}(\xi(t))$ or $Tf(D^{\xi}\eta(t)) = D\xi(t)$ but generally speaking $Tf(D\eta(t)) \neq D\xi(t)$. An analogous fact in true for backward mean derivatives: $Tf(D_*\eta(t)) = D_*^{\eta}(\xi(t))$ or $Tf(D_*^{\xi}\eta(t)) = D_*\xi(t)$ but generally speaking $Tf(D_*\eta(t)) \neq D_*\xi(t)$.

Lemma 9.3 *Let $\xi(t)$ be a solution of equation* (7.18)*. Then $Y^0(t,m) = a(t,m)$ and so $D\xi(t) = a(t,\xi(t))$.*

Proof. Let $m \in M$ and consider a normal chart $\mathcal{U}_n(m)$ of H at m. In this chart the local connector of H at m is equal to zero, i.e., $\xi(t)$ is described by equation (7.19) with $\boldsymbol{\Gamma}_m(A,A) = 0$. Applying Lemma 8.26(i) and Definition 9.1 we then obtain $D\xi(t)_m = a(t,m)$. Since here both sides of the equation are vectors, the equality remains true in all charts. Using these arguments for all m we obtain the formula $D\xi(t) = a(t,\xi(t))$. □

Recall that an invariant equation, independent of a choice of connection on M, is an Itô equation $(\hat{a}, A)$, a cross-section of an Itô bundle (see Section 7.2). If a connection is specified on M, we can pass to the canonically corresponding Itô vector field (a, A) and obtain the Itô equation in Belopolskaya-Daletskii form (7.18) whose solutions are solutions of $(\hat{a}, A)$ and vice versa. We also use a connection for determining the forward mean derivatives on M.

Lemma 9.4 *For a solution $\xi(t)$ of the Itô equation $(\hat{a}, A)$ its forward mean derivative $D\xi(t)$ with respect to a connection H satisfies the equality*

$$D^{\mathsf{H}}\xi(t) = \hat{a}(t,\xi(t)) + \frac{1}{2}\mathrm{tr}\boldsymbol{\Gamma}_{\xi(t)}(A(t,\xi(t)), A(t,\xi(t)) = \mathcal{H}(\mathcal{A})$$

where $\mathcal{A}$ is the generator of the flow of equation $(\hat{a}, A)$, $\mathcal{H} : \tau M \to TM$ is the mapping generated by the connection H by formula (2.45) *and $\boldsymbol{\Gamma}$ is the local connector of H.*

Proof. By formula (7.17), $\hat{a}(t,m) = a(t,m) - \frac{1}{2}\mathrm{tr}\boldsymbol{\Gamma}_{\xi(t)}(A(t,\xi(t)), A(t,\xi(t))$. Thus the equality $D^{\mathsf{H}}\xi(t) = \hat{a}(t,\xi(t)) + \frac{1}{2}\mathrm{tr}\boldsymbol{\Gamma}_{\xi(t)}(A(t,\xi(t)), A(t,\xi(t))$ follows from Lemma 9.3. The fact that

$$a(t,m) = \hat{a}(t,\xi(t)) + \frac{1}{2}\mathrm{tr}\boldsymbol{\Gamma}_{\xi(t)}(A(t,\xi(t)), A(t,\xi(t)) = \mathcal{H}(\mathcal{A})$$

follows from Lemma 7.26. □

From Lemmas 9.3 and 9.4 it follows that if we apply the same connection both for the transition from $(\hat{a}, A)$ to (a, A) (and hence to equation (7.18)) and for determining the mean derivative, we obtain for a solution $\xi(t)$ that $D\xi(t) = a(t,\xi(t))$. Moreover, if we change the connection, the Itô vector field

(a, A) canonically corresponding to $(\hat{a}, A)$ and the forward mean derivative $D\xi(t)$ will be changed but the equality $D\xi(t) = a(t, \xi(t))$ for those new values will remain true.

For the sake of simplicity, if $\xi(t)$ is a solution of (7.18) we rename the vector $Y_*^0(t, m)$ as $a_*(t, m)$, thus $D_*\xi(t) = a_*(t, \xi(t))$.

Let H be a connection on M. Let (a, A) be an Itô vector field on M. Denote by ∇A the covariant derivative of the field $A(t, m)$ with respect to the connection H. ∇A is a field of bilinear operators $\nabla A(t, m)(\cdot, \cdot) : T_m M \times \mathbb{R}^n \to T_m M$. Consider the field $\nabla A(t, m)(A\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^n \to T_m M$ and the related vector field

$$\operatorname{tr} \nabla A(A)(t, m) = \operatorname{tr} \nabla A(t, m)(A(t, m)(\cdot), \cdot) \tag{9.1}$$

(see the description of the trace in formula (7.9)). Determine on M the following equation

$$\begin{aligned} \mathrm{d}\xi(t) = \exp_{\xi(t)}(&a(t, \xi(t))\mathrm{d}t + \operatorname{tr} \nabla A(A)(t, \xi(t))\mathrm{d}t \\ &- A(t, \xi(t))D_*^{\xi} w(t)\mathrm{d}t + A(t, \xi(t))\mathrm{d}w(t)), \end{aligned} \tag{9.2}$$

where $(a(t, m)\mathrm{d}t + \operatorname{tr} \nabla A(A)(t, m)\mathrm{d}t - A(t, m)D_*^{\xi} w(t)\mathrm{d}t + A(t, m)\mathrm{d}w(t))$ denotes the class of stochastic processes in $T_m M$ consisting of the solutions of all the equations of the form

$$\begin{aligned} X(t + r) = &\int_t^{t+r} \tilde{a}(s, X(s))\mathrm{d}s + \int_t^{t+r} \operatorname{tr} \tilde{A}'(\tilde{A}(s, X(s))\mathrm{d}s \\ &- \int_t^{t+r} \tilde{A}(s, X(s))D_*^X(s)\mathrm{d}s + \int_t^{t+r} \tilde{A}(s, X(s))\mathrm{d}w(s), \end{aligned} \tag{9.3}$$

here $r > 0$, $a(s, X)$ and $A(s, X))$ are analogous to those in Definition 7.27 with the additional assumption that $A(s, X)$ is smooth, and $A' : T_m M \times \mathbb{R}^n \to T_m M$ is the ordinary derivative of A in the vector space $T_m M$.

Let us represent (9.2) in local coordinates in the same manner as (7.18) was represented in the form (7.19). Note the formula $(\exp{}'_m \tilde{A})' = \exp{}''_m(\cdot, \tilde{A}(\cdot)) + \exp{}'_m \tilde{A}'(\cdot, \cdot)$, (where the primes denote derivatives) and the equalities $\exp{}'_m(0) = I$ and $\exp{}''_m(0)(\cdot, \cdot) = -\boldsymbol{\Gamma}_m(\cdot, \cdot)$ that follow from formula (7.20). Thus $\exp_m$ sends the vector $\operatorname{tr} \tilde{A}'(\tilde{A}(t, 0))$, tangent to $T_m M$, to the vector $\exp{}'_m \operatorname{tr} \tilde{A}(\tilde{A}(t, 0)(\cdot), \cdot) = \operatorname{tr} A'(t, m)(A(\cdot), \cdot) + \operatorname{tr} \boldsymbol{\Gamma}_m(A, A) = \operatorname{tr} \nabla A(A)(t, m)$, tangent to M (this clarifies the notation in (9.2)) and $\tilde{A}(t, 0)D_*^X w(t)$ turns into $A(t, m)D_*^{\xi} w(t)$. Summarizing the above formulae we obtain the presentation for (9.2) in the local coordinates of a chart $\mathcal{U}_\alpha$ as follows:

$$\begin{aligned} \mathrm{d}\xi(t) = &a(t, \xi(t))\mathrm{d}t + \operatorname{tr} \nabla A(A)(t, \xi(t))\mathrm{d}t \\ &- A(t, \xi(t))D_*^{\xi} w(t)\mathrm{d}t - \frac{1}{2}\operatorname{tr} \boldsymbol{\Gamma}_{\xi(t)}(A, A)\mathrm{d}t + A(t, \xi(t))\mathrm{d}w(t). \end{aligned} \tag{9.4}$$

Direct verification shows that (9.4) transforms correctly (covariantly) under changes of coordinates. This means that equation (9.2) is well-defined.

Theorem 9.5 *Let $\xi(t)$, $\xi(0) = m_0$, be a strong solution of (9.4). Then $D_*\xi(t) = a(t, \xi(t))$ for $t \in (0, l]$.*

Proof. For the sake of simplicity consider $\xi(t)$ in the normal chart of the initial point m_0. Let $\xi(t) = \exp_{m_0} X(t)$ where

$$X(t) = \int_0^t \tilde{a}(s, X(s))\mathrm{d}s + \int_0^t \operatorname{tr} \tilde{A}(\tilde{A}(s, X(s)))\mathrm{d}s \\ - \int_0^t \tilde{A}(s, X(s)))D_*^X w(s)\mathrm{d}s + \int_0^t \tilde{A}(s, X(s))\mathrm{d}w(s)$$

is a process in $T_{m_0}M$ which exists by construction. Since $D_*\xi(t)$ is a vector, it suffices to show that $D_*X(t) = \tilde{a}(t, X(t))$. The latter is a consequence of (8.29). □

Definition 9.6. The *backward stochastic differential*

$$a(t, m)_*\mathrm{d}_*t + A(t, m)\mathrm{d}_*w_*^{\xi}(t)$$

of a process $\xi(t)$ determined by the Itô vector field (a_*, A) is a class of stochastic processes in the tangent space T_mM consisting of solutions of the (backward) stochastic equations

$$X(t - r) = \int_{t-r}^t \tilde{a}(s, X(s))\mathrm{d}s + \int_{t-r}^t \tilde{A}(s, X(s))\mathrm{d}_*w_*^X(s), \tag{9.5}$$

where $r > 0$, $\tilde{a}(t, m)$ and $\tilde{A}(t, m)$ are analogous to the corresponding terms in Definition 7.27, with the additional assumption that $\tilde{A}(t, m)$ is smooth.

Definition 9.7. An *Itô equation in backward differentials on M* is an expression of the form

$$\mathrm{d}_*\xi(t) = \exp_{\xi(t)}(a_*(t, \xi(t))\mathrm{d}_*t + A(t, \xi(t))\mathrm{d}_*w_*^{\xi}(t)). \tag{9.6}$$

This means that for each t in the domain of $\xi(t)$ the process $\xi(t - r)$, $r > 0$, a.s. coincides with a process from the class $\exp_{\xi(t)}(a_*(t, \xi(t))\mathrm{d}_*t + A(t, \xi(t))\mathrm{d}_*w_*^{\xi}(t))$ until $\xi(t - r)$ leaves some neighborhood of $\xi(t)$.

From formula (8.29) it follows that (9.5) is equivalent to the following equation in T_mM of type (9.3)

$$X(t - r) = \int_{t-r}^t \tilde{a}(s, X(s))\mathrm{d}s + \int_{t-r}^t \operatorname{tr} \tilde{A}'(\tilde{A}(s, X(s))\mathrm{d}s \\ - \int_{t-r}^t \tilde{A}(s, X(s)) \circ D_*^X w(s)\mathrm{d}s + \int_{t-r}^t \tilde{A}(s, X(s))\mathrm{d}w(s). \tag{9.7}$$

Let us describe (9.6) in the local coordinates of a chart $\mathcal{U}_\alpha$. To do this consider (9.7) and then make a transition to the corresponding expression of type (9.4) in local coordinates. Then replace $\mathrm{tr}\,\nabla A(A)(t,m)$ by $\mathrm{tr}\,A'(t,m)(A(\cdot),\cdot)+\mathrm{tr}\,\boldsymbol{\Gamma}_m(A,A)$ (see the derivation of equation (9.4)) and

$$\int_{t-r}^{t}\mathrm{tr}\,A'(s,m)(A(\cdot),\cdot)\mathrm{d}s-\int_{t-r}^{t}A(s,m)\circ D_*^X w(s)\mathrm{d}s+\int_{t-r}^{t}A(s,m)\mathrm{d}w(s)$$

by $\int\limits_{t-r}^{t}A(s,m)\mathrm{d}_*w_*^X(s)$ (cf. formula (8.29)). So we obtain the formula

$$\mathrm{d}_*\xi(t)=a_*(t,\xi(t))\mathrm{d}_*t+\frac{1}{2}\mathrm{tr}\,\boldsymbol{\Gamma}_{\xi(t)}(A,A)\mathrm{d}_*t+A(t,\xi(t))\mathrm{d}_*w_*^{\xi}(t) \tag{9.8}$$

which is the description of (9.6) in local coordinates that we have been searching for.

It should be pointed out that equation (9.6) plays an auxiliary role (as equation (8.30) does in linear spaces). However, it can be considered as an invariant form of formula (9.8) that is convenient for applications.

Let a process $\xi(t)$ be a solution of equation (7.18), $t\in[0,T]$. Fix $t\in[0,T]$. From the above arguments we obtain the following:

Theorem 9.8 *The process $\eta(t)$ satisfying the Itô equation in backward differentials (9.6) with $a_*(t,m)=a(t,m)-\mathrm{tr}\,\nabla A(A\cdot,\cdot)+A(t,m)D_*^{\xi}w(t)$, such that $\eta(t)=\xi(t)$, has the same backward mean derivative at t as $\xi(\cdot)$.*

Thus for small enough $s<t$ such $\eta(s)$ approximates $\xi(s)$.

Remark 9.9. Note the description of mean derivatives for diffusion processes on Riemannian manifolds given in [190]. There, the expressions in local coordinates include Christoffel symbols (i.e., the local connectors) and lead to formulas similar to (7.19) and (9.8). However that presentation is not connected with stochastic differential equations. We have shown that Ito equations in Belopolskaya-Daletskii form are naturally compatible with the machinery of mean derivatives.

We introduce the notation $\hat{a}_*(t,m)=a_*(t,\xi(t))+\frac{1}{2}\mathrm{tr}\,\boldsymbol{\Gamma}_{\xi(t)}(A,A)$. Taking into account formula (2.19), as in the proof of Lemma 7.25 we obtain that $\boldsymbol{\Gamma}_m(X,Y)=-\varphi''_{\alpha n}(X,Y)$ where $\varphi_{\alpha n}$ is the change of coordinates from the normal chart to another chart $\mathfrak{U}_\alpha$. Thus analogously to the proof of Lemma 7.25 one can easily see that under the change of coordinates $\varphi_{\beta\alpha}$ between the charts $\mathfrak{U}_\alpha$ and $\mathfrak{U}_\beta$ the triple $(m,(\hat{a}_*,A))$ transforms according formula (7.16) in the form

$$(m^\alpha,(\hat{a}_*^\alpha,A^\alpha))\mapsto\left(\varphi_{\beta\alpha}m^\alpha,\left(\varphi'_{\beta\alpha}\hat{a}_*^\alpha-\frac{1}{2}\mathrm{tr}\,\varphi''_{\beta\alpha}(A^\alpha,A^\alpha),\varphi'_{\beta\alpha}(A^\alpha)\right)\right) \tag{9.9}$$

Note that (9.9) can be obtained from Lemma 8.33 by replacing f by $\varphi_{\beta\alpha}$.

Thus $(\hat{a}_*,A)$ is a backward Itô equation according to Definition 7.23.

Recall that the process $\xi(t)$ is described by the Itô equation $(\hat{a}, A)$ corresponding to (7.18).

Definition 9.10. The Itô equation $(\hat{a}, A)$ and the backward Itô equation $(\hat{a}_*, A)$ introduced above are said to be *coupled* to each other.

Denote by $\hat{\mathcal{A}}_*$ the backward generator of $(\hat{a}_*, A)$ coupled with $(\hat{a}, A)$ that describes the process $\xi(t)$. In local coordinates it is clearly expressed in the form

$$\hat{\mathcal{A}}_* = -\hat{a}^i_* \frac{\partial}{\partial q^i} + \frac{1}{2}(AA^*)^{ij} \frac{\partial^2}{\partial q^i \partial q^j}. \tag{9.10}$$

Lemma 9.11 *$D^{\mathsf{H}}_* \xi(t) = -\mathcal{H}(\hat{\mathcal{A}}_*)$ where $\mathcal{H}$ is the mapping generated as in formula* (2.45) *by the connection* H*, with respect to which the mean derivatives are calculated.*

Proof. Note that $D^{\mathsf{H}}_* \xi(t) = a_*(t, \xi(t))$ and $\hat{a}_*(t, \xi(t)) = a_* + \frac{1}{2}\mathrm{tr}\, \boldsymbol{\Gamma}_{\xi(t)}(A, A) = \hat{a}^k_* \frac{\partial}{\partial q^k} + \Gamma^k_{ij}(AA_*)^{ij} \frac{\partial}{\partial q^k}$. From formulae (2.45) and (9.10) we obtain that

$$\mathcal{H}(\hat{\mathcal{A}}_*) = -\hat{a}^k_* \frac{\partial}{\partial q^k} + \Gamma^k_{ij}(AA^*)^{ij} \frac{\partial}{\partial q^k}$$
$$= -a^k_* \frac{\partial}{\partial q^k} - \Gamma^k_{ij}(AA_*)^{ij} \frac{\partial}{\partial q^k} + \Gamma^k_{ij}(AA_*)^{ij} \frac{\partial}{\partial q^k} = -a_* \frac{\partial}{\partial q^k}. \qquad \square$$

Lemma 9.11 is "symmetric" to Lemma 9.4.

9.2 Current and Osmotic Velocities

Now consider the current velocity $D_S \xi(t) = \frac{1}{2}(a(t, \xi(t)) + a_*(t, \xi(t)) = v^\xi(t, \xi(t))$ where $v^\xi(t, m) = \frac{1}{2}(a(t, m) + a_*(t, m))$ is the regression.

Theorem 9.12 *$v^\xi(t, m)$ is a vector tangent to M, independent of the choice of connection, with respect to which the forward and backward mean derivatives are calculated.*

Proof. Indeed,

$$\begin{aligned} v^\xi(t, m) &= \frac{1}{2}(a(t, m) + a_*(t, m)) \\ &= \frac{1}{2}\left(a(t, m) - \frac{1}{2}\mathrm{tr}\, \boldsymbol{\Gamma}_{\xi(t)}(A, A) + \frac{1}{2}\mathrm{tr}\, \boldsymbol{\Gamma}_{\xi(t)}(A, A) + a_*(t, m)\right) \\ &= \frac{1}{2}(\hat{a}(t, m) + \hat{a}_*(t, m)). \end{aligned}$$

Under the change of coordinates $\varphi_{\beta\alpha}$ between charts $\mathcal{U}_\alpha$ and $(\mathcal{U})_\beta$ the transformation rule for $\hat{a}(t, m)$ is described in formula (7.12) while $\hat{a}_*(t, m)$ transforms according to formula (7.16). Thus

$$
\begin{aligned}
v^{\xi}(t,m)^{\beta} &= \frac{1}{2}(\hat{a}(t,m)^{\beta} + \hat{a}_*(t,m)^{\beta}) \\
&= \frac{1}{2}\left(\varphi'_{\beta\alpha}\hat{a}(t,m)^{\alpha} + \frac{1}{2}\mathrm{tr}\,\varphi''_{\beta\alpha}(A^{\alpha},A^{\alpha}) + \varphi'_{\beta\alpha}\hat{a}_*(t,m)\right. \\
&\qquad \left. - \frac{1}{2}\mathrm{tr}\varphi''_{\beta\alpha}(A^{\alpha},A^{\alpha})\right) \\
&= \frac{1}{2}\varphi'_{\beta\alpha}(\hat{a}(t,m)^{\alpha} + \hat{a}_*(t,m)^{\alpha}) = \varphi'_{\beta\alpha}v^{\xi}(t,m)^{\alpha}.
\end{aligned}
$$

Hence under coordinate changes $v^{\xi}(t,m)$ transforms by formula (1.1) and so it is a tangent vector, independent of the choice of connection. □

For current velocity (and osmotic velocity below) there is a simple generalization of the presentation that we used in Section 8.1 for $\mathbb{R}^n$ in the case of a smooth non-degenerate diffusion term of a diffusion process.

Let $A(m) : \mathbb{R}^k \to T_mM$ be smooth, autonomous and have rank equal to $\dim M$ at every $m \in M$. Then $\alpha(m) = A(m)A^*(m)$ is smooth, symmetric and non-degenerate. Denote the matrix of α by (α^{ij}). Its inverse (α_{ij}) is smooth, symmetric and non-degenerate, hence it determines on M a Riemannian metric, which we denote by $\alpha(\cdot,\cdot)$. Denote by $\rho^{\xi}(t,x)$ the probability density of $\xi(t)$ with respect to the volume form $\mathrm{d}t \wedge \Lambda_{\alpha} = \sqrt{\det(\alpha_{ij})}\,\mathrm{d}t \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n$ (see Section 8.1).

Theorem 9.13 *For $v^{\xi}(t,m)$ and $\rho^{\xi}(t,m)$ the following generalization of formula* (8.19) *(equation of continuity) holds*

$$
\frac{\partial \rho^{\xi}(t,x)}{\partial t} = -\mathrm{Div}(v^{\xi}(t,x)\rho^{\xi}(t,x)), \tag{9.11}
$$

where Div *denotes divergence with respect to the Riemannian metric $\alpha(\cdot,\cdot)$.*

Theorem 9.13 generalizes formulae (8.19) and (8.21). The proof of Theorem 9.13 is a modification of that for Lemma 8.18 (cf. the proof in [190]).

Consider also the osmotic velocity $D_A\xi(t) = u^{\xi}(t) = \frac{1}{2}(a(t,\xi(t)) - a_*(t,\xi(t))$ whose regression $u^{\xi}(t,m)$ takes the form $u^{\xi}(t,m) = \frac{1}{2}(a(t,m) - a_*(t,m))$. In the case under consideration the following generalization of formulae (8.18) and (8.20) holds:

Theorem 9.14 *For $u^{\xi}(t,m)$ the equality*

$$
u^{\xi}(t,x) = \frac{1}{2}\mathrm{Grad}\log\rho^{\xi}(t,x) = \mathrm{Grad}\log\sqrt{\rho^{\xi}(t,x)} \tag{9.12}
$$

is valid where Grad *denotes the gradient with respect to the Riemannian metric $\alpha(\cdot,\cdot)$.*

The proof of Theorem 9.14 is a modification of that for Lemma 8.17 (cf. the proof in [190]).

9.3 Mean Derivatives of Vector Fields Along Stochastic Processes

The construction of mean derivatives for vector fields along a stochastic process needs a modification that is typical in the transition from vector spaces to manifolds.

Let $Y(t,m)$ be a vector field on M. Consider the invariant mean derivatives

$$\begin{aligned} DY(t,\xi(t)) &= \lim_{\Delta t\to+0} E^{\xi}_t\left(\frac{Y(t+\Delta t,\xi(t+\Delta t))-Y(t,\xi(t))}{\Delta t}\right),\\ D_*Y(t,\xi(t)) &= \lim_{\Delta t\to+0} E^{\xi}_t\left(\frac{Y(t,\xi(t))-Y(t-\Delta t,\xi(t-\Delta t))}{\Delta t}\right), \end{aligned} \tag{9.13}$$

taking values in TTM. Specify a connection H and denote by $K: TTM \to TM$ its connector (see Definition 2.13). Introduce the *covariant mean derivatives* of $Y(t,m)$ along $\xi(t)$ by analogy with the ordinary covariant derivatives via the formulae

$$\begin{aligned} \mathbf{D}Y(t,\xi(t)) &= K\circ \lim_{\Delta t\to+0} E^{\xi}_t\left(\frac{Y(t+\Delta t,\xi(t+\Delta t))-Y(t,\xi(t))}{\Delta t}\right)\\ &= K\circ DY(t,\xi(t)),\\ \mathbf{D}_*Y(t,\xi(t)) &= K\circ \lim_{\Delta t\to+0} E^{\xi}_t\left(\frac{Y(t,\xi(t))-Y(t-\Delta t,\xi(t-\Delta t))}{\Delta t}\right)\\ &= K\circ D_*Y(t,\xi(t)). \end{aligned} \tag{9.14}$$

For an Itô process $\xi(t)$ on M (see Definition 7.82) denote by $\Gamma_{s,t}$ the operator of parallel translation along $\xi(\cdot)$ from $\xi(t)$ to $\xi(s)$ (see the definition and notation in Section 7.7.1). Let $Y(t,m)$ be a C^2-smooth vector field on M. It is easy to see that the covariant mean derivatives $\mathbf{D}Y(t,\xi(t))$ and $\mathbf{D}_*Y(t,\xi(t))$ defined by formulae (9.14) can be equivalently described in this case by the formulae

$$\begin{aligned} \mathbf{D}Y(t,\xi(t)) &= \lim_{\Delta t\to+0} E^{\xi}_t\left(\frac{\Gamma_{t,t+\Delta t}Y(t+\Delta t,\xi(t+\Delta t))-Y(t,\xi(t))}{\Delta t}\right),\\ \mathbf{D}_*Y(t,\xi(t)) &= \lim_{\Delta t\to+0} E^{\xi}_t\left(\frac{Y(t,\xi(t)-\Gamma_{t,t-\Delta t}Y(t-\Delta t,\xi(t-\Delta t))}{\Delta t}\right) \end{aligned} \tag{9.15}$$

Recall that the vector field Y can be considered as a mapping $Y: M \to TM$ with the additional condition $\pi Y = \mathrm{id}$ where $\pi: TM \to M$ is the natural projection and id is the identity mapping (see Definition 1.4). In particular the tangent mapping $TY = (Y, \mathrm{d}Y)$ sends TM to TTM. Let $m \in M$. The restriction of the differential $\mathrm{d}Y$ at T_mM is the derivative (a linear operator) $Y': T_mM \to T_{(m,Y(m))}TM$. Denote by Y'' the second derivative (a bilinear mapping) that sends $T_mM \times T_mM$ to $T_{(m,Y(m))}TM$.

Let $\xi(t)$ be a process on M that is given by an Itô equation $(\hat{a}, A)$ where $A(m)$ is autonomous, smooth and has maximal rank at every m as a mapping $A(m) : \mathbb{R}^k \to \mathbb{R}^n$ for some $k \geq n$. Then, as above, $\alpha(m) = A(m)A^*(m)$ has in local coordinates a symmetric, smooth and non-degenerate matrix (α^{ij}) and its inverse (α_{ij}) defines on M the Riemannian metric $\alpha(\cdot, \cdot)$. Below in this construction we deal with the Levi-Civitá connection H of this Riemannian metric. In particular, in formulae (9.14) K is the connector of this connection, and all covariant derivatives, the Laplace-Beltrami operator and all forward and backward mean derivatives of $\xi(t)$ are calculated with respect to this connection.

Let $\xi(t)$ have mean derivatives $D\xi(t) = a(t, \xi(t))$ and $D_*\xi(t) = a_*(t, \xi(t))$, i.e., the vector fields $a(t, m)$ and $a_*(t, m)$ are the regressions of $D\xi(t)$ and $D_*\xi(t)$, respectively. Then taking into account the forward and backward Itô formulae (6.10) and (6.13) as well as the construction of mean derivatives one can easily see that $DY(t, \xi(t))$ is a vector in $T_{(\xi(t), Y(t,\xi(t)))}TM$ of the form

$$\frac{\partial Y}{\partial t} + Y'(a(t, \xi(t)) + \frac{1}{2}\mathrm{tr}Y''(I, I)$$

and $D_*Y(t, \xi(t))$ is a vector in the same space of the form

$$\frac{\partial Y}{\partial t} + Y'(a_*(t, \xi(t)) - \frac{1}{2}\mathrm{tr}Y''(I, I).$$

By the Definition 2.22 of the covariant derivative $K(Y'(a(t, m)) = \nabla_{a(t,m)}Y$ and $K(Y'(a_*(t, m))) = \nabla_{a_*(t,m)}Y$. One can also easily derive that $K(\frac{1}{2}\mathrm{tr}Y''(I, I) = \frac{1}{2}\nabla^2 Y$ where ∇^2 is the Laplace-Beltrami operator. Thus from the above formulae we obtain the following description of the regressions $\mathbf{D}Y$ and $\mathbf{D}_*Y$ of the covariant derivatives (9.14):

$$\mathbf{D}Y = \frac{\partial Y}{\partial t} + \nabla_a Y + \frac{1}{2}\nabla^2 Y, \tag{9.16}$$

$$\mathbf{D}_*Y = \frac{\partial Y}{\partial t} + \nabla_{a_*} Y - \frac{1}{2}\nabla^2 Y.$$

Formulae (9.16) and (9.17) are natural analogs of (8.24) and (8.25), respectively.

We now turn to the case where we have to use a Riemannian metric specified *a priori* on M (i.e., not the metric generated by the diffusion coefficient of an equation). In this case we find formulae for covariant mean derivatives along an Itô process, using a modification of the construction of covariant derivatives based on parallel translation. We use Itô processes and parallel translation with respect to the Levi-Civitá connection of the above metric.

Let $\xi(t)$ be an Itô development of the Itô process $\zeta(t)$ given in a certain tangent space by the formula $\zeta(t) = \int_0^t a(s)\mathrm{d}s + \sigma w(t)$ where $a(t)$ satisfies (7.54). Let $D\xi(t) = a(t, \xi(t))$ and $D_*\xi(t) = a_*(t, \xi(t))$. Applying the Itô formulae together with formulae (9.15), we obtain the following modifications of formulae (9.16) and (9.17):

$$\mathbf{D}Y = \frac{\partial Y}{\partial t} + \nabla_a Y + \frac{\sigma^2}{2}\nabla^2 Y, \qquad (9.17)$$
$$\mathbf{D}_* Y = \frac{\partial Y}{\partial t} + \nabla_{a_*} Y - \frac{\sigma^2}{2}\nabla^2 Y.$$

9.4 The Quadratic Mean Derivative

For a stochastic process $\xi(t)$ on a probability space $(\Omega, \mathcal{F}, P)$ with values in a manifold M we introduce its quadratic derivative as follows (cf. Definition 8.10 for the case of linear spaces). Take any chart $\mathcal{U}$ and consider in it the L^1 random variable determined by the rule

$$D_2\xi(t) = \lim_{\triangle t \to +0} E_t^{\xi}\left(\frac{(\xi(t+\triangle t) - \xi(t)) \otimes (\xi(t+\triangle t) - \xi(t))}{\triangle t}\right), \qquad (9.18)$$

where the limit is assumed to exist in $L^1(\Omega, \mathcal{F}, P)$.

Definition 9.15. $D_2\xi(t)$ is called the *quadratic mean derivative* of the process $\xi(t)$ on M at time t.

Notice that for $D_2\xi(t)$ there exists a regression in any chart, i.e., a measurable field $\alpha^0(t,m)$ such that $D_2\xi(t) = \alpha^0(t, \xi(t))$.

An important geometric feature of the quadratic mean derivative is that (like current velocity) it is independent of the choice of connection; its regression is a $(2,0)$-tensor field.

Let $\xi(t)$ be given by an Itô equation $(\hat{a}, A)$ (see Section 7.3), i.e., in particular, it is Markovian. Recall that here $A(t,m)$ is a field of linear operators $A(t,m) : \mathbb{R}^k \to T_m M$ with k sufficiently large.

Lemma 9.16 *Suppose that $\xi(t)$ is given by an Itô equation $(\hat{a}, A)$. Then*

$$D_2\xi(t) = A(t,\xi(t))A^*(t,\xi(t)) = 2(\mathfrak{Q}\mathcal{A})(t,\xi(t)) \qquad (9.19)$$

where A^* is the conjugate operator, $\mathcal{A}$ is the corresponding generator and $\mathfrak{Q}$ is defined by formula (2.44). *$A(t,m)A^*(t,m)$ is a $(2,0)$-tensor field on M. In particular, $D_2\xi(t) = 0$ if and only if $\xi(t)$ has C^1-smooth sample paths.*

Proof. Let H be a connection and represent $(\hat{a}, A)$ as an Itô equation in Belopolskaya-Daletskii form in terms of the Itô vector field (a, A) canonically corresponding to $(\hat{a}, A)$ with respect to H (see Section 7.3). Then in a local chart $\xi(t)$ is expressed in Baxendale form

$$
\begin{aligned}
\xi(t+\triangle t)-\xi(t) = &\int_t^{t+\triangle t} a(s,\xi(s))\mathrm{d}s \\
&-\frac{1}{2}\int_t^{t+\triangle t} \mathrm{tr}\boldsymbol{\Gamma}_{\xi(t)}(A(t,\xi(t)),A(t,\xi(t)))\mathrm{d}s \\
&+\int_t^{t+\triangle t} A(s,\xi(s))\mathrm{d}w(s)
\end{aligned}
$$

where $\boldsymbol{\Gamma}_m(\cdot,\cdot)$ is the local connector of H in that chart. As in the proof of Theorem 8.12, by direct calculation it follows that the components of $(\xi(t+\triangle t)-\xi(t))\otimes(\xi(t+\triangle t)-\xi(t))$ are elements of the matrix $(\xi(t+\triangle t)-\xi(t))(\xi(t+\triangle t)-\xi(t))^*$ where we use the matrix multiplication of the column-vector $(\xi(t+\triangle t)-\xi(t))$ and the row-vector $(\xi(t+\triangle t)-\xi(t))^*$ (i.e., the transpose of $(\xi(t+\triangle t)-\xi(t))$). In particular, this matrix product is a symmetric positive semi-definite matrix.

Taking into account the properties of the Lebesgue and Itô integrals one can see that $(\xi(t+\triangle t)-\xi(t))(\xi(t+\triangle t)-\xi(t))^*$ is approximated by the sum

$$
\begin{aligned}
&a(t,\xi(t))(a(t,\xi(t)))^*(\Delta t)^2 \\
&-\frac{1}{2}a(t,\xi(t))\,(\mathrm{tr}\Gamma(A(t,\xi(t)),A(t,\xi(t))))^*(\Delta t)^2 \\
&+a(t,\xi(t))\Delta t)(A(t,\xi(t))\Delta w(t))^* \\
&-\frac{1}{2}\mathrm{tr}\Gamma(A(t,\xi(t)),A(t,\xi(t)))(a(t,\xi(t)))^*(\Delta t)^2 \\
&+\frac{1}{4}\mathrm{tr}\Gamma(A(t,\xi(t)),A(t,\xi(t)))(\mathrm{tr}\Gamma(A(t,\xi(t)),A(t,\xi(t))))^*(\Delta t)^2 \\
&+(A(t,\xi(t))\Delta w(t))(a(t,\xi(t))\Delta t)^* \\
&+A(t,\xi(t))(A(t,\xi(t)))^*\Delta t.
\end{aligned}
$$

In this expression only the last summand is an infinitesimal of the same order as Δt while the others are infinitesimals of higher order. Then applying formula (9.18) we obtain that $D_2\xi(t)=E_t^{\xi}(A(t,\xi(t))A^*(t,\xi(t)))$. Thus (9.19) follows and it is obvious that this expression is independent of H. Recall (see Lemma 7.2) that AA^* is a $(2,0)$-tensor field.

The fact that $D_2\xi(t)=2\mathfrak{Q}\mathcal{A}$ now follows from the definitions.

From (9.19) it evidently follows that the equality $D_2\xi(t)=0$ means that $A(t,m)=0$ and so the sample paths of $\xi(t)$ are C^1-smooth. If they are C^1-smooth, from the definition of quadratic mean derivative (see Theorem 8.12) we have $D_2\xi(t)=0$ (as always for C^1-curves). □

Recall that $A(t,m)A^*(t,m)$ is the diffusion coefficient of $\xi(t)$. Note also that if AA^* is non-degenerate, smooth and autonomous, it can be considered as a metric $(2,0)$-tensor so that its inverse is a Riemannian metric on M. In particular, this metric can be used to determine forward and backward mean derivatives of $\xi(t)$. These derivatives turn out to have many uses.

So, we can consider differential equations and inclusions in terms of forward (or backward) mean derivatives or current velocities on the one hand and quadratic mean derivatives on the other hand by analogy with Section 8.4. In doing this we have to be able to represent a given $(2,0)$-tensor field $\alpha(t,m)$ in the form $\alpha(t,m) = A(t,m)A^*(t,m)$. Unlike the case of linear spaces (see Section 8.4), on non-parallelizable manifolds there is a topological obstruction for obtaining such a presentation for smooth or continuous $\alpha(t,m)$ with smooth or at least continuous $A(t,m) : \mathbb{R}^n \to T_mM$ where n is the dimension of M (see, e.g., [150]). Nevertheless such a presentation is possible in larger dimensions under some additional assumptions.

Theorem 9.17 *Consider a symmetric* $(2,0)$*-tensor field* $\alpha(t,m)$ *on an* n*-dimensional manifold* M*. There exists a* $k > n$ *such that, if* $\alpha(t,m)$ *is positive definite and continuous (smooth), there exists a continuous (smooth, respectively) field of linear operators* $A(t,m) : \mathbb{R}^k \to T_mM$ *such that the relation* $\alpha(t,m) = A(t,m)A^*(t,m)$ *holds.*

Proof. Denote by (α^{ij}) the matrix of $\alpha(t,m)$ in the local coordinates of some chart $\mathcal{U}_\alpha$. Introduce an arbitrary Riemannian metric $g(\cdot,\cdot)$ on M with matrix (g_{ij}) and denote by $\bar{g}$ the corresponding metric $(2,0)$-tensor with matrix (g^{ij}) (see Notation 1.51 and Remark 1.52). Then the $(2,0)$-tensor field α is represented in the form $\alpha(\cdot,\cdot) = \bar{g}(b(\cdot),\cdot)$ where $b(t,m)(\cdot)$ is a $(1,1)$-tensor field of self-adjoint linear operators acting in the cotangent spaces to M. The existence and uniqueness of b is derived as follows. In the local coordinates of $\mathcal{U}_\alpha$ the expression $\alpha(\cdot,\cdot) = \bar{g}(b(\cdot),\cdot)$ takes the form $\alpha^{ij} = g^{ik}b^j_k$ where b^i_j are the coefficients of $(b)^\alpha$. Since (g_{ij}) is not degenerate, b^i_j are presented in the form $b^i_j = g_{jk}\alpha^{ik}$. Since the metric tensor is C^∞-smooth, the field b is continuous (smooth, or Borel measurable) if α is continuous (smooth, Borel measurable, respectively).

Let $a^1, \dots, a^n$ be a field of orthonormal frames with respect to the metric $\bar{g}(\cdot,\cdot)$ in cotangent spaces at the points of $\mathcal{U}_\alpha$. With respect to these frames b is represented by a matrix $(\bar{b})$ that is symmetric, positive definite and satisfies the conditions for applying the Gauss decomposition as in the proof of Lemma 8.40. Hence $b(t,m)$ is presented in the form $b(t,m) = f(t,m)f^*(t,m)$.

Embed by Nash's Theorem (Theorem 1.46) the manifold M with metric $g(\cdot,\cdot)$ isometrically into a Euclidean space $\mathbb{R}^k$ (see [186]). k can be chosen so that it depends only on M (by Nash's theorem it is determined by the dimension M), i.e., it is the same for all Riemannian metrics on M.

Denote by P_m the orthogonal projector of $\mathbb{R}^k$ onto its subspace T_mM (the tangent space to M at $m \in M$). Let $A(t,m) = f(t,m) \circ P_m : \mathbb{R}^k \to T_mM$. Then one can easily see that $\alpha(t,m) = A(t,m)A^*(t,m)$ and by construction $A(t,m)$ is continuous (smooth, respectively). $\square$

9.5 Mean Derivatives of Itô Processes on Manifolds

Let a Riemannian manifold M satisfy the hypothesis of Theorem 7.76 or be uniformly complete (see Definition 7.85). Under these assumptions the integral approach to stochastic differential equations on manifolds of Section 7.7 is well-posed.

Lemma 9.18 *Let $\xi(t) = R_I z(t)$ where $z(t) = \int_0^t a(s)\mathrm{d}s + \int_0^t A(s)\mathrm{d}w(s)$ is an Itô process in $T_{m_0}M$. Then $\xi(t)$ satisfies the following Itô equation in Belopolskaya-Daletskii form:*

$$\mathrm{d}\xi(t) = \exp_{\xi(t)}(\mathit{\Gamma}_{t,0}a(t)\mathrm{d}t + \mathit{\Gamma}_{t,0}A(t)\mathrm{d}w(t)). \tag{9.20}$$

From the Definition 7.83 of parallel translation it easily follows that Lemma 9.18 is a reformulation of Lemma 7.65. Clearly the equation in local coordinates (i.e., the Baxendale form) for (9.20) is as follows:

$$\mathrm{d}\xi(t) = \mathit{\Gamma}_{t,0}a(t)\mathrm{d}t - \frac{1}{2}\mathrm{tr}\boldsymbol{\mathit{\Gamma}}_{\xi(t)}(\mathit{\Gamma}_{t,0}A(t), \mathit{\Gamma}_{t,0}A(t))\mathrm{d}t + \mathit{\Gamma}_{t,0}A(t)\mathrm{d}w(t) \tag{9.21}$$

(recall that $\mathit{\Gamma}_{t,0}$ is the operator of parallel translation while $\boldsymbol{\mathit{\Gamma}}_m(\cdot,\cdot)$ is the local connector).

Taking into account formulae (9.20) and (9.21), we can apply the results of the preceding sections to the calculation of the mean derivatives of Itô processes in manifolds. In particular, from Lemma 9.3 we have:

Lemma 9.19 *For $\xi(t)$ as in Lemma 9.18, $D\xi(t) = E_t^{\xi}(\mathit{\Gamma}_{t,0}a(t))$.*

Also, from Lemma 9.16 we have:

Lemma 9.20 *For $\xi(t)$ as in Lemma 9.18, $D_2\xi(t) = E_t^{\xi}(\mathit{\Gamma}_{t,0}(A(t)A^*(t)))$.*

For the applications below we have to calculate mean derivatives for an Itô process $\xi(t) = R_I z(t)$ where $z(t)$ is an Itô process in $T_{m_0}M$ for $t \in [0, l]$ of the form $z(t) = \int_0^t a(s)\mathrm{d}s + \sigma w(t)$ where $\sigma > 0$ is a real constant and $a(t)$ satisfies (8.32). Recall (see Section 7.7.3) that in this case we can deal with a stochastically complete M (a less restrictive assumption than above).

So, let M be stochastically complete. For the above $\xi(t)$, formula (9.20) takes the form

$$\mathrm{d}\xi(t) = \exp_{\xi(t)}(\mathit{\Gamma}_{t,0}\, a(t)\mathrm{d}t + \mathit{\Gamma}_{t,0}\, \mathrm{d}w(t)) \tag{9.22}$$

so that (9.21) takes the form

$$\mathrm{d}\xi(t) = \mathit{\Gamma}_{t,0}\, a(t)\mathrm{d}t - \frac{1}{2}\mathrm{tr}\mathit{\Gamma}_{\xi(t)}(I, I)\mathrm{d}t + \mathit{\Gamma}_{t,0}\, \mathrm{d}w(t). \tag{9.23}$$

Thus by Lemma 9.20

$$D_2\xi(t) = \sigma^2 I \tag{9.24}$$

and by Lemma 9.19
$$D\xi(t) = E_t^{\xi}(\Gamma_{t,0}\, a(t)) \tag{9.25}$$
(this also follows from Lemma 9.19) and
$$D_*\xi(t) = E_t^{\xi}(\Gamma_{t,0}\, a(t)) + D_*^{\xi}(\Gamma_{t,0}\, w(t)). \tag{9.26}$$
in any chart in M. Consequently
$$D_S\xi(t) = v^{\xi}(t,\xi(t)) = E_t^{\xi}(\Gamma_{t,0}\, a(t)) + \frac{1}{2}D_*^{\xi}(\Gamma_{t,0}\, w(t)) \tag{9.27}$$
and
$$D_A\xi(t) = u^{\xi}(t,\xi(t)) = -\frac{1}{2}D_*^{\xi}(\Gamma_{t,0}\, w(t)) \tag{9.28}$$
also in any chart.

Lemma 9.21 *Formulae* (9.11) *and* (9.12) *hold for the above mentioned process* $\xi(t) = R_I z(t)$.

Indeed, the fact that the coefficient at $w(t)$ is σI means that the Riemannian metric is generated by the diffusion coefficient as in Section 9.2.

Let $a(t)$ satisfy (8.32) so that the measure μ^z corresponding to $z(t) = \int_0^t a(s)\mathrm{d}s + w(t)$ on the space $(C^0([0,l], T_{m_0}M), \tilde{\mathcal{F}})$ is absolutely continuous with respect to the Wiener measure ν with density (8.33). Then formula (9.26) and consequently formulae (9.27) and (9.28) can be expressed in the more precise form:

Lemma 9.22
$$D_*\xi(t) = E_t^{\xi}(\Gamma_{t,0}a(t)) + E_t^{\xi}\left[\Gamma_{t,0}\left(\frac{z(t)}{t} - \kappa(t)\right)\right] \tag{9.29}$$
$$D_S\xi(t) = E_t^{\xi}(\Gamma_{t,0}a(t)) + \frac{1}{2}E_t^{\xi}\left[\Gamma_{t,0}\left(\frac{z(t)}{t} - \kappa(t)\right)\right] \tag{9.30}$$
$$D_A\xi(t) = -\frac{1}{2}E_t^{\xi}\left[\Gamma_{t,0}\left(\frac{z(t)}{t} - \kappa(t)\right)\right] \tag{9.31}$$
where $\kappa(t)$ *is as defined in Lemma* 8.35.

Proof. In order to derive (9.29) note that the probability of the event $\xi(t) \in A$, where $A \subset M$ is a Borel set, is equal to $\mu^z(\xi(t)^{-1}A)$, where $\xi(t)^{-1}A$ is the set of curves from $C^0([0,l], T_{m_0}M)$ such that the values of their developments at t belong to A. Denote by $\Xi(t)$ the set in $T_{m_0}M$ consisting of the values at t of all curves from $\xi(t)^{-1}A$. Generally speaking, $\mathcal{N}_t^z$ and $\mathcal{N}_t^{\xi}$ do not coincide, in particular $\Xi(t)$ may not belong to $\mathcal{N}_t^z$ (but it certainly belongs to $\mathcal{P}_t^z$ by construction). But we can calculate $\mu^z(\xi(t)^{-1}A)$ by integrating the probability density of $z(t)$ over $\Xi(t)$. This means that the value $\xi(t)$ is distributed as a map from $\tilde{\Omega}$ into $T_{m_0}M$ which is measurable with respect to $\mathcal{N}_t^\xi$ and has the same probability distribution as $z(t)$. Taking into account the relation between the probability distributions and mean derivatives (see Lemma 8.17) as well as the construction of the Itô development, we obtain (9.29). Formulae (9.30) and (9.31) follow immediately from (9.29) and from (9.27) and (9.28), respectively. □

Formulae (9.25) and (9.26), and consequently (9.29) and (9.30), can also be derived from the following general statement. Let M be a Riemannian manifold and $z(t)$ be an Itô process in a certain $T_{m_0}M$ such that $\xi(t) = R_I z(t)$ exists. Let $y(t)$ be another process given on the same probability space.

Theorem 9.23

(i) *$D^y\xi(t)$ exists if and only if $D^y z(t)$ exists and $D^y\xi(t)$ is parallel to $D^y z(t)$ along $\xi(\cdot)$.*
(ii) *$D_*^y\xi(t)$ exists if and only if $D_*^y z(t)$ exists and $D_*^y\xi(t)$ is parallel to $D_*^y z(t)$ along $\xi(\cdot)$.*

This statement for $\xi(t) = R_I z(t)$ is an analog of the characteristic property of the curve $\mathcal{S}v(t)$ from Theorem 3.43. It follows directly from the construction of $R_I z(t)$ (see Section 7.6.1). The principal point here is that the conditional expectation with respect to the same σ-algebra $\mathcal{N}_t^y$ is used in the mean derivatives before and after the parallel translation. Note that $D\xi(t)$ and $Dz(t)$ (as well as $D_*\xi(t)$ and $D_*z(t)$, respectively) are not, generally speaking, parallel to each other along $\xi(\cdot)$ since $\mathcal{N}_t^\xi$ and $\mathcal{N}_t^z$ may not coincide.

9.6 Equations and Inclusions with Mean Derivatives

In this section we introduce differential equations and inclusions with mean derivatives on manifolds and prove some simple existence of solution theorems.

Let M be a Riemannian manifold. In this section we use the Levi-Civitá connection H of the Riemannian metric.

Take $t \in [0, T]$. Consider a vector field $a(t, m)$ and a symmetric positive semi-definite $(2, 0)$-tensor field $\alpha(t, m)$ on M.

By a *first order differential equation with mean derivatives* we mean a system of the form

$$\begin{cases} D\xi(t) = a(t, \xi(t)), \\ D_2\xi(t) = \alpha(t, \xi(t)). \end{cases} \tag{9.32}$$

Definition 9.24. We say that (9.32) has a *weak solution* on $[0, T]$ with initial condition $\xi(0) = m_0$ if there exists a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and a process $\xi(t)$ given on $(\Omega, \mathcal{F}, \mathsf{P})$, taking values in M, such that P-a.s. and for almost all t in $[0, T]$ system (9.32) is satisfied.

We shall mainly look for weak solutions of (9.32) among solutions of some equations of type (7.18). Taking into account Theorem 9.5 and Lemma 9.16 it is clear that the first equation of (9.32) determines the drift of a solution of some equation of type (13.5) while the second equation determines the diffusion coefficient. However, we do not assume *a priori* that a solution of (9.32) is also a solution of an equation of the form (7.18). This is why we do not consider the notion of strong solutions.

For each (t, m) denote by $\mathcal{A}_{a,\alpha}(t, m)$ the generator determined by the formulae $\mathcal{H}\mathcal{A}_{a,\alpha}(t, m) = a(t, m)$ and $\mathcal{Q}\mathcal{A}_{a,\alpha}(t, m) = \alpha(t, m)$ (see formulae (2.45) and (2.44)).

Theorem 9.25 *Assume that in* (9.32) *$a(t, m)$ and $\alpha(t, m)$ are C^1-smooth, $\alpha(t, m)$ is positive definite and $\mathcal{A}_{a,\alpha}(t, m)$ satisfies the conditions of Theorem* 7.43*. Then for any initial condition $\xi(0) = m_0 \in M$ equation* (9.32) *has a weak solution that exists for all $t \in [0, T]$.*

Proof. By Theorem 9.17 we can construct a field of linear operators $A(t, m)$ that is C^1-smooth and such that $\alpha(t, m) = A(t, m)A^*(t, m)$. Consider equation (7.18) with these $a(t, m)$ and $A(t, m)$. Since its coefficients are C^1-smooth (i.e., locally Lipschitz continuous), for any initial condition $\xi(0) = m_0 \in M$, by Theorem 7.36 the equation has a strongly unique local strong solution that exists, by Theorem 7.43, on the entire interval $t \in [0, T]$. From Theorem 9.5 and Lemma 9.16 it follows that this solution satisfies (9.32). □

Now consider inclusions in mean derivatives on M. Let $\mathbf{a}(t, m)$ be a set-valued vector field on M, i.e., for every point $m \in M$ a set $\mathbf{a}(t, m) \subset T_m M$ is specified. Let also $\boldsymbol{\alpha}(t, m)$ be a set-valued symmetric positive semi-definite $(2, 0)$-tensor field on M (this means that for all t and m any tensor from the set $\boldsymbol{\alpha}(t, m)$ is symmetric and positive semi-definite). Consider the problem

$$\begin{cases} D\xi(t) \in \mathbf{a}(t, \xi(t)), \\ D_2\xi(t) \in \boldsymbol{\alpha}(t, \xi(t)). \end{cases} \tag{9.33}$$

Definition 9.26. We say that (9.33) has a *weak solution* on $[0, T]$ with initial condition $\xi(0) = m_0$ if there exists a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and a process $\xi(t)$ given on $(\Omega, \mathcal{F}, \mathsf{P})$, taking values in M, such that P-a.s. and for almost all t in $[0, T]$ the inclusions (9.33) are satisfied.

Denote by $\mathcal{A}_{\mathbf{a},\boldsymbol{\alpha}}(t, m)$ the set-valued second order vector field with images $\mathcal{A}_{\mathbf{a},\boldsymbol{\alpha}}(t, m) = \{\mathcal{A}_{a,\alpha}(t, m) \mid a \in \mathbf{a}(t, m), \alpha \in \boldsymbol{\alpha}(t, m)\}$ determined by formulae $\mathcal{H}\mathcal{A}_{a,\alpha}(t, m) = a(t, m)$ and $\mathcal{Q}\mathcal{A}_{a,\alpha}(t, m) = \alpha(t, m)$ as above (see formulae (2.45) and (2.44)).

Theorem 9.27 *Let $\boldsymbol{\alpha}(t, m)$ and $\mathbf{a}(t, m)$ be an upper semicontinuous set-valued symmetric positive semi-definite $(2, 0)$-tensor field and a vector field on M, respectively, with closed convex images. Let in addition for every compact $\mathbf{K} \subset M$ the sets $\mathbf{a}([0, T], \mathbf{K})$ and $\boldsymbol{\alpha}([0, T], \mathbf{K})$ be compact and assume* *that at every* (t, m) *the generator* $\mathcal{A}_{a,\alpha}$ *from a neighborhood* $\mathcal{V}$ *of the graph of* $\mathcal{A}_{\mathbf{a},\boldsymbol{\alpha}}(t, m)$ *satisfies the conditions of Theorem* 7.43 *with the same proper function* φ*. Then for any initial condition* $\xi(0) = m_0$ *there exists a weak solution of* (9.33)*, well-defined on the entire interval* $[0, T]$*.*

Proof. Introduce a sequence of positive numbers $\varepsilon_q \to 0$. By Theorem 4.11 for every ε_q there exists a single-valued continuous ε_q-approximation $a_q(t, m)$ of $\mathbf{a}(t, m)$ and a single valued continuous ε_q-approximation $\tilde{\alpha}_q(t, m)(\cdot, \cdot)$ of $\boldsymbol{\alpha}(t, m)$ such that the sequences of those approximations point-wise converge to a Borel measurable selection $a(t, m)$ of $\mathbf{a}(t, m)$ and $\alpha(t, m)$ of $\boldsymbol{\alpha}(t, m)$, respectively, as $q \to \infty$.

Note that the tensor fields $\tilde{\alpha}_q(t, m)(\cdot, \cdot)$ are symmetric and positive semi-definite. Introduce another sequence

$$\alpha_q(t, m)(\cdot, \cdot) = \tilde{\alpha}_q(t, m)(\cdot, \cdot) + \varepsilon_q g(m)(\cdot, \cdot)$$

where $g(m)(\cdot, \cdot)$ is the $(2, 0)$-metric tensor (Riemannian metric). It is evident that the tensors $\alpha_q(t, m)(\cdot, \cdot)$ are continuous, positive definite and symmetric and that the sequence $\alpha_q(t, m)(\cdot, \cdot)$ point-wise converges to $\alpha(t, m)(\cdot, \cdot)$ as $q \to \infty$. Since continuous fields can be approximated by smooth ones, without loss of generality we may suppose that all fields $a_q(t, m)$ and $\alpha_q(t, m)(\cdot, \cdot)$ are smooth. For simplicity denote by $\mathcal{A}_q(t, m)$ the generator $\mathcal{A}_{a_q,\alpha_q}(t, m)$.

From Theorem 9.17 it follows that there exist a sufficiently large integer K and a sequence of smooth fields of linear operators $A_q(t, m) : \mathbb{R}^K \to T_m M$ such that $\alpha_q(t, m) = A_q(t, m)A_q^*(t, m)$ for all q, t and m. Note that K depends only on the dimension of M since $\mathbb{R}^K$ is a space in which M with arbitrary Riemannian metric can be isometrically embedded by Nash's Theorem (Theorem 1.46).

Now introduce $\hat{a}_q(t, m)$ which, in a chart on M, has coordinates $\hat{a}_q^k = a^k - \frac{1}{2}\Gamma_{ij}^k \alpha_q^{ij}$ where Γ_{ij}^k are the Christoffel symbols of H and (α_q^{ij}) is the matrix of α_q. Note that $\mathcal{A}_q(t, m)$ is the field of generators for the Itô equation $(\hat{a}_q, A_q)$. By construction, for q large enough $(m, \mathcal{A}_q(t, m))$ belongs to $\mathcal{V}$. Then one can easily derive from the hypothesis that the Itô equations $(\hat{a}_q, A_q)$ satisfy the conditions of Lemma 7.57 and so the equations $(\hat{a}_q, A_q)$ have strong and strongly unique solutions ξ_q, well-defined on the entire interval $[0, T]$, and the set of corresponding measures $\{\mu_q\}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ is weakly compact. Hence we can choose a subsequence that weakly converges to some measure μ. For convenience we suppose that the sequence μ_q itself weakly converges to μ. Denote by $\xi(t)$ the coordinate process on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$. Recall that this means that for every $x(\cdot) \in C^0([0, T], M)$ we have $\xi(t, x(\cdot)) = x(t)$.

We show that $\xi(t)$ is the solution that we are looking for.

Notice that the "present" σ-algebra $\mathcal{N}_t^\xi$ of $\xi(t)$ is a σ-subalgebra of $\mathcal{P}_t^\xi$ and so for conditional expectations the equality

$$E(E(\cdot \mid \mathcal{P}_t) \mid \mathcal{N}_t) = E(\cdot \mid \mathcal{N}_t) \tag{9.34}$$

holds.

Consider the chart $\mathcal{U}$, the global chart in $\mathbb{R}^K$ and the symbols $\widehat{\Gamma}^k_{ij}$ introduced in the proof of Lemma 7.57. Define $\bar{a}_q(t,m,x)$, $\bar{A}_q(t,m,x)$ and $\bar{\alpha}_q(t,m,x)(\cdot,\cdot)$ by formula (7.33) where $\breve{a}$, $\breve{A}$ and $(\tilde{\alpha}^{ij})$ are replaced by $a(t,m)$, $A(t,m)$ and (α^{ij}), respectively. Define $\bar{a}(t,m)$ and $\bar{\alpha}(t,m)(\cdot,\cdot)$ by the same formulae via the replacements $a(t,m)$ and $\alpha(t,m)$, respectively.

As in the proof of Lemma 7.57, one can easily show that the equation

$$\begin{aligned} \mathrm{d}\bar{\xi}_q(t) = {} & \bar{a}_q\left(t,\bar{\xi}_q(t)\right)\mathrm{d}t \\ & -\frac{1}{2}\mathrm{tr}\left(\widehat{\boldsymbol{\Gamma}}_{\bar{\xi}_q(t)}\left(\bar{A}_q\left(t,\bar{\xi}_q(t)\right),\bar{A}_q\left(t,\bar{\xi}_q(t)\right)\right)\right)\mathrm{d}t \\ & +\bar{A}_q\left(t,\bar{\xi}_q(t)\right)\mathrm{d}w(t) \end{aligned} \tag{9.35}$$

(analogous to (7.34)) is well-defined on $\mathbb{R}^K$ and that in $\mathcal{U}$ it transforms into the system

$$\begin{cases} d\xi_q(t) = a_q(t,\xi_q(t))\mathrm{d}t - \frac{1}{2}\mathrm{tr}(\boldsymbol{\Gamma}_{\xi_q}(A_q,A_q))\mathrm{d}t + A_q(t,\xi_q(t))\mathrm{d}w(t), \\ d\bar{\xi}_q(t) = 0 \end{cases} \tag{9.36}$$

where $\mathrm{tr}\boldsymbol{\Gamma}(A_q,A_q) = \Gamma^k_{ij}\alpha^{ij}_q$, and so with probability 1 the solution of (9.35) with initial condition $\bar{\xi}_q(0) = m_0 \in M$ lies in M and coincides with the corresponding solution of $(\hat{a}_q, A_a)$ for all q. Thus the measures on the path space in $\mathbb{R}^K$ corresponding to the solutions are located on $C^0([0,T],M)$ and coincide there with μ_q for all q.

By construction the sequence $\bar{a}_q(t,x(\cdot)) = \bar{a}_q(t,x(t))$ point-wise converges to $\bar{a}(t,x(\cdot)) = \bar{a}(t,x(t))$. Hence it converges almost surely with respect to all $\lambda\times\mu_q$ where λ is the normalized Lebesgue measure on $[0,T]$.

The fact that

$$E\Bigg(\bar{\xi}\left((t+\Delta t)\wedge\theta_p^{x(\cdot)}\right) - x\left(t\wedge\theta_p^{\xi(\cdot)}\right) - \int_{t\wedge\theta_p^{x(\cdot)}}^{(t+\Delta t)\wedge\theta_p^{x(\cdot)}}\left[\bar{a}\left(s,\xi(\cdot)\right) - \frac{1}{2}\mathrm{tr}\left(\widehat{\Gamma}\left(\bar{\alpha}\right)\right)\right]\mathrm{d}s\,\Bigg|\,\mathcal{P}_t\Bigg) = 0$$

is proved by a simple modification of the proof of Lemma 8.47. Introduce compact W_p as in Lemma 7.57. Taking into account (9.34) the fact that $a(t,m)$ and $\alpha(t,m)$ are Borel measurable and that the above arguments are valid for every p and $\cup_p W_p = M$, and so every m belongs to W_p for large enough p, by transition to (9.36) one can easily derive from the last expression that the regression (8.5) takes the form $Y^o(t,m) = a(t,m) - \frac{1}{2}\mathrm{tr}(\boldsymbol{\Gamma}_m(\alpha(t,m))$. Since in the normal chart of every $m\in M$ with respect to H all Christoffel symbols $\Gamma^k_{ij}(m)$ are equal to zero, we obtain from Definition 8.2 that $D^{\mathsf{H}}\xi(t) = a(t,\xi(t))$.

The fact that

$$E\left(\left(\bar{\xi}\left((t+\Delta t)\wedge\theta_p^{x(\cdot)}\right) - \bar{\xi}\left(t\wedge\theta_p^{x(\cdot)}\right)\right)\right.$$
$$\otimes\left(\bar{\xi}\left((t+\Delta t)\wedge\theta_p^{x(\cdot)}\right) - \bar{\xi}\left(t\wedge\theta_p^{x(\cdot)}\right)\right)$$
$$\left. - \int\limits_{t\wedge\theta_p^{x(\cdot)}}^{(t+\Delta t)\wedge\theta_p^{x(\cdot)}} \bar{\alpha}\left(s,\bar{\xi}(\cdot)\right)\mathrm{d}s \right| \mathcal{P}_t\Bigg) = 0$$

is proved by a slight modification of the proof of Lemma 8.48. From the last relation one easily deduces that $D_2\xi(t) = \alpha(t,\xi(t))$. □

9.7 Stochastic Differential Inclusions in Terms of Infinitesimal Generators

In this Section we deal with a slight generalization of the notion of an infinitesimal generator which is well-defined for non-Markovian stochastic processes. For a process $\xi(t)$ with values in a manifold M (in particular, in $\mathbb{R}^n$) we introduce the generator as a field of second order semi-elliptic differential operators acting on the (sufficiently smooth) function f according to the rule

$$\mathcal{A}(t,m)f = \lim_{\Delta t\to+0} E\left(\left.\frac{f(\xi((t+\Delta t))\wedge\tau_m) - f(\xi(t\wedge\tau_m))}{\Delta t}\right|\xi(t)=m\right) \quad (9.37)$$

where τ_m is the Markov time that ξ first hits the boundary of a given chart containing m. The difference between (9.37) and Definition 6.34 is that here we use the regression (see Section 6.1.2) instead of the unconditional expectation. Note that if $\xi(t)$ is Markovian, both (9.37) and Definition 6.34 define the same object.

Clearly the generator defined by (9.37) is a second order tangent vector. As in Lemma 9.4 and Lemma 9.16 one can easily prove that if $\mathcal{A}$ is the generator of $\xi(t)$ for a given connection H, the formulae

$$D^{\mathsf{H}}\xi(t) = (\mathcal{H}\mathcal{A})(t,\xi(t)) \quad (9.38)$$

and

$$D_2\xi(t) = 2(\mathcal{Q}\mathcal{A})(t,\xi(t)) \quad (9.39)$$

hold where $\mathcal{Q}$ and $\mathcal{H}$ are defined by formulae (2.44) and (2.45), respectively.

Suppose that the field of second order tangent vectors $\mathfrak{A}(t,m)$ is set-valued, i.e., in every second order tangent space $\tau_m M$ to the manifold M a set $\mathfrak{A}(t,m)$

depending on $t \in [0, \infty)$ is given. We want to find a stochastic process $\xi(\cdot)$ such that for every t its generator $\mathcal{A}(t, m)$ a.s. satisfies the inclusion

$$\mathcal{A}(t, \xi(t)) \in \mathfrak{A}(t, \xi(t)). \tag{9.40}$$

Problems of this sort naturally arise if the process is described in terms of its generator.

If the set-valued field $\mathfrak{A}(t, m)$ has a continuous selector with some regularity properties, one can find the process having that selector as generator. This process is a solution of (9.40). For example, if $\mathfrak{A}(t, m)$ is lower semi-continuous and has convex closed values, by Michael's theorem it has a continuous selector. If the selector has a positive definite $(2, 0)$-tensor component, the above argument applies, proving the existence of solutions of (9.40).

If the selectors do not exist, the proof of solvability for (9.40) becomes much more complicated. We consider a rather general case of this sort, important for applications, where $\mathfrak{A}(t, m)$ is upper semi-continuous and not necessarily positive definite.

Theorem 9.28 *Let $\mathfrak{A}(t, m)$, $t \in [0, T]$, be an upper semi-continuous set-valued second order vector field on a manifold M with closed convex values such that:*

(i) *for every $t \in [0, T]$, $m \in M$, and for each $\mathcal{A} \in \mathfrak{A}(t, m)$, the $(2, 0)$-tensor $\mathcal{Q}_m\mathcal{A}$ is symmetric and positive semi-definite;*
(ii) *for every compact $\mathbf{K} \in M$ the set $\mathfrak{A}([0, T], \mathbf{K})$ is compact in τM;*
(iii) *there exist a proper function $\psi : M \to \mathbb{R}$, a constant $C > 0$ and a neighborhood $\mathcal{V}$ of the graph of $\mathfrak{A}$ in $[0, T] \times \tau(M)$ such that for every $(t, m, \mathcal{A}) \in \mathcal{V}$ the inequality $|\mathcal{A}\psi| < C$ holds.*

Then for every $m_0 \in M$ there exists a probability space and a stochastic process $\xi(t)$ with initial condition $\xi(0) = m_0$, well-defined for all $t \in [0, T]$, given on the probability space and taking values in M, such that, for its infinitesimal generator, inclusion (9.40) is a.s. satisfied.

Proof. In this proof we are working in the Banach manifold $C^0([0, T], M)$ equipped with the σ-algebra $\mathcal{F}$ generated by cylinder sets. By $\mathcal{P}_t$ we denote the σ-subalgebra of $\mathcal{F}$ generated by cylinder sets with bases over $[0, t] \subset [0, T]$.

Specify an arbitrary complete Riemannian metric $g(\cdot, \cdot)$ on M with chart components g_{ij}. This metric turns M into a metric space with respect to the corresponding Riemannian distance. Denote by H the Levi-Civitá connection of $g(\cdot, \cdot)$. Recall that $g_m(\cdot, \cdot)$ is an inner product in T_mM, $m \in M$. This inner product determines the inner products $\mathbf{g}_m(\cdot, \cdot)$ in the space of $(2, 0)$-tensors at $m \in M$ by the rule $\mathbf{g}_m((\alpha^{ij}), (\beta^{lk})) = g_{ik}g_{jl}\alpha^{ij}\beta^{kl}$. Define the inner products in τ_mM by the formula

$$\mathfrak{g}(\mathcal{A}_1, \mathcal{A}_2) = g(\mathcal{H}\mathcal{A}_1, \mathcal{H}\mathcal{A}_2) + \mathbf{g}(\mathcal{Q}\mathcal{A}_1, \mathcal{Q}\mathcal{A}_2).$$

Thus in all T_mM, τ_mM and the space of $(2,0)$-tensors over M the corresponding Euclidean norms smoothly depending on m are given.

Let $\varepsilon_q \to 0$ be a positive sequence. By Theorem 4.11 for every ε_q there exists a single-valued ε-approximation $\tilde{\mathcal{A}}_q(t,m)$ of $\mathfrak{A}(t,m)$ such that the sequence $\mathcal{A}_q(t,m)$ point-wise converges to a Borel measurable selector $\mathcal{A}(t,m)$ of $\mathfrak{A}(t,m)$ as $q \to \infty$.

Consider the sequences of vector fields $a_q(t,m) = \mathcal{H}\mathcal{A}_q(t,m)$ and of $(2,0)$-tensor fields $\tilde{\alpha}_q(t,m)(\cdot,\cdot) = 2\mathcal{Q}\mathcal{A}_q(t,m)$, respectively. It is evident that $a_q(t,m)$ point-wise converges to $a(t,m) = \mathcal{H}\mathcal{A}(t,m)$, $\tilde{\alpha}_q(t,m)$ point-wise converges to $\alpha(t,m) = 2\mathcal{Q}\mathcal{A}(t,m)$ as $q \to \infty$ and that both $a(t,m)$ and $\alpha(t,m)$ are Borel measurable.

Note that the tensor fields $\tilde{\alpha}_q(t,m)(\cdot,\cdot)$ are symmetric and positive semi-definite. Introduce another sequence

$$\alpha_q(t,m)(\cdot,\cdot) = \tilde{\alpha}_q(t,m)(\cdot,\cdot) + \varepsilon_q g(m)(\cdot,\cdot).$$

It is clear that the tensors $\alpha_q(t,m)(\cdot,\cdot)$ are continuous, positive definite and symmetric and that the sequence $\alpha_q(t,m)(\cdot,\cdot)$ point-wise converges to $\alpha(t,m)(\cdot,\cdot)$ as $q \to \infty$. Since continuous fields can be approximated by smooth fields, without loss of generality we may suppose that all fields $a_q(t,m)$ and $\alpha_q(t,m)(\cdot,\cdot)$ are smooth. Denote by $\mathcal{A}_q(t,m)$ the smooth second order tangent vector field corresponding to the pair $(a_q(t,m), \alpha_q(t,m))$. By construction $\mathcal{A}_q(t,m)$ is a 2ε-approximation of $\mathfrak{A}(t,m)$ and the sequence $\mathcal{A}_q(t,m)$ point-wise converges to $\mathcal{A}(t,m)$ as $q \to \infty$.

Note that the properties of $a_q(t,m)$ and $\alpha_q(t,m)$ are the same as in the proof of Theorem 9.27. Hence by imitating that proof we can show that there exists a stochastic process $\xi(t)$, defined for all $t \in [0,T]$, such that $D^{\mathsf{H}}\xi(t) = a(t,\xi(t)) \in (\mathcal{H}\mathfrak{A})(t,\xi(t))$ and $D_2\xi(t) = \alpha(t,\xi(t)) \in 2(\mathcal{Q}\mathfrak{A})(t,\xi(t))$. By construction this is the solution of (9.40) that we are looking for. $\square$

The inclusion (9.40) most often arises in applications in a linear space. For the case in hand, we prove an existence theorem of another sort, whose hypothesis is formulated in terms of estimates of Itô type.

Let $\mathfrak{A}(t,m)$ be a set-valued second order vector field in $\mathbb{R}^n$. Then we can consider the set-valued vector field $\mathbf{a}(t,m)$, the vector part of $\mathfrak{A}(t,m)$, and its tensor part, the set-valued symmetric $(2,0)$-tensor field $\boldsymbol{\alpha}(t,m)$ taking values in positive semi-definite tensors.

Theorem 9.29 *Let the set-valued vector field* $\mathbf{a}(t,x)$ *be upper semi-continuous, have closed convex values and satisfy the estimate*

$$\|\mathbf{a}(t,x)\| < K(1 + \|x\|) \tag{9.41}$$

for some $K > 0$.

Let the set-valued $(2,0)$*-tensor field* $\boldsymbol{\alpha}(t,x)$ *be upper semi-continuous, take closed convex values in symmetric positive semi-definite tensors and be such that for each* $\alpha(t,x) \in \boldsymbol{\alpha}(t,x)$ *the estimate*

$$\|\mathrm{tr}\alpha(t,x)\| < K(1+\|x\|)^2 \tag{9.42}$$

holds for some $K>0$.

Then for every $x_0 \in \mathbb{R}^n$ there exists a probability space and a stochastic process $\xi(t)$ on the probability space with initial condition $\xi(0)=x_0$, well-defined for all $t\in[0,T]$, taking values in $\mathbb{R}^n$, such that for its infinitesimal generator the inclusion (9.40) is a.s. satisfied.

Proof. Consider a sequence of positive numbers $\varepsilon_q \to 0$. As in the proof of Theorem 9.28 we can construct a sequence $\mathcal{A}_q(t,x)$ of smooth ε_q-approximations of $\mathfrak{A}(t,x)$ that point-wise converges to a Borel measurable selector $\mathcal{A}(t,x)$ of $\mathfrak{A}(t,x)$ as $q\to\infty$. As in the proof of Theorem 9.28, introduce a sequence of smooth vector fields $a_q(t,x)$ point-wise converging to a Borel measurable selector $a(t,x)$ of $\mathbf{a}(t,x)$, and the sequence $\tilde{\alpha}_q(t,x)(\cdot,\cdot)$ of smooth $(2,0)$-tensor fields point-wise converging to a Borel measurable selector $\alpha(t,x)$ of $\boldsymbol{\alpha}(t,x)$. Construct the sequence $\alpha_q(t,x)(\cdot,\cdot)=\tilde{\alpha}_q(t,x)(\cdot,\cdot)+\varepsilon_q I(\cdot,\cdot)$ where $I(\cdot,\cdot)$ is the $(2,0)$-tensor field with unit matrix at every $x\in\mathbb{R}^n$. Each $\alpha_q(t,x)(\cdot,\cdot)$ is smooth, symmetric and positive definite, then by Lemma 8.40 for every q there exists a smooth field of linear operators $A_q(t,x):\mathbb{R}^n\to\mathbb{R}^n$ such that $\alpha_q(t,x)=A_q(t,x)A_q^*(t,x)$ where $A^*(t,x)$ is transposed to $A(t,x)$. Note that $\mathrm{tr}\alpha_q(t,x)$ equals the sum of the squares of the elements of $A_q(t,x)$, i.e. it is the square of the Euclidean norm on $A_q(t,x)$ in the corresponding space of matrices. Thus from (9.42) it follows that

$$\|A_q(t,x)\| \le K(1+\|x\|) \tag{9.43}$$

for all q. Note that by construction all $a_q(t,x)$ satisfy the estimate of type (9.41) for all q.

Consider the Itô equations

$$\mathrm{d}\xi_q(t) = a_q(t,\xi_q(t))\mathrm{d}t + A_q(t,\xi_q(t))\mathrm{d}w(t). \tag{9.44}$$

Applying Theorem 6.26 and the fact that the coefficients are smooth, one can derive from the hypothesis and the above argument that every equation (9.44) has a unique solution $\xi_q(t)$ with initial condition $\xi_q(0)=x_0$, well-defined on the entire interval $[0,T]$. Denote by μ_q the measure on $(C^0([0,T],\mathbb{R}^n),\tilde{\mathcal{F}})$ corresponding to $\xi_q(\cdot)$ where $\tilde{\mathcal{F}}$ is the σ-algebra generated by cylinder sets. Taking into account (9.41) and (9.43) one derives from Lemma 6.28 that the set $\{\mu_q\}$ is weakly compact. Thus we can select a subsequence that weakly converges to some measure μ. Denote by $\xi(t)$ the coordinate process on the probability space $(C^0([0,T],\mathbb{R}^n),\mathcal{F},\mu)$. The fact that $\xi(t)$ satisfies (8.41) with the above $a(t,m)$ and $\alpha(t,m)$, where D denotes the forward mean derivative with respect to the Levi-Civitá connection of the Euclidean metric in $\mathbb{R}^n$, is proved by analogy with the proof of Theorem 8.46. Hence $\mathcal{A}(t,x)$ is the generator of $\xi(\cdot)$. Since, by construction, $\mathcal{A}(t,\xi(t))\in\mathfrak{A}(t,\xi(t))$ a.s., this means that $\xi(t)$ is the solution we are looking for. $\square$

Chapter 10
Stochastic Analysis on Groups of Diffeomorphisms

Everywhere in this chapter we use the notions and notation introduced in Chapter 5. We describe a stochastic differential equation of a special sort on those groups of diffeomorphisms that arise in the applications to viscous hydrodynamics described in Section 16.4 below (see, e.g., [100, 104, 113]). This class of equations is characterized by the fact that they involve finite-dimensional Wiener processes. It should be pointed out that a theory of stochastic differential equations on infinite-dimensional manifolds involving infinite-dimensional Wiener processes exists (see, e.g., [23, 35, 66]) and is used in viscous hydrodynamics (see, e.g., [39]). However, the description of this theory requires a complicated functional-analytic machinery that is not included in our exposition. For simplicity of presentation, we restrict ourselves to the finite-dimensional version of the theory since, in applications, the theories yield very similar results.

10.1 The General Case

Consider an n-dimensional compact Riemannian manifold M without boundary. Let $s > \frac{n}{2} + 1$, then the group $\mathcal{D}^s(M)$ of Sobolev H^s-diffeomorphisms of M is well-defined as are all the geometric objects on it as described in Chapter 5, in particular, the weak Riemannian metric (5.1), its Levi-Civitá connection with connector (5.2) and the corresponding exponential mapping $\overline{\exp}$ from Remark 5.8.

Using Nash's Theorem (Theorem 1.46) embed M isometrically into a Euclidean space $\mathbb{R}^N$ for some sufficiently large N. Construct the field of linear operators $\mathbf{A}(t,m) : \mathbb{R}^N \to T_mM$ introduced in Example 7.4 (see also Example 7.40). For simplicity we shall suppose that the $(1,1)$-tensor field $B(t,m)$ (see Example 7.4) is C^∞-smooth.

Let $X \in \mathbb{R}^N$. Applying to this vector all operators of the field $\mathbf{A}(t,m)$, we obtain the C^∞-vector field $\mathbf{A}(t,m)X$ on M, i.e., a vector in the tangent space

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_10,

$T_e\mathcal{D}^s(M)$. Thus the field $\mathbf{A}(t,m)$ generates the linear mapping $\bar{\mathbf{A}}(t) : \mathbb{R}^N \to T_e\mathcal{D}^s(M)$ that acts according to the rule $\bar{\mathbf{A}}(t)X = \mathbf{A}(t,m)X$. Applying this mapping by right translation to all points of the group, we obtain the right-invariant field of linear operators $\bar{\mathbf{A}}(t,\eta) : \mathbb{R}^N \to T_\eta\mathcal{D}^s(M)$ given by the formula $\bar{\mathbf{A}}(t,\eta)X = TR_\eta\bar{\mathbf{A}}(t)X$. In particular, this mapping can be applied to a Wiener process $w(t)$ in $\mathbb{R}^N$ and so it sends the Wiener process into all tangent spaces to $\mathcal{D}^s(M)$.

By construction the field $\bar{\mathbf{A}}(t,\eta)$ is right-invariant and, since $\mathbf{A}(t,m)$ is C^∞-smooth, from the results of Section 5.1 (see ω-lemma 5.2) it follows that $\bar{\mathbf{A}}(t,\eta)$ is also C^∞-smooth.

Let $a(t,m)$ be a vector field on M of Sobolev class H^{s+1}. Regard $a(t,m)$ as a tangent vector in the space $T_e\mathcal{D}^s(M)$. Denote by $\bar{a}(t,\eta)$ the right-invariant vector field on $\mathcal{D}^s(M)$ generated by this vector. By Theorem 5.4 the vector field $\bar{a}(t,\eta)$ is C^1-smooth.

Thus on $\mathcal{D}^s(M)$ we can consider the following Itô stochastic differential equation in Belopolskaya-Daletskii form

$$\mathrm{d}\bar{\xi}(t) = \overline{\exp}_{\bar{\xi}(t)}\left(\bar{a}(t,\xi(t))\mathrm{d}t + \bar{\mathbf{A}}(t,\xi(t))\mathrm{d}w(t)\right). \tag{10.1}$$

Theorem 10.1 *For every $g \in \mathcal{D}^s(M)$ equation (10.1) has a unique strong solution $\bar{\xi}_{0,g}(t)$ with initial condition $\bar{\xi}_{0,g}(0) = g$ that exists for all $t \in [0,+\infty)$.*

Proof. Define on $\mathcal{D}^s(M)$ the strong Riemannian metric (5.12). Let $\mathcal{W}$ be a neighborhood of e in $\mathcal{D}^s(M)$ that is covered by the mapping $\overline{\exp}_e$ according to Theorem 5.10. Introduce in $\mathcal{W}$ a normal chart of the Levi-Civitá connection of the weak Riemannian metric (5.1). The strong norm of the local connector $\boldsymbol{\Gamma}_\eta(\cdot,\cdot)$ for this connection in this chart, being the norm of a quadratic operator, is a continuous function of the point $\eta \in \mathcal{W}$ and at e this function equals zero since the connector itself equals zero. Hence, there exists an open set $\mathcal{U} \subset \mathcal{W}$ such that at every point of this set the above-mentioned norm is less than any *a priori* given number $C > 0$. Since $\mathcal{U}$ is open, it contains a ball $V_e(r)$ centered at e of some radius $r > 0$ with respect to the strong Riemannian distance on $\mathcal{D}^s(M)$ generated by the metric (5.12).

Now for every point $\eta \in \mathcal{D}^s(M)$ determine the chart in its neighborhood as the right shift by η of the normal chart $\mathcal{W}$. The existence of a local solution $\bar{\xi}_{0,g}(t)$ (up to the Markov time that $V_g(r)$ first hits the boundary in the chart at g, constructed above) follows from that fact that the coefficients of equation (10.1) are smooth.

Since the metric (5.12) is right-invariant, the atlas, constructed by the above-mentioned right shifts, is by construction a uniform Riemannian metric for the strong metric (5.12). On the balls $V_\eta(r)$ in the charts of this atlas, the norm of the local connector $\boldsymbol{\Gamma}$ is bounded by C since the Levi-Civitá connection of the weak metric (5.1) is right-invariant (Theorem 5.6). Clearly the right-invariant Itô vector field $(\bar{a},\bar{\mathbf{A}})$ on $\mathcal{D}^s_\mu(\mathcal{T}^n)$ is bounded with respect to the metric (5.12). Thus equation (10.1) satisfies the hypothesis of Theorem 7.41. □

Denote $\bar{\xi}_e(t)$ by $\xi(t)$. From the fact that equation (10.1) is right-invariant it follows that $\bar{\xi}_g(t) = \xi(t) \circ g$. It is not hard to see that $\xi(t)$ is a general solution of the stochastic differential equation (7.23) on M. In other words, for every $m \in M$ the process $\xi(t)(m)$ is a solution of (7.23) on M with the initial condition m at $t = 0$.

Note that $\xi(t)$ exists by Theorem 7.36 since M is compact and the field $(a, \mathbf{A})$ is smooth. However, the infinite-dimensional equation (10.1) gives us additional information about the solution $\xi(t)$: for $\omega \in \Omega$ the mapping $\xi(t, \omega) : M \to M$ is a.s. an H^s-diffeomorphism.

An analogous construction can also be realized for equations in Stratonovich form. On $\mathcal{D}^s(M)$ consider the equation

$$\mathrm{d}\bar{\eta}(t) = \bar{a}(t, \eta(t))\mathrm{d}t + \bar{\mathbf{A}}(t, \eta(t)) \circ \mathrm{d}w(t) \tag{10.2}$$

with the same $\bar{a}(t, \eta)$, $\bar{\mathbf{A}}(t, \eta)$ and $w(t)$ as in equation (10.1).

Theorem 10.2 *For every $g \in \mathcal{D}^s(M)$, equation (10.2) has a unique strong solution $\bar{\eta}_{0,g}(t)$ with initial condition $\bar{\eta}_{0,g}(0) = g$ that exists for all $t \in [0, +\infty)$.*

The proof of Theorem 10.2 follows the same argument as that in Theorem 10.1 with the following modification. In the normal chart $\mathcal{W}$ at e of the Levi-Civitá connection of the weak metric (5.1) the strong norm of the operator $\frac{1}{2}\mathrm{tr}\bar{\mathbf{A}}'(t, \eta)(\bar{\mathbf{A}}(t, \eta)(\cdot), \cdot)$ is a continuous function of $\eta \in \mathcal{W}$. Let $\|\frac{1}{2}\mathrm{tr}\bar{\mathbf{A}}'(t, e)(\bar{\mathbf{A}}(t, e)(\cdot), \cdot)\| < C$. As $\mathcal{U} \subset \mathcal{W}$ we take an open set such that at each of its points $\|\frac{1}{2}\mathrm{tr}\bar{\mathbf{A}}'(t, \eta)(\bar{\mathbf{A}}(t, \eta)(\cdot), \cdot)\|$ is less than the constant C. The rest of the argument follows without modification.

The solution of (10.2), starting at time 0 from e, is the general solution of the finite-dimensional equation (7.3).

The construction of the Itô vector field $(\bar{a}, \bar{\mathbf{A}})$ and of equation (10.2) on $\mathcal{D}^s(M)$, generated by the Itô field $(a, \mathbf{A})$ and equation (7.3) on the finite-dimensional manifold M, is a version of a general construction due to K.D. Elworthy (see [66]).

10.2 The Case of a Flat Torus

In the case where the Riemannian manifold M is a flat n-dimensional torus $\mathcal{T}^n$ (see Section 5.2) we can consider stochastic differential equations with a special diffusion term. Equations with such a term will be used below in some models of mathematical physics. It is also possible to introduce this type of equation on the group $\mathcal{D}^s_\mu(\mathcal{T}^n)$ of diffeomorphisms preserving the volume.

In this section we deal only with Itô equations in Belopolskaya-Daletskii form. The transition to equations in Stratonovich form can be achieved by analogy with that in the previous section.

Let $\mathcal{T}^n$ be an n-dimesional flat torus, i.e., the Riemannian metric $\langle\cdot,\cdot\rangle$ on $\mathcal{T}^n$ is inherited from $\mathbb{R}^n$ under factorization with respect to the integral lattice.

Consider the group $\mathcal{D}^s(\mathcal{T}^n)$ of H^s-diffeomorphisms of the torus. Using the flat metric $\langle\cdot,\cdot\rangle$ on $\mathcal{T}^n$, introduce on $\mathcal{D}^s(\mathcal{T}^n)$ the weak Riemannian metric $(\cdot,\cdot)$ by formula (5.1), the corresponding Levi-Civitá connection (5.2), its covariant derivative $\bar{\nabla}$, the exponential mapping $\overline{\exp}$ and the other geometric objects as they were described in Sections 5.1 and 5.2.

Consider the mapping A from Definition 5.16(ii). It is clear that A is jointly C^∞-smooth in all variables and that for any given $X\in\mathbb{R}^n$ the vector field $\mathsf{A}(X)$ on $\mathcal{T}^n$ is constant (its coordinates with respect to the basis $\frac{\partial}{\partial q^1},\ldots,\frac{\partial}{\partial q^n}$, mentioned in Remark 1.40, are constant). In particular this means that this vector field is C^∞-smooth and divergence-free (see Section 5.2).

Thus A generates the mapping $\bar{\mathsf{A}}:\mathcal{D}^s(\mathcal{T}^n)\times\mathbb{R}^n\to T\mathcal{D}^s(\mathcal{T}^n)$ where $\bar{\mathsf{A}}_e:\mathbb{R}^n\to T_e\mathcal{D}^s_\mu(\mathcal{T}^n)\subset T_e\mathcal{D}^s(\mathcal{T}^n)$ is given by the expression $\bar{\mathsf{A}}_e(X)=\mathsf{A}(X)$ and for $g\in\mathcal{D}^s(\mathcal{T}^n)$ the mapping $\bar{\mathsf{A}}_g:\mathbb{R}^n\to T_g\mathcal{D}^s(\mathcal{T}^n)$ is constructed from A by the right shift: $\bar{\mathsf{A}}_g(X)=TR_g\bar{\mathsf{A}}_e(X)=(\mathsf{A}\circ g)(X)$. Since A is C^∞-smooth, from Theorem 5.4 it follows that $\bar{\mathsf{A}}$ is also jointly C^∞-smooth in $X\in\mathbb{R}^n$ and $g\in\mathcal{D}^s(\mathcal{T}^n)$, i.e., in particular, for all $X\in\mathbb{R}^n$ the right-invariant vector field $\bar{A}(X)$ on $\mathcal{D}^s(\mathcal{T}^n)$ is C^∞-smooth.

Let $\sigma>0$ be a real number and $a(t,m)$ be an H^α-vector field on $\mathcal{T}^n$ where $t\in[0,l]$ and $\alpha>s$ is an integer. Denote by $\bar{a}(t,g)$ the right invariant vector field on $\mathcal{D}^s(\mathcal{T}^n)$ generated by $a(t,m)$ as a vector of $T_e\mathcal{D}^s(\mathcal{T}^n)$. For the above-mentioned α the smoothness of the field $\bar{a}(t,g)$ is no coarser than C^1. The pair $(\bar{a},\bar{\mathsf{A}})$ is an Itô vector field on $\mathcal{D}^s(\mathcal{T}^n)$.

Consider a Wiener process $w(t)$ in $\mathbb{R}^n$ given on a probability space $(\Omega,\mathcal{F},\mathsf{P})$. According to the description of stochastic differential equations on Hilbert manifolds given at the end of Section 7.3, the Belopolskaya-Daletskii approach is well-posed in this setting and we can consider on $\mathcal{D}^s(\mathcal{T}^n)$ the stochastic differential equation of type (7.18) in the form:

$$\mathrm{d}\bar{\xi}(t)=\overline{\exp}_{\bar{\xi}(t)}\left(\bar{a}(t,\bar{\xi}(t))\mathrm{d}t+\sigma\bar{\mathsf{A}}(t,\bar{\xi}(t))\mathrm{d}w(t)\right). \tag{10.3}$$

Theorem 10.3 *For every $g\in\mathcal{D}^s(\mathcal{T}^n)$ there exists a unique strong solution $\bar{\xi}_g(t)$ of (10.3) with initial condition $\bar{\xi}_g(0)=g$ which is well-defined for all $t\in[0,l]$, where $l>0$ is an arbitrary a priori specified real number.*

Proof. Introduce the normal chart in a neghbourhood of e in $\mathcal{D}^s(\mathcal{T}^n)$ by applying $\overline{\exp}$. Note that at every point of this chart the local connector (5.2) equals zero since the connection is generated by the Euclidean connection on the torus $\mathcal{T}^n$. Take a strong right-invariant Riemannian metric on $\mathcal{D}^s(\mathcal{T}^n)$ (say, generated by (5.12)). Since the above-mentioned normal chart is an open set, there exists a real number $r>0$ such that the ball $V_e(r)$ (with radius r with respect to the strong Riemannian distance on $\mathcal{D}^s(\mathcal{T}^n)$) centered at e is contained in this neighborhood. Then in a neighborhood of each point $g\in\mathcal{D}^s(\mathcal{T}^n)$ we determine the chart by applying the right shift R_g to the ball $V_e(r)$. In such a manner we obtain an atlas on $\mathcal{D}^s(\mathcal{T}^n)$ that is uniformly Riemannian for the strong metric, and in the charts of this atlas the local connectors of connection (5.2) equal zero since the connection is right-invariant (see Theorem 5.6). By construction, the right-invariant Itô vector field $(\bar{a}, \sigma\bar{\mathsf{A}})$ is uniformly bounded with respect to the strong Riemannian metric. Thus Theorem 7.41 can be applied to (10.3). $\square$

Denote $\bar{\xi}_e(t)$ by $\xi(t)$. From the fact that (10.3) is right-invariant, it follows that $\bar{\xi}_g(t) = \xi(t) \circ g$. It is not hard to see that $\xi(t)$ is the general solution of the following stochastic differential equation on $\mathcal{T}^n$:

$$\mathrm{d}\xi(t) = \exp_{\xi(t)}(a(t, \xi(t))\mathrm{d}t + \sigma\mathsf{A}\mathrm{d}w(t)). \tag{10.4}$$

In other words, for every $m \in \mathcal{T}^n$ the process $\xi(t)(m)$ is a solution of (10.4) on $\mathcal{T}^n$ with initial condition m at $t = 0$.

Note that $\xi(t)$ exists by Theorem 7.36, since $\mathcal{T}^n$ is compact and the field $(\bar{a}, \bar{A})$ is smooth. However, as in the previous Section, the infinite-dimensional equation (10.3) gives us additional information on the solution $\xi(t)$: for $\omega \in \Omega$ the mapping $\xi(t, \omega) : \mathcal{T}^n \to \mathcal{T}^n$ is a.s. an H^s-diffeomorphism of the torus $\mathcal{T}^n$.

Remark 10.4. In equations (10.3) and (10.4) we used the general notation of Itô equations in Belopolskaya-Daletskii form. Nevertheless it should be pointed out that since the connection on $\mathcal{D}^s(\mathcal{T}^n)$ is generated by the flat connection on the torus, the corresponding exponential map is like that on a linear space. So, without loss of generality we can employ the same notation that is used for Itô equations in linear spaces. However, this is not the case for equations on $\mathcal{D}^s_\mu(\mathcal{T}^n)$ which we consider below. This is why it is sometimes more convenient to to consider the latter equations as equations on $\mathcal{D}^s(\mathcal{T}^n)$ subjected to the constraint $\bar{\beta}$ obtained by right translations of $T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$ at all points of $\mathcal{D}^s(\mathcal{T}^n)$, as introduced below in Section 16.2.

Theorem 10.5 *Let $\sigma > 0$ be a real constant.*

(i) *For every $\omega \in \Omega$ and $t \in [0, l]$ the vector field $\mathsf{A}(\sigma w(t, \omega))$ on $\mathcal{T}^n$, where $w(t)$ is a Wiener process in $\mathbb{R}^n$, is divergence-free, i.e., $\mathsf{A}(\sigma w(t))$ is a stochastic process in $T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$.*

(ii) *For every $\omega \in \Omega$ and $t \in [0, l]$ the mapping $W^{(\sigma)}_\omega(t) = \overline{\exp}_e \mathsf{A}(\sigma w(t, \omega)) : \mathcal{T}^n \to \mathcal{T}^n$ is a volume-preserving H^s-diffeomorphism of $\mathcal{T}^n$, i.e., $W^{(\sigma)}(t)$ is a stochastic process in $\mathcal{D}^s_\mu(\mathcal{T}^n)$.*

Proof. Let $\omega \in \Omega$ and $t \in [0, l]$. As mentioned above, the vector field $\mathsf{A}(\sigma w(t, \omega))$ on $\mathcal{T}^n$ is constant (has constant coordinates with respect to the basis $\frac{\partial}{\partial q^1}, \dots, \frac{\partial}{\partial q^n}$) and so it is C^∞ and divergence-free. The mapping $W^{(\sigma)}_\omega(t)$ sends $m \in \mathcal{T}^n$ to $\exp_m \mathsf{A}(w(t, \omega))$ where $\exp_m : T_m\mathcal{T}^n \to \mathcal{T}^n$ is the exponential mapping of the flat torus $\mathcal{T}^n$. This means that every m is sent to $m + \mathsf{A}(\sigma w(t, \omega)) = m + \sigma w(t, \omega)$ modulo factorization with respect to the integral lattice, i.e., all points of the torus $\mathcal{T}^n$ under the mapping $W_\omega^{(\sigma)}(t)$ carry out the same shift as that generated by the shift of the space $\mathbb{R}^n$ by $\sigma w(t,\omega)$. $W_\omega^{(\sigma)}(t)$ is clearly volume-preserving. □

For ease of reference we highlight the formula for the action of the diffeomorphism $W_\omega^{(\sigma)}(t)$ on $\mathcal{T}^n$:

$$W_\omega^{(\sigma)}(t)(m) = m + \sigma w(t,\omega). \tag{10.5}$$

Let $a(t,m)$ be a divergence-free vector field on $\mathcal{T}^n$. In the rest of this section we assume $a(t,m)$ to be H^α-smooth where $\alpha > s$. Thus the right invariant vector field $\bar{a}$ on the group $\mathcal{D}_\mu^s(\mathcal{T}^n)$ of volume-preserving H^s-diffeomorphisms of $\mathcal{T}^n$ is at least C^1-smooth (see Section 5.1). From Theorem 10.5(i) it follows that the mapping $\bar{\mathsf{A}}$ can be restricted to $\mathcal{D}_\mu^s(\mathcal{T}^n)$ and so we can consider the restriction $\bar{\mathsf{A}} : \mathcal{D}_\mu^s(\mathcal{T}^n) \times \mathbb{R}^n \to T\mathcal{D}_\mu^s(\mathcal{T}^n)$. In particular, $\bar{\mathsf{A}}$ sends the Wiener process $w(t)$ on $\mathbb{R}^n$ to the tangent spaces to $\mathcal{D}_\mu^s(\mathcal{T}^n)$.

Thus we can consider on $\mathcal{D}_\mu^s(\mathcal{T}^n)$ the following stochastic differential equation

$$\mathrm{d}\tilde{\xi}(t) = \widetilde{\exp}_{\tilde{\xi}(t)}\left(\bar{a}(t,\tilde{\xi}(t))\mathrm{d}t + \sigma\bar{\mathsf{A}}(\tilde{\xi}(t))\mathrm{d}w(t)\right), \tag{10.6}$$

where $\widetilde{\exp}$ is the exponential mapping of the spray $\mathcal{S}$ on $\mathcal{D}_\mu^s(\mathcal{T}^n)$ (see Section 5.1) and $\sigma > 0$.

Theorem 10.6 *For every $g \in \mathcal{D}_\mu^s(\mathcal{T}^n)$ there exists a unique strong solution $\tilde{\xi}_g(t)$ of (10.6) with initial condition $\tilde{\xi}_g(0) = g$ which is well-defined for all $t \in [0,l]$.*

Proof. Introduce the strong Riemannian metric (5.11) on $\mathcal{D}_\mu^s(\mathcal{T}^n)$. Recall that this metric is right-invariant on $\mathcal{D}_\mu^s(\mathcal{T}^n)$. Let $\mathcal{W}$ be a neighborhood of e in $\mathcal{D}_\mu^s(\mathcal{T}^n)$ which is covered by the mapping $\widetilde{\exp}_e$ according to Theorem 5.10. Consider the normal chart at e in $\mathcal{W}$. The strong norm of the local connector $\boldsymbol{\Gamma}_\eta(\cdot,\cdot)$, being a quadratic operator, is a continuous function of $\eta \in \mathcal{W}$ in this chart. Moreover, at e we have $\boldsymbol{\Gamma}_e(\cdot,\cdot) = 0$. Hence there exists an open set $\mathcal{U} \subset \mathcal{W}$ such that at each of its points the above-mentioned norm is less than an *a priori* given constant $C > 0$. Since $\mathcal{U}$ is open, it contains the ball $V_e(r)$ centered at e and having some radius $r > 0$ with respect to the strong Riemannian distance on $\mathcal{D}_\mu^s(\mathcal{T}^n)$ generated by the metric (5.11).

Now we define a chart at a neighborhood of each point $g \in \mathcal{D}_\mu^s(\mathcal{T}^n)$ by the right shift of the normal chart $\mathcal{W}$ at g. The atlas constructed in this way is obviously uniformly Riemannian for the strong metric (5.11). Moreover, on the balls $V_g(r)$ in the charts of this atlas the norm of the local connector $\boldsymbol{\Gamma}$ of the Levi-Civitá connection of the metric (5.1) on $\mathcal{D}_\mu^s(\mathcal{T}^n)$ is bounded by C since the connection is right-invariant (Theorem 5.6). The right-invariant Itô vector field $(\bar{a},\bar{\mathsf{A}})$ on $\mathcal{D}_\mu^s(\mathcal{T}^n)$ is bounded with respect to the right-invariant metric (5.11). Thus, equation (10.6) satisfies the conditions of Theorem 7.41. □

Part III
Applications to Mathematical Physics

Chapter 11
Newtonian Mechanics

11.1 A Geometric Language for Newtonian Mechanics

Let M be a finite-dimensional manifold. Recall that on the manifold TM there is a vertical distribution V (a sub-bundle of the second tangent bundle TTM) whose fibers consist of vectors tangent to the fibers of TM. The vectors belonging to V are said to be *vertical* (see Section 2.1).

Definition 11.1. A 1-form on TM is said to be *horizontal* if it vanishes on vertical vectors.

The horizontal 1-form $\alpha\big(t,(m,X)\big) \in T^*_{(m,X)}TM$ gives rise to a 1-form $\tilde{\alpha}(t,m,X) \in T^*_m M$ depending on $X \in T_m M$ via the formula

$$\tilde{\alpha}(t,m,X)(Y) = \alpha\left(t,(m,X)\right)\left(T\pi^{-1}Y\big|_{T_{(m,X)}TM}\right). \tag{11.1}$$

We define a *mechanical system* to be the following collection of data:

1) a *configuration space*, i.e., a smooth manifold M;
2) *kinetic energy*, i.e., a smooth function $\mathcal{K}$ on the tangent bundle TM;
3) a *force field*, i.e., a horizontal 1-form $\alpha(t,m,X)$ on TM, which in general is time-dependent.

The tangent space TM is called the *coordinate-velocity phase space* and the cotangent space T^*M is called the *coordinate-momentum phase space* (see [134]). In what follows, we consider only mechanical systems with quadratic kinetic energy, i.e., $\mathcal{K}(X) = \langle X, X\rangle/2$, where $X \in TM$ and $\langle\cdot,\cdot\rangle$ is a Riemannian metric on M.

Since $\alpha\big(t,(m,X)\big)$ is horizontal, the form $\tilde{\alpha}(t,m,X)$ is well-defined. Formula (11.1) gives a one-to-one correspondence between horizontal 1-forms $\alpha\big(t,(m,X)\big)$ on TM and 1-forms $\tilde{\alpha}(t,m,X)$ on M. The latter will also be called a force field.

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_11,

Let $\bar{\alpha}(t, m, X)$ be the vector field on M (depending on $X \in T_m M$) physically equivalent to the 1-form $\tilde{\alpha}(t, m, X)$ with respect to the Riemannian metric $\langle \cdot, \cdot \rangle$, which gives rise to the kinetic energy of the system. In other words, $\langle \bar{\alpha}(t, m, X), Y \rangle = \tilde{\alpha}(t, m, X)(Y)$ for any $Y \in T_m M$.

Definition 11.2. A vector field $a(t, m, X)$ where $t \in \mathbb{R}$, $m \in M$ and $X \in T_m M$ such that $\pi a(t, m, X) = \pi(m, X) = m$ is called a *vector force field.*

One can easily see that $\bar{\alpha}(t, m, X)$, as defined above, is an example of a vector force field.

Remark 11.3. In the remainder of the book, with the exception of Chapter 16, we shall consider only mechanical systems with a finite-dimensional configuration space. The force fields are usually introduced as vector fields and the passage to 1-forms is left to the reader as a simple exercise.

The motion of a mechanical system is governed by Newton's second law, i.e., the equation:

$$\frac{\mathrm{D}}{\mathrm{d}t}\,\dot{m}(t) = \bar{\alpha}\left(t, m(t), \dot{m}(t)\right), \tag{11.2}$$

where $\frac{\mathrm{D}}{\mathrm{d}t}$ is the covariant derivative of the Levi-Civitá connection of the metric $\langle \cdot, \cdot \rangle$ (see Section 2.6). Recall also that a curve $m(t)$ is a solution of (11.2) if and only if the curve $(m(t), \dot{m}(t))$ in TM is an integral curve of the vector field

$$\mathcal{Z} + \bar{\alpha}(t, m, X)^l, \tag{11.3}$$

where $\mathcal{Z}$ is the spray of the Levi-Civitá connection of $\langle \cdot, \cdot \rangle$ and $\bar{\alpha}(t, m, X)^l$ is the vertical lift of $\bar{\alpha}(t, m, X)$ to the space $V_{(m,X)} \subset T_{(m,X)}TM$.

A curve $m(t)$ on M is called a *trajectory* of the mechanical system if it is a solution of (11.2). For any initial conditions $m(0) = m_0$ and $\dot{m}(0) = X_0 \in T_{m_0}M$, there exists a trajectory $m(t)$ on a sufficiently small interval of time provided that, for example, $\tilde{\alpha}(t, m, X)$ satisfies the Carathéodory condition [74]. To see this, observe that, since Z is smooth, the field (11.3) on TM satisfies the Carathéodory condition and the local existence of a solution follows from the classical existence theorem for ordinary differential equations. Below we investigate the existence of trajectories for more general force fields (for instance, in the case of a discontinuous force field, this question is related to the passage from (11.2) to differential inclusions). Assume, for example, that $\bar{\alpha}(t, m, X)$ is locally Lipschitz. Then the trajectory is unique for any given initial conditions.

The existence of trajectories on $(-\infty, \infty)$ may be analyzed using the methods developed in Section 3.1. Let us point out that if the Riemannian metric which gives rise to the kinetic energy is complete, then, by the Hopf-Rinow theorem, the trajectories of the system with zero force field (i.e., geodesics) are defined on $(-\infty, \infty)$. From the physical point of view, this means that a free particle in the system does not escape to infinity in finite time. Note that the assumptions that the Riemannian metric is complete and that a solution of the Cauchy problem for trajectories is unique are easy to justify for systems arising in physics. Here, however, we usually do not require in advance that the solution should be unique.

The cotangent vector p physically equivalent to the velocity $\dot{m}$ with respect to the metric $\langle\cdot,\cdot\rangle$ (i.e., where $p = \langle\dot{m},\cdot\rangle$) is called the *momentum* of the system. Consider the operator $\mathcal{L}\colon TM \to T^*M$ sending a vector $X \in T_mM$ to the covector $\langle X,\cdot\rangle \in T^*_mM$. The operator $\mathcal{L}$ is called the *inertia operator* (or the *inertia tensor*). In particular, $p = \mathcal{L}(\dot{m})$. In terms of the inertia operator, the kinetic energy of the system is given by the relation $\mathcal{K}(X) = \big(\mathcal{L}(X)\big)(X)/2$.

The manifold M is often equipped with a canonical Riemannian metric $(\cdot,\cdot)$ (e.g., the standard inner product in $\mathbb{R}^3$), which is not related to the metric $\langle\cdot,\cdot\rangle$ giving rise to $\mathcal{K}$. In this case, we may identify vectors and covectors by means of $(\cdot,\cdot)$. Then $\langle X,Y\rangle = (\mathfrak{L}X,Y)$ where $\mathfrak{L}$ is a $(1,1)$-tensor field of self-adjoint linear operators in tangent spaces. $\mathfrak{L}$ is also called the inertia tensor; in fact this tensor is physically equivalent to $\mathcal{L}$. Sometimes it is convenient to define the kinetic energy not by the metric $\langle\cdot,\cdot\rangle$ but by the self-adjoint operator $\mathfrak{L}\colon TM \to TM$ once the canonical metric $(\cdot,\cdot)$ is given.

11.2 Mechanical Systems on Lie Groups

Consider a mechanical system which has a Lie group as the configuration space with the kinetic energy given by a right- or left-invariant metric. We shall call such a system a mechanical system on the group.

The best known example of a mechanical system on a group is the system describing the motion of a rigid body which rotates about a stationary point in $\mathbb{R}^3$. It is easy to see that the configuration space of this system, i.e., the set of all possible positions of the rigid body with a stationary point, is the special orthogonal group $SO(3)$. Choosing an orthonormal basis in $\mathbb{R}^3$, we may identify $SO(3)$ with the group of all orthogonal matrices with unit determinant. Recall that the Lie algebra $\mathfrak{so}(3)$ of $SO(3)$ at $e = \mathrm{id}$ is identified with $\mathbb{R}^3$ (see Section 1.2) so that after this identification, the Killing form $(A,B) = \mathrm{tr}\,(A\circ B)$ and the commutator $[A,B] = A\circ B - B\circ A$ on $\mathfrak{so}(3)$ become the standard inner and vector products on $\mathbb{R}^3$, respectively. The Riemannian metric obtained from the Killing form by left translations turns out to be bi-invariant. This metric is used as the canonical metric to express the kinetic energy via the inertia tensor.

Let $g(t)$ be a trajectory on $SO(3)$ corresponding to the motion of the rigid body. The velocity vector $\dot{g}(t) \in T_{g(t)}SO(3)$ can be translated to the algebra $\mathfrak{so}(3) = T_eSO(3)$ (and so to $\mathbb{R}^3$) in two different (and non-commuting) ways: by left translation and by right translation (see Section 1.2). The vectors

$$\omega_c(t) = TL^{-1}_{g(t)}\dot{g}(t) \quad \text{and} \quad \omega_s(t) = T\mathbb{R}^{-1}_{g(t)}\dot{g}(t)$$

belong to $\mathfrak{so}(3)$, which we have identified with the space $\mathbb{R}^3$ where the body moves (see above). The vector ω_c is the angular velocity with respect to the body coordinates (i.e., the coordinate system "attached" to the body and moving along with it) and the vector ω_s is the angular velocity in the space coordinates (i.e., with respect to a coordinate system fixed in space).

Let $\mathfrak{L}\colon \mathbb{R}^3 \to \mathbb{R}^3$ be the inertia operator (tensor) of the rigid body. Recall that $\mathfrak{L}$ depends on the shape of the body and on the distribution of its mass. The operator $\mathfrak{L}$ is self-adjoint with respect to the standard inner product on $\mathbb{R}^3$. Let us define a new inner product $\langle A, B\rangle_e = (\mathfrak{L}A, B)$ on $\mathfrak{so}(3)$. This inner product gives rise to the left-invariant Riemannian metric $\langle \cdot, \cdot \rangle$, which determines the kinetic energy by the general formula $\mathcal{K}(X) = \langle X, X\rangle/2$. Note that the choice of the left-invariant metric is motivated by physics: the kinetic energy depends on the angular velocity with respect to the body's coordinates, but not on the position of the body in space.

Remark 11.4. In Chapter 16 we consider a mechanical system on the infinite-dimensional group of diffeomorphisms. This system describes the hydrodynamics of an ideal incompressible fluid and the energy is given by a right-invariant (weak) Riemannian metric.

The classical Euler equation of motion of a rigid body (with a stationary point) describes the time variation of the angular velocities ω_c and ω_s, as well as the angular momenta $M_c = \mathcal{L}(\omega_c)$ and $M_s = \mathcal{L}(\omega_s)$. Thus, the Euler equation is an equation in the algebra $\mathfrak{so}(3)$ or in the dual space $\mathfrak{so}(3)^*$. Throughout this book, we adopt the following terminology of [47]. The equations in $SO(3)$ are said to describe the motion in the *material coordinates* or in the *Lagrangian representation*. The equations for ω_s in $\mathfrak{so}(3)$ are said to be with respect to the *space coordinates* or in the *Eulerian representation*, whereas the equations for ω_c are in the *body coordinates* or in the *convective representation*. Similar terminology is used for the corresponding angular momenta equations.

Analogous terms are used to describe a mechanical system on an arbitrary Lie group G. Thus, Newton's equation on G is called the equation of motion in the material coordinates (the Lagrangian representation). The corresponding equation on the Lie algebra is called the Euler equation. If the equation is obtained by right translations, then it is said to be with respect to the space coordinates (the Eulerian representation), while for left translations it is with respect to the body coordinates (the convective representation), etc. This also applies to the so-called Eulerian and Lagrangian specifications in hydrodynamics.

Mechanical systems on groups, especially the Euler equation on Lie algebras, have been studied very intensively in the last few decades and there is a substantial literature devoted to the subject. A detailed introduction to this field may be found in Appendix 2 of [3].

11.3 Conservative Mechanical Systems

Definition 11.5. A mechanical system is called *conservative* if its force field is independent of velocities (and is usually autonomous) and is equal to $-\mathrm{d}\mathcal{U}$, the negative differential of a certain real function $\mathcal{U}$ on M which we call the *potential energy*.

Thus the vector force field of a conservative mechanical system is $-\mathrm{grad}\,\mathcal{U}$ where the gradient is calculated with respect to a Riemannian metric defining the kinetic energy, i.e., $\langle X, \mathrm{grad}\,\mathcal{U}\rangle = \mathrm{d}\mathcal{U}(X)$ for any vector field X on M.

For a conservative system Newton's law (11.2) is transformed into

$$\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) = -\mathrm{grad}\,\mathcal{U}. \tag{11.4}$$

Recall that $\langle X, \mathrm{grad}\,\mathcal{U}\rangle = X\mathcal{U}$ for any vector field X on M, where $X\mathcal{U}$ denotes the derivative of the function $\mathcal{U}$ along the vector field X. In particular, for a trajectory $m(t)$ (i.e. a solution of (11.4)) this means that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{U}(m(t)) = \langle \dot{m}, \mathrm{grad}\,\mathcal{U}\rangle. \tag{11.5}$$

Recall also that the gradient of $\mathcal{U}$ depends both on $\mathcal{U}$ and on the Riemannian metric $\langle\cdot,\cdot\rangle$, i.e., on the form of the kinetic energy.

Now we can derive the principal laws of dynamics as theorems.

There is a notable difference in the definitions of kinetic and potential energy: the kinetic energy is a function on TM (indeed, it assigns the value $\frac{1}{2}\langle X, X\rangle$ to any vector $X \in TM$) while the potential energy is a function on M. This is not convenient because, for example, we cannot sum these functions. To overcome this obstacle we extend $\mathcal{U}$ to TM by the formula

$$\mathcal{U} \circ \pi : TM \to \mathbb{R} \tag{11.6}$$

where $\pi : TM \to M$ is the natural projection.

Definition 11.6. The function $E = \mathcal{K} + \mathcal{U} \circ \pi : TM \to R$ is called the *total energy*.

Theorem 11.7 (Conservation of energy law). *The total energy is constant along any trajectory $m(t)$ of a conservative mechanical system.*

Proof. Since the Levi-Civitá connection is Riemannian, from (2.30) it follows that $\frac{\mathrm{d}}{\mathrm{d}t}\langle \dot{m}(t), \dot{m}(t)\rangle = \langle \frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t), \dot{m}(t)\rangle + \langle \dot{m}(t), \frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t)\rangle = 2\langle \frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t), \dot{m}(t)\rangle$. Taking into account formula (11.5) we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}E(m(t)) &= \frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{K}(\dot{m}(t)) + \mathcal{U}(m(t))) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\langle \dot{m}(t), \dot{m}(t)\rangle + \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{U}(m(t)) \\ &= \left\langle \frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t), \dot{m}(t)\right\rangle + \langle \mathrm{grad}\,\mathcal{U}, \dot{m}(t)\rangle .\end{aligned}$$

Since by (11.4) $\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) = -\mathrm{grad}\,\mathcal{U}$, we get $\frac{\mathrm{d}}{\mathrm{d}t}E(m(t)) = 0$. $\square$

11.4 Hamilton's Principle of Least Action

The *Lagrangian* of a natural mechanical system is the function $L = \mathcal{K} - \mathcal{U} \circ \pi$.

Let us recall some notions and formulae from the calculus of variations. A function whose argument is a curve (for the sake of definiteness we assume all curves to be parametrized by t in the interval $t \in [0, T]$), not a point, is (in the calculus of variations) called a *functional*.

The *functional of action* of the above-mentioned Lagrangian L is

$$A(m(t)) \mid_0^T = \int_0^T L(m(t))\mathrm{d}t = \int\limits_0^T \left(\frac{1}{2} \langle \dot{m}(t), \dot{m}(t) \rangle - \mathcal{U}(m(t)) \right) \mathrm{d}t. \quad (11.7)$$

We shall also deal with two more functionals: the *functional of length* $s(m(t)) = \int_0^T \|\dot{m}(t)\| \mathrm{d}t = \int_0^T \sqrt{\langle \dot{m}(t), \dot{m}(t) \rangle} \mathrm{d}t$ (cf. formula (1.19)) and the so-called *functional of action* A_0 (which is a particular case of (11.7) with $\mathcal{U} = 0$)

$$A_0(m(t)) = \int_0^T \|\dot{m}(t)\|^2 \mathrm{d}t = \int_0^T \langle \dot{m}(t), \dot{m}(t) \rangle \mathrm{d}t. \quad (11.8)$$

For a curve $m(t)$ its *variation* $m(t, s)$ is a smooth mapping from $[0, T] \times (-\varepsilon, \varepsilon)$ into M such that $m(t, 0) = m(t)$. The term "variation" refers to the fact that by varying s, we vary the curve $m(t)$, i.e. we obtain curves near to it. As in Section 2.6 (see Lemma 2.57), we consider the vector fields $\frac{\partial}{\partial t}$ and $\frac{\partial}{\partial s}$ on the image of a variation. In particular, the field of vectors $\frac{\partial}{\partial s}$ along the curve $m(t) = m(t, 0)$ will be denoted by $W(t)$ and called the *field of variation*.

Definition 11.8. A variation $m(t, s)$ is called a *variation with fixed end-points* if $m(0, s) = \text{const}$ and $m(T, s) = \text{const}$ for all s.

In particular, for a variation of a curve with fixed ends we have

$$\frac{\partial}{\partial s}(0, s) = 0 \qquad \frac{\partial}{\partial s}(T, s) = 0. \quad (11.9)$$

The *variation of a functional* with respect to the variation of a curve is an analog of the derivative of a function in ordinary analysis. It is denoted by $\frac{\delta}{\mathrm{d}s}$. To find the variation of a functional one should substitute the variation of the curve into the functional and differentiate the obtained expression with respect to s at $s = 0$.

The *extremal* (an analog of the extremum) is a curve at which the variation of the functional with respect to each variation of the curve equals zero. An *extremal with fixed end-points* of a functional is a curve at which the variation of the functional with respect to each variation of the curve with fixed ends equals zero.

Lemma 11.9 *The variation of the functional of action A_0 at a smooth curve $m(t)$ with respect to the variation $m(t, s)$ with fixed ends is equal to*

$$\frac{\delta}{\mathrm{d}s} A_0 = -2 \int_0^T \left\langle \frac{\mathrm{D}}{\mathrm{d}t} \dot{m}(t), W(t) \right\rangle \mathrm{d}t, \tag{11.10}$$

where $W(t)$ is the field of variation $m(t,s)$.

We refer the reader to [181] for a proof of Lemma 11.9. Note that in [181] the variation is found in the class of piecewise smooth curves, from which (11.10) follows as a particular case. This equation suffices for our purposes, since the trajectories of mechanical systems are smooth.

Theorem 11.10 *The geodesics of the Levi-Civitá connection, and only these geodesics, are extremals with fixed ends of the functional of action A_0.*

Proof. Let $m(t)$ be a geodesic, i.e., $\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) = 0$. Then expression (11.10) equals zero for any $W(t)$. Thus the variation of A_0 at $m(t)$ equals zero with respect to each variation of this curve with fixed ends. By definition this means that $m(t)$ is an extremal of A_0 with fixed ends.

Let $m(t)$ be an extremal with fixed ends, i.e., expression (11.10) equals zero for each $W(t)$. This can happen only if the second factor under the integral is zero. Hence $\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) = 0$ for all t, i.e., $m(t)$ is a geodesic. □

Theorem 11.11 *The geodesics of the Levi-Civitá connection, and only these geodesics, are extremals with fixed ends of the functional of length.*

Proof. First we derive an analog of formula (11.10):

$$\begin{aligned}
\frac{\partial}{\partial s}\int_0^T \left\| \frac{\partial}{\partial t} \right\| \mathrm{d}t_{|s=0} &= \frac{\partial}{\partial s}\int_0^T \sqrt{\left\langle \frac{\partial}{\partial t}, \frac{\partial}{\partial t} \right\rangle} \mathrm{d}t_{|s=0} \\
&= \int_0^T \frac{\partial}{\partial s}\sqrt{\left\langle \frac{\partial}{\partial t}, \frac{\partial}{\partial t} \right\rangle} \mathrm{d}t_{|s=0} \\
&= \frac{1}{2}\int_0^T \frac{\frac{\partial}{\partial s}\left\langle \frac{\partial}{\partial t}, \frac{\partial}{\partial t} \right\rangle}{\sqrt{\left\langle \frac{\partial}{\partial t}, \frac{\partial}{\partial t} \right\rangle}} \mathrm{d}t_{|s=0}.
\end{aligned}$$

Recall that the length (unlike the action A_0) does not depend on the parametrization. Hence we can change the parametrization t in the variation $m(t,s)$ so that $\|\frac{\partial}{\partial t}(t,s)\| = \sqrt{\langle \frac{\partial}{\partial t}, \frac{\partial}{\partial t} \rangle}$ does not depend on t, i.e., is equal to some constant $C(s)$. Then $\frac{\partial}{\partial s}\int_0^T \|\frac{\partial}{\partial t}\| \mathrm{d}t_{|s=0} = \frac{1}{2C(s)}\int_0^T \frac{\partial}{\partial s}\langle \frac{\partial}{\partial t}, \frac{\partial}{\partial t}\rangle \mathrm{d}t_{|s=0}$, which differs from (11.10) only by the presence of the factor $\frac{1}{2C(s)}$ before the integral. Thus the remaining argument coincides word-for-word with that of the proof of Theorem 11.10. □

So, unlike the straight lines of Euclidean spaces, the geodesics of the Levi-Civitá connection realize only local minimums of length. In addition we should mention that sometimes several geodesics can connect two points and yet it can happen that there exists no geodesic with length equal to the Riemannian distance between those points. Conditions under which such geodesics exist are given by the Hopf-Rinow Theorem (Theorem 3.68).

The next statement is often regarded as the central law of mechanics (see, e.g., [171]).

Theorem 11.12 (Hamilton's principle of least action). *The trajectories of a natural mechanical system, and only these trajectories, are the extremals of the functional of action* (11.7) *with fixed ends.*

Proof. The variation of A with respect to the variation $m(t,s)$ of the curve $m(t)$, $t \in [0,T]$, is calculated as follows:

$$\frac{\partial}{\partial s}\int_0^T \left(\frac{1}{2}\langle m_t(t,s), m_t(t,s)\rangle - \mathcal{U}(m(t,s))\right) \mathrm{d}t_{|s=0}$$
$$= \frac{\partial}{\partial s}\int_0^T \left(\frac{1}{2}\langle m_t(t,s), m_t(t,s)\rangle \mathrm{d}t_{|w=0} - \frac{\partial}{\partial w}\int_0^T \mathcal{U}(m(t,s))\right) \mathrm{d}t_{|s=0}.$$

The first integral on the right-hand side here is one half of the variation of the functional A_0 and so we can substitute formula (11.9) divided by 2 for its value. For the second integral we have:

$$\frac{\partial}{\partial s}\int_0^T \mathcal{U}(m(t,s)))\mathrm{d}t_{|s=0} = \int_0^T \frac{\partial}{\partial s}\mathcal{U}(m(t,s)))\mathrm{d}t_{|s=0},$$

which is equal to $\int_0^t \langle \operatorname{grad}\mathcal{U}, \frac{\partial}{\partial s}m(t,s)\rangle \mathrm{d}t_{|s=0}$ by the definition of gradient. Recall (see above) that $\frac{\partial}{\partial w}m(t,s)_{|s=0}$ is the field of variation denoted by $W(t)$. Thus, the formula of variation with fixed ends of A takes the form

$$-\int_0^T \left\langle \frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t), W(t)\right\rangle \mathrm{d}t - \int_0^t \langle \operatorname{grad}\mathcal{U}, W(t)\rangle \,\mathrm{d}t. \tag{11.11}$$

Now let $m(t)$ be a trajectory, i.e., $\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) = -\operatorname{grad}\mathcal{U}$ by (11.4). In this case, by (11.11) the variation of A is equal to zero for any W, i.e., $m(t)$ is an extremal. If (11.11) equals zero for every $W(t)$, $\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) = -\operatorname{grad}\mathcal{U}$ and so the extremal is a trajectory. □

For conservative systems we also have the principle of least action in Maupertuis form. According to this principle, trajectories with total energy h are simply the geodesics of a new metric (the so-called *Jacobi metric*):

$$(h - \mathcal{U})\langle \cdot, \cdot \rangle. \tag{11.12}$$

Note that the Jacobi metric (11.12) is Riemannian only when $h > \mathcal{U}$ everywhere on M. Hence, in this case, the analysis of trajectories reduces to the description of geodesics on M with the metric given by (11.12). If there are points where the value of $\mathcal{U}$ is greater than or equal to h, then the trajectories lie in the domain $\Omega_h = \{m \in M \mid \mathcal{U}(m) < h\}$ called the *domain of possible motion* (with boundary). The geometry of this domain has been studied by many authors in connection with certain problems of mechanics. Particularly active research in this area was initiated by the works of Kozlov. In [169] the reader may find a detailed review of this subject.

Remark 11.13. A more complex case is where one has the so-called gyroscopic force in addition to the potential one. A *gyroscopic force field* has the form $\tilde{\alpha}_1(m, X) = \omega(X, \cdot)$, where ω is a 2-form on M. (Recall that the 1-form $\tilde{\alpha}_1(m, X)$ takes the value $\omega_m(X, Y)$ on $Y \in T_m M$.) Usually one assumes that the form ω is closed or even exact. An example of a gyroscopic force is the action of a magnetic field on a charge. Substantial progress in the study of such systems was achieved by Novikov [193]. Another approach was developed in [169].

11.5 Noether's Theorem

Recall the following:

Definition 11.14. A non-constant smooth function on a manifold, on which a differential equation is given, is called the *first integral* or an *integral of motion* of the equation if it is constant along every solution of the equation. Sometimes one says that the equation has a *conservation law*.

Examples of first integrals (conservation laws) are the Hamiltonian for a Hamiltonian system (see Remark 3.69) and the total energy for a conservative mechanical system (see Theorem 11.7).

In this section, for the case of conservative mechanical systems, we prove one of the most important theorems in contemporary natural science, Noether's theorem, which gives a standard method of finding conservation laws.

Definition 11.15. A *one-parameter diffeomorphism group* g_s is a mapping $g : \mathbb{R} \times M \to M$, jointly smooth in all variables, for which the following conditions are satisfied:

(i) for each $s \in \mathbb{R}$ the mapping $g_s : M \to M$ is a diffeomorphism (i.e., g_s is one-to-one and smooth and g_s^{-1} is also smooth);
(ii) for each $m \in M$ and all $s, u \in \mathbb{R}$ the relation $g_{s+u}(m) = g_s(g_u(m))$ holds (this relation is usually written as $g_{s+u} = g_s \circ g_u$).

From (ii) above it is clear that $g_0(m) = m$ for each $m \in M$ and that $g_s^{-1} = g_{-s}$.

Denote by $g(m)$ the set of points obtained from m by the action of g_s for all $s \in R$; such a set is called a *trajectory* or *orbit* of the one-parameter diffeomorphism group g_s.

For $m \in M$ consider the vector $Y_m = \frac{\mathrm{d}}{\mathrm{d}s} g_s(m)|_{s=0}$. Assigning, for each $m \in M$, the vector Y_m to the point m we obtain a smooth vector field Y on M that is called the *generator* of the group g_s. Note that the orbits of the group, and only the orbits, are integral curves of its generator Y.

Recall that on a compact manifold M every smooth autonomous vector field is a generator of a one-parameter diffeomorphism group that is the flow of the vector field. In the general case of non-compact manifolds, a smooth vector field is a generator of a one-parameter diffeomorphism group if the flow is complete.

Definition 11.16. A group g_s is said to *preserve* a non-constant function $f : M \to R$ if f is constant along the orbits of g_s.

It is clear that g_s preserves a smooth function f if and only if for every $m \in M$ the relation $\frac{\mathrm{d}}{\mathrm{d}s} f(g_s(m))|_{s=0} = Yf|_{g=0} = 0$ holds where Y is the generator of g_s.

Let g_s be a one-parameter diffeomorphism group on M. Then it is easy to see that Tg_s is a one-parameter diffeomorphism group on the tangent bundle TM where Tg_s is the tangent mapping to g_s. Let $L = \mathcal{K} - \mathcal{U} \circ \pi : TM \to R$ be the Lagrangian of a conservative mechanical system.

Definition 11.17. A one-parameter diffeomorphism group g_s is said to *preserve the Lagrangian* L if Tg_s preserves L according to Definition 11.16.

Theorem 11.18 (E. Noether) *Let on the configuration space M of a conservative mechanical system with Lagrangian $L = \mathcal{K} - \mathcal{U} \circ \pi$ there be a one-parameter diffeomorphism group g_s that preserves the Lagrangian L. Then this system has a conservation law.*

Proof. Consider a trajectory $m(t)$ of the mechanical system for $t \in [0, l]$. Recall that $m(t)$ satisfies Newton's second law (11.4). Applying the diffeomorphisms in g_s, for $s \in (-\varepsilon, \varepsilon)$, to $m(t)$, we obtain the set $N = g_s m(t)$ in M as the image of the mapping $m : [0, l] \times (-\varepsilon, \varepsilon) \to M$, $m(t, s) = g_s m(t)$. Consider the vector fields $\frac{\partial}{\partial t} = \frac{\partial}{\partial t} m(t, s)$ and $\frac{\partial}{\partial s} = \frac{\partial}{\partial s} m(t, s)$ on N. It is clear that $\frac{\partial}{\partial s}$ coincides with the generator Y of the group g_s.

Lemma 11.19 *For any given $s_* \in (-\varepsilon, \varepsilon)$ the curve $m(t, s_*)$ is a trajectory of the mechanical system.*

Proof. [of Lemma 11.19] By Hamilton's principle of least action (Theorem 11.12) the trajectory $m(t)$ is an extremal with fixed ends of the action functional with Lagrangian L. Since g_s preserves L, the values of the action functional on the curves under the action of g_s are preserved. Thus $m(t, s_*) = g_{s_*} \gamma(t)$ is also an extremal, i.e., it is a trajectory of the mechanical system. $\square$

By the hypothesis, g_s preserves the Lagrangian L, i.e., for all t the value $L(m(t,s)) = \mathcal{K}(\frac{\partial}{\partial t}) - \mathcal{U}(m(t,s))$ is constant in s, i.e.,

$$\frac{\partial}{\partial s}L(m(t,s)) = \frac{\partial}{\partial s}\mathcal{K}\left(\frac{\partial}{\partial t}\right) - \frac{\partial}{\partial s}\mathcal{U}(m(t,s)) = 0.$$

Taking into account the definition of $\mathcal{K}$, formula (2.29) defining the notion of a Riemannian connection and Lemma 2.57 on the second covariant derivative (recall that they are both valid for the Levi-Civitá connection), we obtain that $\frac{\partial}{\partial s}\mathcal{K}(\frac{\partial}{\partial t}) = \frac{\partial}{\partial s}\frac{1}{2}\langle\frac{\partial}{\partial t}, \frac{\partial}{\partial s}\rangle = \langle\nabla_{\frac{\partial}{\partial s}}\frac{\partial}{\partial t}, \frac{\partial}{\partial t}\rangle = \langle\nabla_{\frac{\partial}{\partial t}}\frac{\partial}{\partial s}, \frac{\partial}{\partial t}\rangle$.

Applying consequently formula (11.5), Newton's second law (11.4) and Remark 2.27 we obtain

$$\begin{aligned}\frac{\partial}{\partial s}\mathcal{U}(m(t,s)) &= \mathrm{d}\mathcal{U}\left(\frac{\partial}{\partial s}\right) = \left\langle\frac{\partial}{\partial s}, \operatorname{grad}\mathcal{U}\right\rangle = -\left\langle\frac{\partial}{\partial s}, \frac{\mathrm{D}}{\mathrm{d}t}\frac{\partial}{\partial t}\right\rangle \\ &= -\left\langle\frac{\partial}{\partial s}, \nabla_{\frac{\partial}{\partial t}}\frac{\partial}{\partial t}\right\rangle.\end{aligned}$$

Thus

$$0 = \frac{\partial}{\partial s}L(m(t,s)) = \left\langle\nabla_{\frac{\partial}{\partial t}}\frac{\partial}{\partial s}, \frac{\partial}{\partial t}\right\rangle + \left\langle\frac{\partial}{\partial s}, \nabla_{\frac{\partial}{\partial t}}\frac{\partial}{\partial t}\right\rangle.$$

By formula (2.29) this expression equals $\frac{\partial}{\partial t}\langle\frac{\partial}{\partial s}, \frac{\partial}{\partial t}\rangle$. Hence, $\frac{\partial}{\partial t}\langle\frac{\partial}{\partial s}, \frac{\partial}{\partial t}\rangle = 0$, i.e., $\langle\frac{\partial}{\partial s}, \frac{\partial}{\partial t}\rangle$ is constant in t.

We have shown that along every trajectory $m(t)$ of the mechanical system the product $\langle Y, \dot{m}\rangle$ is constant where $\dot{m}$ is the velocity vector of the trajectory and Y is the generator of the one-parameter diffeomorphism group g_s. $\square$

As the first application of Noether's theorem we prove the momentum conservation law for some special systems.This law is not as universal as the conservation of energy law since it is not valid for all conservative systems.

Consider the motion of two material particles in $\mathbb{R}^3$. Let x be the vector in $\mathbb{R}^3$ corresponding to the first particle and y the vector corresponding to the second particle. It is clear that the system's state at a given time is described by the pair x, y of vectors in $\mathbb{R}^3$, i.e., the configuration space is $\mathbb{R}^6 = \mathbb{R}^3 \times \mathbb{R}^3$. We allow the particles to be located at the same point $\mathbb{R}^3$ (otherwise the configuration space would have the form $(\mathbb{R}^3 \times \mathbb{R}^3)\setminus\triangle$ where $\triangle$ is the diagonal set $\{(x,y) \in \mathbb{R}^3 \times \mathbb{R}^3 | x = y\}$).

The kinetic energy here is the sum of the kinetic energies of both particles. It depends only on the velocities and does not depend on the positions of the particles. This situation is natural for Euclidean spaces, where it is possible to translate the vectors along the space.

Suppose that the force depends only on the distance between the particles and does not depend on the positions of the particles (this situation occurs, for example, in Newton's law of gravitation and in Coulomb's law in electrostatics). Denote the force by $\bar{\alpha}(t, x^1 - y^1, x^2 - y^2, x^3 - y^3)$. It is clear that the Euclidean distance in $\mathbb{R}^3$ is a function of the differences of the same coordinates.

Suppose in addition that the force field $\bar{\alpha}(t, x^1 - y^1, x^2 - y^2, x^3 - y^3) = -\operatorname{grad} \mathcal{U}(x^1 - y^1, x^2 - y^2, x^3 - y^3)$ where $\mathcal{U}$ is a smooth real-valued function on $\mathbb{R}^3$.

Consider on the configuration space $\mathbb{R}^6$ of this system the one-parameter diffeomorphism group of the form

$$g_s(x^1, x^2, x^3, y^1, y^2, y^3) = (x^1 + s, x^2, x^3, y^1 + s, y^2, y^3).$$

Diffeomorphisms of this group do not change the differences $x^1 - y^1$, $x^2 - y^2$, $x^3 - y^3$. For this reason, and since $\mathcal{U}$ and $\mathcal{K}$ do not depend on the location of a point in configuration space, this group preserves the Lagrangian. The generator takes the form $\frac{\partial}{\partial x^1} + \frac{\partial}{\partial y^1}$, and by Noether's theorem the system has a conservation law of the form $\langle \frac{\partial}{\partial x^1} + \frac{\partial}{\partial y^1}, \dot{m} \rangle = \mathrm{m}_1 \dot{x}^1 + \mathrm{m}_2 \dot{y}^1$. Notice that $\mathrm{m}_1 \dot{x}^1 + \mathrm{m}_2 \dot{y}^1$ is the first coordinate of the complete momentum of the system that by definition (see [3]) equals $\mathrm{m}_1 \dot{x} + \mathrm{m}_2 \dot{y}$ where m_1 and m_2 are the masses of the particles. By analogy, introducing the one-parameter diffeomorphism groups by adding t to the other coordinates of particles, we obtain that the other coordinates of the complete momentum are also integrals of motion. This means that the complete momentum of the system is preserved.

There is another example where a one-parameter diffeomorphism group preserves the Lagrangian. This is a mechanical system on a group such that the Riemannian metric defining the kinetic energy is left-invariant and the force equals zero. For the system of a rigid body with stationary point (see Section 11.2) the conservation law that arises here, by Noether's theorem, is known as the *conservation law of angular momentum*. In Section 16.3 below we present the proof of an analogous fact on an infinite-dimensional group of diffeomorphisms that has a hydrodynamical interpretation, as well as a certain infinite-dimensional version of Noether's theorem.

Remark 11.20. Taking into account Noether's theorem, it is possible to derive the energy conservation law from the fact that both the potential and kinetic energies of a conservative mechanical system do not depend on time, i.e., they are preserved under shifts in t of the form $t + s$.

11.6 Geometric Mechanics with Linear Constraints

This section is an introduction to the modern geometric approach to mechanics with constraints, which goes back to the paper [71] by Faddeev and Vershik. (See also [224, 225, 226] and the bibliography therein.) Here we focus on systems with quadratic kinetic energy, as in previous sections, and linear constraints only. Note that the papers mentioned above are devoted to Lagrangian mechanics with more general Lagrangians and possibly non-linear constraints.

11.6.1 The notion of a linear mechanical constraint

Consider a mechanical system, in the sense of Section 11.1, on a configuration space M.

Definition 11.21. A *linear constraint* in the system is a smooth distribution (i.e., a sub-bundle of the tangent bundle) $\boldsymbol{\beta}$ on M in the sense of Definition 1.41.

In what follows we call a linear constraint simply a *constraint.*

Definition 11.22. A tangent vector is called *admissible* if it lies in the distribution $\boldsymbol{\beta}$. A curve in M is *admissible* if all its tangent vectors are admissible.

A constraint $\boldsymbol{\beta}$ imposes a restriction on the motion of the system: all its trajectories must be admissible.

Let $P\colon TM \to \boldsymbol{\beta}$ be the operator of orthogonal projection (with respect to the Riemannian metric on M that determines the kinetic energy) of tangent spaces onto their subspaces $\boldsymbol{\beta}$, i.e., we have $P_m\colon T_mM \to \boldsymbol{\beta}_m$ for every $m \in M$. Let us define the *reduced covariant derivative* $\bar{\nabla} a$ on admissible vectors by the formula $\bar{\nabla}_X Y = P\nabla_X Y$, where ∇ is the covariant derivative of the Levi-Civitá connection. Denote by $\frac{\bar{\mathrm{D}}}{\mathrm{d}t} = P\frac{\mathrm{D}}{\mathrm{d}t}$ the reduced covariant derivative along a curve. Let $\bar{\alpha}(t, m, X)$ be the vector force field of the mechanical system. The equation of motion of the mechanical system with the constraint $\boldsymbol{\beta}$ is the following analog of Newton's equation:

$$\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\,\dot{m}(t) = P\bar{\alpha}(t, m, \dot{m}). \tag{11.13}$$

In the same way as for (11.2) and (3.17), one may show that a curve $m(t)$ is a solution of (11.13) if and only if its derivative $\dot{m}(t)$, regarded as a curve in the total space of the bundle $\boldsymbol{\beta}$, is an integral curve of the vector field

$$Y = TP(Z) + \big(P\bar{\alpha}(t, m, \dot{m})\big)^l. \tag{11.14}$$

It is not hard to see that $T\pi Y(m, X) = X \in \boldsymbol{\beta}_m$ and $Y \in T_{(m,X)}\boldsymbol{\beta}$.

If the distribution $\boldsymbol{\beta}$ is involutory (i.e., integrable, by Frobenius' Theorem 1.44), then the constraint is said to be *holonomic*, and (11.13) turns into Newton's equation (11.2) on the integral manifolds of the distribution. Thus, a system with a holonomic constraint reduces to one with no constraint on a manifold of lower dimension.

When the distribution $\boldsymbol{\beta}$ fails to be involutory (i.e., it is not integrable), the constraint is called *non-holonomic*. In this case, some extra effort is needed to study the mechanical system.

Definition 11.23. A constraint $\boldsymbol{\beta}$ is *totally non-holonomic* if the Lie brackets of the admissible vector fields generate the entire Lie algebra of vector fields on M.

Remark 11.24. One may introduce the notion of the degree of non-holonomity [225], which we do not consider here. Note also that linear constraints are admissible and ideal in the sense of [224].

11.6.2 Reduced connections

Consider the orthonormal frame bundle $\bar{\pi}\colon \mathbf{O}^{\boldsymbol{\beta}}(M) \to M$ of $\boldsymbol{\beta}$ (i.e., $b \in \mathbf{O}^{\boldsymbol{\beta}}_m$ is an orthonormal frame in $\boldsymbol{\beta}_m$). It is clear that $\mathbf{O}^{\boldsymbol{\beta}}(M)$ is a principal bundle with structure group $\mathbf{O}(k)$, $k = \dim \boldsymbol{\beta}_m$.

Theorem 11.25 *The reduced covariant derivative $\bar{\nabla}$ has all four properties of the regular covariant derivative described in Theorem* 2.24.

Proof. Since the operator P is linear on the fibers of TM, only the fourth property deserves a proof. For admissible X, Y and a smooth function f, we have

$$\bar{\nabla}_X(fY) = P\nabla_X(fY) = P\big(f\nabla_X Y + (Xf)Y\big) = f\bar{\nabla}_X Y + (Xf)Y,$$

because $PY = Y$. □

Thus, $\bar{\nabla}$ is the covariant derivative on admissible vectors and, in particular, it gives rise to the parallel translation of admissible vectors along admissible curves. The definition of such a parallel translation is quite similar to the standard one. Since P is orthogonal, it is clear that the parallel translation preserves the inner product on fibers of $\boldsymbol{\beta}$ and (2.29) holds for $\bar{\nabla}$. Therefore, on the fibers of $\boldsymbol{\beta}$, the parallel translation of orthonormal frames is defined along admissible curves. Consider now the sub-bundle $\bar{H}$ of $T\mathbf{O}^{\boldsymbol{\beta}}(M)$, which is defined as follows. The fiber of $\bar{H}$ over a point $(m, b) \in \mathbf{O}^{\boldsymbol{\beta}}(M)$ is formed by "infinitesimal" parallel translations of the frame b. It is easy to check that the sub-bundle is invariant with respect to the right action of $\mathbf{O}(k)$ and the fibers of $\bar{H}$ have zero intersection with the vertical subspaces $\bar{V}_{(m,b)}$. Thus, $\bar{H}$ can be thought of as an analog of a connection.

Definition 11.26. The sub-bundle $\bar{H}$ is called the *reduced connection.*

Remark 11.27. If the constraint is holonomic, the reduced connection is the Levi-Civitá connection on the integral manifolds with respect to the induced Riemannian metric. (Compare the construction of the reduced connection with that of the connection on the adapted frame bundle [161].)

Theorem 11.28 *Let $X_1,\ X_2, \dots, X_k$ be orthonormal admissible vector fields in a chart $\mathfrak{U}$ (i.e., at every $m \in \mathfrak{U}$ the vector fields $X_1, \dots, X_k$ form an orthonormal basis in $\boldsymbol{\beta}_m$). Then $\bar{\nabla}_{X_i} X_j = \overset{\circ}{\Gamma}{}^l_{ij} X_l$, where $\overset{\circ}{\Gamma}{}^l_{ij}$ are the tetrad Christoffel symbols (see Remarks* 2.37 *and* 2.56*) taken for an orthonormal frame in $\mathfrak{U}$ which contains $X_1, \dots, X_k$ as a subframe.*

The result follows from Remark 2.56.

Corollary 11.29 *The reduced connection of a non-holonomic constraint depends on the Riemannian metric on the entire manifold M rather than only on the restriction of the metric to $\boldsymbol{\beta}$.*

Remark 11.30. A variety of open problems concerning reduced connections, as well as their additional properties, is discussed in, e.g., [224]. Here we only point out that the torsion tensor of a reduced connection cannot be defined in the standard way. The reason is that even though the Christoffel symbols of the reduced connection are symmetric by definition, the difference

$$\bar{\nabla}_X Y - \bar{\nabla}_Y X - [X,Y] = P\big([X,Y]\big) - [X,Y]$$

is zero only if the distribution is involutory.

11.6.3 Length minimizing and least constrained non-holonomic geodesics

Let $\boldsymbol{\beta}$ be a non-holonomic constraint on M.

Definition 11.31. An admissible curve $m(t)$ in M is called a *least constrained non-holonomic geodesic* if it satisfies the equation

$$\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\,\dot{m}(t) = 0.$$

It is clear that the least constrained geodesics are, in fact, trajectories of constrained mechanical systems with a zero force field. Definition 11.31 is analogous to the standard definition of geodesics of a connection.

Definition 11.32. An admissible curve $m(t)$ is a *length minimizing non-holonomic geodesic* if it is an extremum of the action functional

$$\int_0^T \big\langle \dot{m}(t), \dot{m}(t)\big\rangle\,\mathrm{d}t,$$

an analog of the functional A_0, see (11.8), on the space of admissible curves with fixed end-points (for simplicity of presentation we consider the curves parametrized by the interval [0,T] as in (11.8)).

For a non-holonomic constraint, the notions of least constrained and length minimizing geodesics are not equivalent. Moreover, if the constraint is non-holonomic, the equation of the least constrained geodesics is not equivalent to any variational principle, even if the force is conservative (cf. Theorem 11.12). A more detailed discussion of this matter can be found, for example, in [224].

Remark 11.33. The two notions of geodesics we discuss here are due to Heinrich Hertz, who was apparently the first to notice that Newton's equation and the variational principle become non-equivalent to each other for a system with constraint [224, 225].

Theorem 11.34 (Chow-Rashevsky, see [225, 226]). *Let a constraint $\boldsymbol{\beta}$ be totally non-holonomic. Then for any two points $m_0, m_1 \in M$, there exists an admissible curve which joins m_0 and m_1.*

Corollary 11.35 (see [226]) *Let a constraint be totally non-holonomic. Then any two points in M can be joined by a length minimizing non-holonomic geodesic.*

Remark 11.36. The differential equation of length minimizing geodesics (see, e.g, [225]) involves admissible vectors as well as their annihilators (i.e., vectors in TM orthogonal to $\boldsymbol{\beta}$). Therefore, once the beginning $m_0 \in M$ of a length minimizing geodesics is specified, the space of initial conditions has dimension n. Thus, as mentioned above, if the constraint is totally non-holonomic, length minimizing geodesics (beginning at m_0) fill the entire manifold M. This question is discussed in more detail in [225, 226].

Theorem 11.37 *On a complete Riemannian manifold, the reduced connection is complete in the sense that all non-holonomic least constrained geodesics extend to $(-\infty, \infty)$.*

Indeed, since the Riemannian norm of all "velocity vectors" of a least constrained geodesic is constant, the Riemannian length of the curve is bounded on every finite interval. Thus, the arc of the geodesic taken over a finite interval is relatively compact because the manifold is complete. This, in turn, means that the geodesic extends to $(-\infty, \infty)$.

Corollary 11.38 (see [224]) *Let the constraint $\boldsymbol{\beta}$ on M be totally non-holonomic. Then any two points of M can be joined by a piecewise least constrained geodesic.*

Remark 11.39. The corollary is sharp: two generic points in M cannot be joined by a least constrained geodesic even if the constraint is totally non-holonomic. The equation of least constrained geodesics is a second order differential equation on the total space of the bundle $\boldsymbol{\beta}$. (The equation is given by the vector field $TP(Z)$ from (11.14).) Thus, the space of initial conditions of least constrained geodesics starting at a given point $m_0 \in M$ has dimension $k = \dim \boldsymbol{\beta}_m$ and the geodesics cannot fill the entire manifold M.

In the contemporary mathematical literature the study of least-constraint geodesics (or, more generally, of solutions of equations of type (11.13)) is called *non-holonomic dynamics* while the study of length minimizing geodesics (or more generally, of variational problems with non-holonomic constraints) is called *vakonomic dynamics*.

Remark 11.40. We should draw the reader's attention to the paper [37] where some geometric fundamentals of non-holonomic and vakonomic dynamics are investigated. For non-holonomic dynamics a certain invariant, the torsion of a special extension of a reduced connection, is found such that if it equals zero, the distribution $\boldsymbol{\beta}$ of the constraint is integrable. If the orthogonal complements to $\boldsymbol{\beta}_m$ form an integrable distribution, a (local) Ehresmann connection is constructed whose torsion characterizes the vakonomic dynamics. If this characteristic equal zero, the distribution $\boldsymbol{\beta}$ is also integrable.

Remark 11.41. We refer the reader to [107, Section 16] where stochastic differential equations in Belopolskaya-Daletskii form with constraints are introduced. The main idea is to use the exponential mapping of the reduced connection instead of the general exponent to get a constrained analog of the equations from Section 7.3. The constrained analogs of stochastic integral operators and the equations of Section 7.7 are also presented in [107, Section 16]. In their construction the parallel translation with respect to the reduced connection along the corresponding stochastic processes is applied. Note that the constrained Belopolskaya-Daletskii and integral approaches to SDEs with constraint have to be applied even for constraint equations in $\mathbb{R}^n$, assuming of course that the constraint is not trivial (i.e., does not consists of subspaces parallel to each other).

11.7 Mechanical Systems with Discontinuous Forces and Systems with Control. Differential Inclusions

Consider a mechanical system with a discontinuous force field. Such fields appear, for example, in systems with dry friction, switching, or with motion in several media having different resistance forces. When the configuration space is linear, the following method is often used to study systems with a discontinuous force. First, one extends the discontinuous force field to a set-valued vector field with convex images. Then one passes from (11.2) to a differential inclusion whose solutions are trajectories of the system (see [74]). This approach in linear space is knows as *Filippov's method*. In this section, we develop a similar method for non-linear configuration spaces [165].

The equation of motion of a mechanical system with feedback control may also be reduced to a differential inclusion. In this case, the set-valued force, a subset in every tangent space, is formed by all values of the force for all possible values of the controlling parameter at a given point.

The requisite notions and results on set-valued mappings we use here can be found in Chapter 4.

Consider a locally bounded vector field f on a finite-dimensional manifold M. The vector field f is not assumed to be continuous, nor even measurable.

For each point m_0, let us define a subset $R(m_0)$ of $T_{m_0}M$ as follows. The set $R(m_0)$ is formed by the limits of all sequences $f(m_k)$ as $m_k \to m_0$ with $m_k \neq m_0$. It is easy to see that

$$R(m_0) = \bigcap_{\epsilon>0} \left\{ \mathrm{cl} \left[\left(\bigcup_{m \in \mathfrak{U}_\epsilon} f(m) \right) \setminus f(m_0) \right] \right\},$$

where $\mathfrak{U}_\varepsilon$ is the ε-neighborhood of the point m_0 and cl denotes the closure.

Definition 11.42. The set $F(m_0) = \overline{\mathrm{co}} R(m_0) \subset T_{m_0}M$, where $\overline{\mathrm{co}}$ denotes the convex hull, is called the *essential extension* of the field f at m_0.

Definition 11.43. A set-valued map $f : \mathbb{R} \times TM \multimap TM$ such that for any point $(m, X) \in TM$ (meaning that $X \in T_m M$, i.e., X is a tangent vector to M at the point $m \in M$) the relation $\pi f(t, m, X) = \pi(m, X) = m$ holds is called a *set-valued vector force field* (cf. the Definition 11.2 of vector force fields).

The essential extension F is a set-valued mapping which assigns a subset of $T_{m_0}M$ to $m_0 \in M$. One can easily see that F is a set-valued vector field. Note that $F = f$ if f is continuous.

Theorem 11.44 *The set-valued vector field F is upper semicontinuous.*

Proof. Let $\delta > 0$ be a real number. Fix a metric ρ on TM which gives rise to a topology equivalent to that on the tangent bundle. Denote the δ-neighborhoods of $R(m_0)$ and $F(m_0)$ by $\mathbb{R}^\delta(m_0)$ and $F^\delta(m_0)$, respectively. We prove that for any δ and any $m \in M$, there exists a neighborhood $U(m) \subset M$ of m such that $R(m') \subset \mathbb{R}^\delta(m)$ for every $m' \in U(m)$ and, therefore, $F(m') \subset F^\delta(m)$.

By the definition of the set $R(m)$, there exists a neighborhood $U(m)$ of m such that for all $m' \in U(m)$ we have $\rho(f(m'), R(m)) < \delta$. Then there exists an open neighborhood $V(m') \subset U(m)$ of the point m' such that the inequality $\rho\big(f(m''), R(m')\big) < \delta$ is satisfied for every $m'' \in V(m')$. Pick a sequence $m''{}_k \to m'$ in $V(m')$. We have

$$\lim \rho\big(f(m''_k), R(m)\big) = \rho \lim\big(f(m''_k), R(m)\big) < \delta \ .$$

Hence, $R(m') \subset \mathbb{R}^\delta(m)$ and $F(m') \subset F^\delta(m)$. □

Now consider a mechanical system with configuration space M and kinetic energy $\mathcal{K}(X) = \langle X, X\rangle/2$, where $\langle\cdot,\cdot\rangle$ is the Riemannian metric on M. Let $\bar{\alpha}(t, m, X)$ be a force field that we require only to be locally bounded in all variables. (As above, we do not assume that $\bar{\alpha}$ is continuous or even measurable.) Consider the vector field $Z(m, X) + \bar{\alpha}(t, m, X)^l$ (i.e., a second order differential equation on M; see Section 3.3), where Z is the geodesic spray of the Levi-Civitá connection of $\langle\cdot,\cdot\rangle$ and $\bar{\alpha}(t, m, X)^l$ is the vertical lift of $\bar{\alpha}(t, m, X)$ to the point $(m, X) \in TM$ (see (4.3)). It is easy to see that the essential extension (with respect to all variables) of $Z(m, X) + \bar{\alpha}(t, m, X)$ may be written in the form

$$Z(m, X) + \mathfrak{a}(t, m, X)^l, \tag{11.15}$$

where $\mathfrak{a}(t, m, X)^l$ is the vertical lift of the essential extension $\mathfrak{a}(t, m, X)$ of $\bar{\alpha}(t, m, X)$ to the point (m, X). Note that $\mathfrak{a}(t, m, X) = \overline{\text{co}}Q(t, m, X)$, where $Q(t, m, X)$ is the set of limit points of all sequences $\bar{\alpha}(t_k, m_k, X_k)$ such that $(t_k, m_k, X_k) \to (t, m, X)$, $X_k \in T_{m_k}M$, and $(t_k, m_k, X_k) \neq (t, m, X)$.

From now on, we focus on the differential inclusion in TM given by the formula

$$\dot{\gamma}(t) \in Z\left(\gamma(t)\right) + \mathfrak{a}\left(t, \gamma(t)\right)^l. \tag{11.16}$$

Definition 11.45. A solution of (11.16) is an absolutely continuous curve $\gamma(t)$ in TM which almost everywhere satisfies (11.16).

Alternatively, making use of covariant derivatives, we consider the following differential inclusion on M:

$$\frac{\mathrm{D}}{\mathrm{d}t}\,\dot{m}(t) \in \mathfrak{a}\left(t, m(t), \dot{m}(t)\right). \tag{11.17}$$

Definition 11.46. A solution of (11.17) is a C^1-curve $m(t)$ in M such that $\dot{m}(t)$ is absolutely continuous and (11.17) is almost everywhere satisfied.

Taking into account (11.15) and the definition of $\frac{\mathrm{D}}{\mathrm{d}t}$, it is easy to check that (11.16) and (11.17) are equivalent. More precisely, this means that $m(t)$ is a solution of (11.17) if and only if $\dot{m}(t)$, regarded as a curve in TM, is a solution of (11.16).

Definition 11.47. A solution of (11.17) is called a *trajectory* of the mechanical system with discontinuous force field $\bar{a}$.

It is easy to see that Definition 11.47 is justified from the physical point of view. As mentioned above, for a flat configuration space the reasons for supporting the definition are discussed, for example, in [74].

The right-hand side of (11.16) is an upper semicontinuous set-valued vector field with convex images. This implies that locally there exists a solution of the Cauchy problem for (11.16) (see Chapter 4). Thus, for any initial conditions $m \in M$ and $X \in T_mM$, inclusion (11.17) has a solution on a sufficiently small interval.

Note that an interesting question for applications in physics is whether or not the local solution of (11.17) is unique. Certain uniqueness conditions are found in [74].

Another class of mechanical systems involving inclusions like (11.17) is the class of mechanical systems with feedback control. Let the force field

$\bar{\alpha}(t, m, X, u)$ depend on the parameter $u \in U$. We define the set-valued vector field $\mathfrak{a}(t, m, X)$ on TM as

$$\mathfrak{a}(t, m, X) = \bigcup_{u \in U} \bar{\alpha}(t, m, X, u).$$

Now we have to assume that this field is upper semicontinuous and has closed convex images. The solution of (11.17) is a trajectory of the control system for a time-dependent control $u(t)$. Let us point out that, since the configuration space is non-linear, we cannot assume the control force to be independent of time, coordinates or velocity, as is usually the case for linear spaces. Some very particular examples where such an assumption does make sense will be considered in Section 11.9 and in Chapter 12.

Note that in systems with control the inclusions of type (11.17) with lower semicontinuous $\mathfrak{a}$ can also arise. Let us consider an example of such an inclusion.

Consider a set-valued bounded and Hausdorff continuous vector force field $\mathfrak{a}(t, m, X)$ with convex closed values.

Definition 11.48. A point a of a convex set $\mathfrak{a}$ is called *extreme* if there does not exist an open interval of a straight line in $\mathfrak{a}$ that contains a. Denote by Ext $\mathfrak{a}(t, m, X)$ the set-valued vector force field whose values at all points (t, m, X) consist of extreme points of $\mathfrak{a}(t, m, X)$.

Lemma 11.49 *For a set-valued bounded Hausdorff continuous vector force field* $\mathfrak{a}(t, m, X)$ *with convex closed values the set-valued vector force field* $\operatorname{Ext}\mathfrak{a}(t, m, X)$ *is lower semicontinuous.*

Lemma 11.49 is a well-known statement from set-valued analysis. A proof can be found in [222, Lemma 2.1.1] and in [48, Proposition 6.2]. Note that $\operatorname{Ext}\mathfrak{a}(t, m, X)$ is bounded and may not have convex values.

Definition 11.50. We say that a trajectory $m(t)$ of a mechanical system with Hausdorff continuous vector force field $\mathfrak{a}(t, m, X)$ *occurs under extremal values of controlling force* if almost everywhere $\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t)$ belongs to Ext $\mathfrak{a}(t, m(t), \dot{m}(t))$.

Thus for a mechanical system with control given by (11.17), with set valued vector force field as above, the problem of the existence of solutions that occurs under extremal values of controlling force is reduced to the differential inclusion

$$\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) \in \operatorname{Ext}\ \mathfrak{a}(t, m(t), \dot{m}(t)) \tag{11.18}$$

with a lower semicontinuous right-hand side having non-convex values.

11.8 Integral Equations of Geometric Mechanics. The Velocity Hodograph

In this chapter we use the integral operators with parallel translation introduced in Section 3.2 to find integral equations equivalent to Newton's equation of geometric mechanics (see Section 11.1). One of these equations describes the velocity hodograph in the sense of [217]. This is an ordinary integral equation in a specified tangent space. We also introduce analogous integral equations for a system with constraint. Our approach is based on the results of [88, 90, 94, 98].

Integral versions of Newton's equation and, in particular, the equation of the velocity hodograph turn out to be useful in the study of certain qualitative problems concerning the behavior of mechanical systems, for example, the existence of special trajectories. It is important to emphasize that the equation of the velocity hodograph is an integral equation in a linear space, and therefore standard methods may be applied to study it. Integral equations are used in Chapter 12 and also in Chapters 14 and 15 where we work with versions of integral equations for random force fields.

11.8.1 General constructions

Consider a mechanical system as in Section 11.1. We assume here that the Riemannian metric $\langle\cdot,\cdot\rangle$ is complete (and so a trajectory of a free particle cannot escape to infinity in finite time) and that the vector force field $\bar{\alpha}(t,m,X)$ is jointly continuous in all variables. The case of discontinuous force fields will be studied in Chapter 12.

Since the metric is complete, we can use the operator $\mathcal{S}$ introduced in Section 3.2

Let $\varGamma\bar{\alpha}\big(t,m(t),\dot{m}(t)\big)$ denote the curve in $T_{m_0}M$ such that the vector $\varGamma\bar{\alpha}(t,m(t),\dot{m}(t))$ is parallel to $\bar{\alpha}(t,m(t),\dot{m}(t))$ along $m(\cdot)$ for every t. Specify a point $m_0 \in M$ and a vector C in $T_{m(0)}M$ and consider the integral equation

$$m(t) = \mathcal{S}\left(\int_0^t \varGamma\bar{\alpha}\left(\tau, m(\tau), \dot{m}(\tau)\right)\,\mathrm{d}\tau + C\right) \tag{11.19}$$

on $I = [0,l]$.

Theorem 11.51 *The solutions of* (11.2) *with the initial conditions* $m(0) = m_0$ *and* $\dot{m}(0) = C$ *coincide with the solutions of* (11.19).

Proof. It is easy to show that for a given $v \in C^0(I, T_{m_0}M)$, the C^2-curve

$$m(t) = \mathcal{S}\left(\int_0^t v(\tau)\,\mathrm{d}\tau + C\right)$$

is the only one satisfying the conditions $m(0) = m_0$, $\dot{m}(0) = C$ and such that for every $t \in I$ the vector $\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t)$ is parallel to $v(t)$ along $m(\cdot)$. To see this, let for some $t \in I$ the curve $\bar{m}(\cdot)$ in $T_{m(t)}M$ be obtained by the parallel translation of $\dot{m}$ to the point $m(t)$. Then, by Theorem 2.32 we have

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\bar{m}(t+\tau)_{|\tau=0} = \frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t).$$

It is clear that the vector $\int_0^t v(\tau)\,\mathrm{d}\tau + C \in T_{m_0}$ is parallel to $\bar{m}(t)$ along $m(\cdot)$. In other words, the vector $\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t)$ is parallel to $v(t)$ along $m(\cdot)$. Let us set $v(t) = \varGamma\bar{\alpha}\big(t, m(t)\dot{m}(t)\big)$. Then (11.19) means that the vector $\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t)$ can be obtained by transporting $\bar{\alpha}\,(t, m(t), \dot{m}(t)) \in T_{m(t)}M$ along $m(\cdot)$ first to the point $m_0 = m(0)$ and then back to $m(t)$. The theorem follows. $\square$

Let $m(t)$, $t \in I$, be a trajectory of the mechanical system, i.e., a solution of (11.2).

Definition 11.52. The *velocity hodograph* of the trajectory $m(t)$ is the curve $v\colon I \to T_{m(0)}M$ such that $v(t)$ is parallel to $\dot{m}(t)$ along $m(\cdot)$.

It is not hard to see that the velocity hodograph of a solution of (11.19) satisfies the equation

$$v(t) = \int_0^t \varGamma\bar{\alpha}\left(\tau, \mathcal{S}v(\tau), \frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{S}v(\tau)\right)\,\mathrm{d}\tau + C, \tag{11.20}$$

It is clear that if v is a solution of (11.20), then $\mathcal{S}v$ is a solution of (11.19), i.e., by Theorem 11.51, a trajectory of the mechanical system.

Remark 11.53. The notion of a velocity hodograph in the sense of Definition 11.52 was introduced by Synge in [217] where analogs of the standard properties of the hodograph were proved for some mechanical systems (see also [218]). The hodograph equation (11.20) originally appeared in [88].

Remark 11.54. If we have a mechanical system on a group, then it is natural to pick the initial condition $m(0) = e$. Thus, (11.20) becomes an equation in the Lie algebra similar to the Euler equation in the body or space coordinates. All three equations are equivalent to Newton's equation on the configuration space. However, for an arbitrary configuration space, (11.20) is the only one among these three that makes sense.

Let us denote the operator which sends $v \in C^0(I, T_{m_0}M)$ to

$$\int_0^t \varGamma\bar{\alpha}\left(\tau, \mathcal{S}v(\tau), \frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{S}v(\tau)\right)\,\mathrm{d}\tau + C \in C^0(I, T_{m_0}M)$$

by $\int \varGamma \circ \bar{\alpha} \circ \mathcal{S}_C$.

Theorem 11.55 *The operator*

$$\int \Gamma \circ \bar{\alpha} \circ \mathcal{S}_C \colon C^0(I, T_{m_0}M) \to C^0(I, T_{m_0}M)$$

is completely continuous.

Proof. Since $\mathcal{S}$, $\bar{\alpha}$ and Γ are continuous, so is the operator. Let $\mathcal{U}_K$ be the ball in $C^0(I, T_{m_0}M)$ of radius K centered at the origin. Because $\bar{\alpha}$ is continuous, Theorem 3.46 and Lemma 3.53 imply that $\left(\int \Gamma \circ \bar{\alpha} \circ \mathcal{S}_C\right)(\mathcal{U}_K)$ is compact. $\square$

11.8.2 An integral formalism of geometric mechanics with constraints

Let the configuration space M be a complete Riemannian manifold equipped with a constraint $\boldsymbol{\beta}$, which may be non-holonomic (Section 11.6). To develop an adequate integral formalism, we use parallel translation of admissible vectors along admissible curves. Such a parallel translation arises from the reduced connection $\bar{H}$.

Let $m(t)$, $t \in I$, be an admissible C^1-curve and $X(t,m)$ an admissible vector field on M. Denote by $\Gamma^{\boldsymbol{\beta}} X\big(t, m(t)\big)$ the curve in $\boldsymbol{\beta}_{m(0)}$ such that the vector $X\left(t, m(t)\right)$ at $m(0)$ is parallel to $\Gamma^{\boldsymbol{\beta}} X\left(t, m(t)\right)$ along $m(\cdot)$ under the reduced connection. The properties of the operator $\Gamma^{\boldsymbol{\beta}}$ are similar to those of Γ studied in Section 3.2.2.

As in Section 11.6.2 consider the orthonormal frame bundle $\bar{\pi}\colon \mathbf{O}^{\boldsymbol{\beta}}(M) \to M$ of $\boldsymbol{\beta}$ (i.e., $b \in \mathbf{O}^{\boldsymbol{\beta}}_m$ is an orthonormal frame in $\boldsymbol{\beta}_m$).

Consider the map $\bar{E}\colon \mathbf{O}^{\boldsymbol{\beta}} \times \mathbb{R}^k \to \bar{H}$, $k = \dim \boldsymbol{\beta}_m$, defined by the formula $\bar{E}_b(X) = T\pi^{-1}(bX) \mid_{\bar{H}_b}$, where $b \in \mathbf{O}^{\boldsymbol{\beta}}_m(M)$ is regarded as an orthogonal operator from $\mathbb{R}^k$ to β_m (cf. Definition 2.68 of basic vector fields in the non-holonomic case). It is easy to see that $\bar{E}$ is smooth and fiber-wise linear. Let $v(t)$ be a continuous curve in $\boldsymbol{\beta}_{m_0}$. Fix $b_0 \in \mathbf{O}\boldsymbol{\beta}_{m_0}(M)$ and consider the time-dependent vector field $\bar{E}\big(b_0^{-1}v(t)\big)$ on $\mathbf{O}^{\boldsymbol{\beta}}(M)$. By definition, this vector field is smooth in b for a fixed t and continuous in t for a fixed $b \in \mathbf{O}^{\boldsymbol{\beta}}(M)$. Hence, for this vector field, the Cauchy problem has a solution. As above, it is easy to show that the integral curves of $\bar{E}\left(b_0^{-1}v(t)\right)$ extend to the entire interval $I = [0, l]$. Consider the integral curve $b_0(t)$ beginning at b_0 and its projection $\mathcal{S}^{\boldsymbol{\beta}} v(t) = \bar{\pi} b_0(t)$. It is clear that $\mathcal{S}^{\boldsymbol{\beta}} v(\cdot)$ is an admissible curve and, in addition, for every $t \in I$ the vector $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}^{\boldsymbol{\beta}} v(t)$ is parallel to $v(t)$ along $\mathcal{S}\boldsymbol{\beta}_v(\cdot)$ with respect to the reduced connection $\bar{H}$.

The following result can be proved in the same way as Theorem 3.56.

Theorem 11.56 *Let $X(t,m)$ be an admissible vector field which is jointly continuous in all its variables. An admissible curve $m(t)$ is an integral curve* *of $X(t,m)$ if and only if it satisfies the equation*

$$m(t) = \mathcal{S}^{\boldsymbol{\beta}} \circ \Gamma^{\boldsymbol{\beta}} X\left(t, m(t)\right). \tag{11.21}$$

Theorem 11.57 *A continuous curve $v(t) \subset \beta_{m_0}$ is a solution of the equation $v(t) = \Gamma^{\boldsymbol{\beta}} X\big(t, \mathcal{S}^{\boldsymbol{\beta}} v(t)\big)$ if and only if $\mathcal{S}^{\boldsymbol{\beta}} v(t)$ satisfies (11.21).*

The operators $\mathcal{S}^{\boldsymbol{\beta}}$, $\Gamma^{\boldsymbol{\beta}}$ and their compositions have the same compactness and continuity properties as the integral operators from Section 3.2.

Now consider the integral equation

$$m(t) = \mathcal{S}^{\boldsymbol{\beta}} \left(\int_0^t \Gamma^{\boldsymbol{\beta}} P \bar{\alpha}\left(\tau, m(\tau), \dot{m}(\tau)\right) \mathrm{d}\tau + C \right), \tag{11.22}$$

where C is a vector in $\boldsymbol{\beta}_{m_0}$. Taking into account the relationship between parallel translations and covariant derivatives, we get the following result.

Theorem 11.58 *An admissible curve with the initial conditions $m(0) = m_0$ and $\dot{m}(0) = C$ satisfies (11.13) (i.e., $m(\cdot)$ is a trajectory of the system with the constraint β) if and only if it is a solution of (11.22).*

It is clear that the equation of the velocity hodograph of a solution of (11.22) is

$$v(t) = \int_0^t \Gamma^{\boldsymbol{\beta}} P \bar{\alpha} \left(\tau, \mathcal{S}^{\boldsymbol{\beta}} v(\tau), \frac{\mathrm{d}}{\mathrm{d}\tau} \mathcal{S}^{\boldsymbol{\beta}} v(\tau) \right) \mathrm{d}\tau + C \tag{11.23}$$

on the space of continuous curves in $\boldsymbol{\beta}_{m_0}$.

Remark 11.59. We emphasize that even if M is the Euclidean space $\mathbb{R}^n$, the integral operators considered in this subsection (e.g., $\mathcal{S}^{\boldsymbol{\beta}}$, $\mathcal{S}^{\boldsymbol{\beta}} \circ \Gamma^{\beta}$, etc.) cannot be reduced to their classical analogs (antiderivatives, the Urysohn-Volterra operator, etc.) unless the distribution $\boldsymbol{\beta}$ is trivial, i.e., all the spaces $\boldsymbol{\beta}_m$ are parallel (in the Euclidean sense) to $\boldsymbol{\beta}_0 \subset T_0\mathbb{R}^n = \mathbb{R}^n$.

11.9 Mechanical Interpretation of Parallel Translation and Systems with Delayed Control Forces

In this section, following [91, 92, 93, 94], we study a certain type of differential equation with delay on Riemannian manifolds. In these equations one of the terms on the right-hand side is obtained by parallel translation to the corresponding point along a solution. The equations are analogous to those differential equations on Euclidean space with discrete delay or where the right-hand side depends only on time. Our analysis of the equations in terms of geometric mechanics is based on the mechanical interpretation of parallel translation.

The mechanical interpretation of Riemannian parallel translation was discovered by Johann Radon and described by Blaschke in [159]. A similar idea was independently used by Synge [217] in order to define the hodograph for a geometric mechanical system. (See Remark 11.53.)

Radon proved that for a pendulum moving in the configuration space of a mechanical system, the direction of oscillation is a parallel translation along its trajectory. In other words, the coordinate system attached to a gyroscope (e.g., the stationary system for a flat configuration space) is parallel along a trajectory.

Consider the motion of a mechanical system with a force field $\bar{\alpha}$ and a "control" force $\varPhi$. The latter depends on time, the velocity and on the coordinates of the point. Faithfully modeling the delays that occur in real life systems, $\varPhi$ acts after time h. The equation of motion of such a system is as follows:

$$\frac{\mathrm{D}}{\mathrm{d}t}\,\dot{m}(t) = \bar{\alpha}\left(t, m(t), \dot{m}(t)\right) + \big\|\varPhi\left(t-h, m(t-h), \dot{m}(t-h)\right), \tag{11.24}$$

where $\|$ denotes the Riemannian parallel translation along the solution.

Similarly, one may consider the evolution of a system with a velocity field V and delayed control "velocity" W, i.e., a system given by the equation

$$\dot{m}(t) = V\left(t, m(t)\right) + \big\|W\left(t-h, m(t-h)\right). \tag{11.25}$$

If M is a Euclidean space, (11.24) and (11.25) are quite simple differential equations with discrete constant delay [4]. However, if M is not flat, (11.24) and (11.25) have much more complex properties.

First, since the parallel translation is defined along C^1-curves and depends on a curve and its derivative, (11.25) is an equation of neutral type [4].

Secondly, the first order equation corresponding to (11.24) has distributed delay. The reason is that if M is not flat, the parallel translations of a vertical vector in TM do not necessarily coincide with the lift of the parallel translation in M. (Here the manifold TM is equipped with the standard metric arising from the metric on M.)

Finally, note that the first order equation, which is equivalent to (11.24) and thus, as we have just explained, has distributed delay, is again an equation of neutral type on TM because the right hand side is neither continuous in the C^0-topology nor defined on arbitrary curves.

Let $\phi\colon [-h, 0] \to M$ be a C^1-curve.

Definition 11.60. A continuous curve $m(\cdot)\colon [-h, \varepsilon) \to M$, $\varepsilon > 0$ is a solution of (11.24) (respectively, (11.25)) on the interval $[-h, \varepsilon)$ with initial condition ϕ if $m(\cdot)$ is C^1-smooth on $(0, \varepsilon)$, coincides with ϕ on $[-h, 0]$, and satisfies (11.24) (respectively, (11.25)) on $[\,0, \varepsilon)$.

It is useful to first analyze the particular cases of (11.24) and (11.25) where the control force depends on time only. These mechanical systems are given by the equations

$$\frac{\mathrm{D}}{\mathrm{d}t}\,\dot{m}(t) = \bar{\alpha}\,(t, m(t), \dot{m}(t)) + \big\|\Phi(t) \tag{11.26}$$

$$\dot{m}(t) = V\,(t, m(t)) + \big\|W(t) \tag{11.27}$$

In this case Φ and W take values in the tangent space to M at the initial point m_0.

It is essential that (11.24) and (11.25) can be reduced to (11.26) and (11.27), respectively. To see this, specify an initial condition $\phi \in C^1([-h, 0], M)$ and consider the following $T_{\phi(0)}M$-valued functions of $\theta \in [-h, 0]$: $\Phi(\theta) = \|\Phi(\theta+h, \phi(\theta+h), \dot{\phi}(\theta+h))$ and $W(\theta) = \|W(\theta+h, \phi(\theta+h))$. Let $t = \theta + h$. It is clear that the solutions of (11.26) and (11.27) with control forces $\Phi(t)$ and $W(t)$ coincide with those of (11.24) and (11.25), respectively.

Note that the equations above make sense only if their solutions are C^1-smooth, for the parallel translation is not defined otherwise.

Theorem 11.61 *Let $V(t, m)$ be a continuous vector field on M and $W(t)$ a continuous curve in $T_{m_0}M$. Then:*

(i) *equation (11.27) has a local C^1-solution;*
(ii) *the solution is unique provided that for any $t \in I$ the field $V(t, m)$ is locally Lipschitz continuous in m.*

Proof. Let $\mathbf{O}(M)$ be the orthonormal frame bundle over M and H the Levi-Civitá connection on $\mathbf{O}(M)$. The tangent map of the natural projection $\pi\colon \mathbf{O}(M) \to M$ induces the isomorphism $T\pi\colon H_b \to T_{\pi b}M$ at every point $b \in \mathbf{O}(M)$. Hence, at every $b \in \mathbf{O}(M)$, we obtain the vector $T\pi^{-1}V(t, \pi b) \in H_b \subset T_b\mathbf{O}(M)$. These vectors form a horizontal vector field on $\mathbf{O}(M)$ (i.e., belonging to H).

Let $\mathcal{O}$ be an orthonormal frame in $T_{m_0}M$. The frame $\mathcal{O}$ gives rise to the isomorphism $\mathcal{O}\colon \mathbb{R}^n \to T_{m_0}M$ and, therefore, we have a horizontal time-dependent basic vector field $E(\mathcal{O}^{-1}W(t))$ on $\mathbf{O}(M)$ (see Definition 2.68).

Consider the vector field $\tilde{V}(t, b) = T\pi^{-1}V(t, \pi b) + E(\mathcal{O}^{-1}W(t))$ on $\mathbf{O}(M)$. By the existence theorem for ordinary differential equations, this vector field has an integral curve γ^*, $\gamma^*(0) = \mathcal{O}$, defined on the interval $[\,0, \epsilon)$. Since the vector field $\tilde{V}$ is horizontal, the frame $\gamma^*(t)$ is parallel to $\mathcal{O}$ along $\gamma = \pi\gamma^*$ for every $t \in [\,0, \varepsilon)$. By definition, we have $\dot{\gamma}^*(t) = T\pi^{-1}[V(t, \gamma(t)) + \gamma^*(t)(\mathcal{O}^{-1}W(t))]$, where the last term is the vector with coordinates $\mathcal{O}^{-1}W(t)$ in the basis $\gamma^*(t)$. Hence, $\dot{\gamma}(t) = V(t, \gamma(t)) + \gamma^*(t)\mathcal{O}^{-1}(W(t))$. Furthermore, by the definition of parallel translation, the last term is parallel to $W(t)$ along $\gamma(\cdot)$. This means that $\gamma(t)$ is a solution of (11.27).

Note that the field $E\big(\mathcal{O}^{-1}W(t)\big)$ is smooth in b for a fixed t. If $V(t, m)$ is locally Lipschitz in m (for every fixed t), then $\tilde{V}$ is locally Lipschitz in b and the resulting solution is unique. $\square$

Corollary 11.62 *Assume that the fields $V(t,m)$ and $W(t,m)$ are jointly continuous in all variables. Then for any C^1-initial condition $\phi\colon [-h,0] \to M$, there exists a solution of (11.25). The solution is unique if V is locally Lipschitz in m for any fixed t.*

Theorem 11.63 *Assume that the field $\bar{\alpha}$ satisfies the Carathéodory condition (see, e.g., [74]) and $\Phi(t)$ is an integrable function with values in $T_{m_0}M$. Then for any initial condition $C \in T_{m_0}M$:*

(i) *equation (11.26) has a local C^1-smooth solution;*
(ii) *the solution is unique if $\bar{\alpha}$ is locally Lipschitz continuous for all t.*

Proof. Consider the direct product $\mathbf{O}(M) \times \mathbb{R}^n$, $n = \dim M$, equipped with the right action of $\mathbf{O}(n)$ given by the formula $(b,x)g = (bg, g^{-1}x)$, where $b \in \mathbf{O}(M)$, $x \in \mathbb{R}^n$ and $g \in \mathbf{O}(n)$. The quotient space of $\mathbf{O}(M) \times \mathbb{R}^n$ under the right action can be naturally identified with TM. We denote the natural projection $\mathbf{O}(M) \times \mathbb{R}^n \to TM$ by λ (see Section 2.7 and Notation 1.36).

Let $(b,x) \in \mathbf{O}(M) \times \mathbb{R}^n$. It is easy to see that $T\lambda$ induces an isomorphism of $H_b \in T_b\mathbf{O}(M)$ with the horizontal space in $T_{\lambda(b,x)}TM$ and an isomorphism of $V_x = T_x\mathbb{R}^n$ with the vertical space $V_{\lambda(b,x)}$. (Recall that the latter is the tangent space to $T_{\pi b}M$, where $\pi\colon \mathbf{O}(M) \to M$ is the natural projection.)

Pick an orthonormal basis $\mathcal{O}$ in $T_{m_0}M$ and define the function $\mathcal{O}^{-1}\Phi(t)$ with values in $\mathbb{R}^n$ to be the coordinates of $\Phi(t)$ with respect to the basis $\mathcal{O}$. As in the proof of Theorem 11.61, the function gives rise to the horizontal vector field $E(\mathcal{O}^{-1}\Phi(t))$ on $\mathbf{O}(M)$. We see that any basis $b \in \mathbf{O}(M)$ gives rise to a vector $T\pi E_b(\mathcal{O}^{-1}\Phi(t))$ in the tangent space $T_{\pi b}M$.

Consider the vector fields A, B, and C on $\mathbf{O}(M) \times \mathbb{R}^n$ such that for any $(b,x) \in \mathbf{O}(M) \times \mathbb{R}^n$ we have $A_{(b,x)} = T\lambda^{-1} Z_{\lambda(b,x)} \in H_b$, where Z is the spray of the Levi-Civitá connection on M and

$$
\begin{aligned}
B_{(b,x)}(t) &= T\lambda^{-1}\left(\bar{\alpha}\left(t, \bar{\pi}\lambda(b,x), \lambda(b,x)\right)\right) \in V_x,\\
C_{(b,x)}(t) &= T\lambda^{-1}\left(T\pi E_b\left(\mathcal{O}^{-1}\Phi(t)\right)\right) \in V_x.
\end{aligned}
$$

By definition, A is a smooth field. Since $\bar{\alpha}$ satisfies the Carathéodory condition, so does $B(t)$. By the hypothesis of the Theorem, the field $C(t)$ is smooth on $\mathbf{O}(M) \times \mathbb{R}^n$ for every fixed t and is measurable in t for any fixed (b,x). Therefore, the field $A + B(t) + C(t)$ satisfies the hypothesis of the classical theorem which guarantees the existence of a local solution of the Cauchy problem. Furthermore, if $\bar{\alpha}$ is locally Lipschitz in t, then the hypothesis of the uniqueness theorem is also satisfied (see [62], Theorems 1 and 2 of Section 1). Note that local solutions are, by construction, absolutely continuous curves.

Let $(b(t), x(t))$ be the local solution with initial condition $(\mathcal{O}, \mathcal{O}^{-1}C)$. The curve $\lambda(b(t), x(t))$ is absolutely continuous in TM and, hence, the tangent vector

$$
Y(t) = T\lambda\left(A + B(t) + C(t)\right) = Z_{\lambda(b(t),x(t))} + T\lambda\left(B(t) + C(t)\right)_{\lambda(b(t),x(t))}
$$

exists for almost all t. The vector $Z_{\lambda(b(t),x(t))}$ belongs to the connection and both vectors $T\lambda(B(t))_{\lambda(b(t),x(t))}$ and $T\lambda(C(t))_{\lambda(b(t),x(t))}$ are in the vertical subspace. Hence, $T\pi Y(t) = T\pi Z$. On the other hand, since Z is the spray, $T\bar{\pi} Z_{\lambda(b(t),x(t))} = \lambda(b(t), x(t))$. As one can easily see, this means that the curve $\lambda(b(t), x(t))$ in TM has the form $(\gamma(t), \dot{\gamma}(t))$, where $\gamma(t) = \pi\lambda(b(t), x(t))$ is a C^1-curve. In particular, the parallel translation is defined along γ.

By definition, the projection of $(b(t), x(t))$ to $\mathbf{O}(m)$ is horizontal, and so $b(t)$ is a parallel frame field along γ. Hence, for every t the vector $T\lambda(C(t)) \in T_{\gamma(t)}M$ is parallel to $\Phi(t)$ along γ. Taking into account the definition of the covariant derivative, we see that γ satisfies (11.26). It is clear that $\gamma(0) = m_0$ and $\dot{\gamma}(0) = C$. □

Corollary 11.64 *Assume that $\bar{\alpha}$ satisfies the Carathéodory condition and $\Phi(t, m, X)$ is jointly measurable in all variables. Then for any C^1-curve $\phi\colon [-h, 0] \to M$, there exists a local solution of* (11.24) *with initial condition ϕ provided that $\|\Phi(t, \phi(t), \dot{\phi}(t))\|$ is integrable on $[-h, 0]$. If for every fixed t the vector field $\bar{\alpha}$ is locally Lipschitz in (m, X), then the solution is unique.*

Theorem 11.65 *Let the Riemannian metric $\langle\cdot,\cdot\rangle$ be complete. Assume also that for some point $m_0 \in M$ the inequalities $\|\bar{\alpha}(t, m, X)\| \le \Psi(t)L(\rho(m_0, m))$ and, respectively, $\|V(t, m)\| \le \Psi(t)L(\rho(m_0, m))$ hold where the function L is defined in Section* 3.1.4 *and satisfies* (3.16)*, ρ is the Riemannian distance on M, and Ψ is a positive function integrable on finite intervals. Then the solutions of* (11.24) *and* (11.25)*, respectively, are defined on $[-h, \infty)$.*

Proof. Without loss of generality we may assume that $\inf L = C > 0$. (Otherwise, we simply replace L by $L + C$.) Let us rewrite (11.24) and (11.25) in their equivalent forms (11.26) and (11.27), respectively, and consider the complete Riemannian metric $\langle\cdot,\cdot\rangle^*$ introduced in Section 3.1.4. In this metric, $\|V(t, m)\|^* \le \Psi(t)$. The norm of $W(t)$ with respect to $\langle\cdot,\cdot\rangle$, being a continuous function, is bounded on $[-h, 0]$ by a constant $K > 0$. It is easy to show that any local solution $m(t)$ of (11.27) satisfies the inequality

$$\|\dot{m}(t)\| < \Psi(t) + \frac{K}{C},$$

where the norm $\|\cdot\|$ is taken with respect to $\langle\,,\rangle$. Thus, $m(t)$ extends to $[-h, h]$. Covering any given interval I by intervals of length h, one can prove that the solution extends to I. For (11.24) the proof is similar. □

We conclude this section by noting that the shift operators along solutions of (11.24), (11.25) and some other neutral type equations were studied in [28, 91, 92, 93]. The existence of fixed points of these operators (i.e., periodic solutions for a periodic right-hand side) was proved by the methods of [28, 33].

Chapter 12
Accessible Points and Sub-Manifolds of Mechanical Systems. Controllability

12.1 Discussion of the Problem

In this chapter we study the question of whether or not two points m_0 and m_1 in the configuration space of a mechanical system can be connected by a trajectory. It is known (see, e.g., [144]) that for a second order differential equation (i.e., in particular, for Newton's law) on Euclidean space such a trajectory exists provided that the right-hand side of the differential equation is bounded and continuous. More precisely, for any two points m_0 and m_1 and any interval $[a, b]$, there exists a solution $m(t)$ such that $m(a) = m_0$ and $m(b) = m_1$. When the right-hand side is linearly bounded, some similar results are known for small time intervals.

Note that if the right-hand side is quadratic in velocity, even in $\mathbb{R}^n$ there may be pairs of points that cannot be connected by a solution. In Section 12.2 we give a simple example of this phenomenon.

The situation becomes much more complicated for a non-linear configuration space. In Section 12.2, we illustrate this with four examples of mechanical systems on the two-dimensional sphere. In the first example, the force field is smooth and independent of time and velocity (and so it is bounded). However, none of the trajectories beginning at the South Pole reach the North Pole. In the second example, the force field is still bounded, autonomous, and smooth but now depends on the velocity. In this case there is no trajectory connecting any two antipodal points on the sphere. In the third example, we consider a gyroscopic force on S^2 (hence, the force field is linear in velocity). The behavior of trajectories in this system turns out to be quite similar to the second example. The same behavior occurs in the fourth example where the force field is quadratic in velocity.

There is a deep geometric reason for the difference in the behavior of trajectories on flat and "curved" configuration spaces. The points on a sphere in the above-mentioned examples are conjugate along all geodesics joining them. Since in the flat case conjugate points are absent, this is not true in $\mathbb{R}^n$.

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_12,

Below we show that if the force field on a complete Riemannian manifold has less than quadratic growth in velocity, then for any two points m_0 and m_1, there exists a trajectory joining m_0 and m_1 provided that the points are not conjugate along some geodesic. This is true only for a small enough time interval, even if the force field is uniformly bounded (however, we show that from our construction it follows that for uniformly bounded forces on flat configuration spaces the required solution exists on every finite time interval). For force fields quadratic in velocity we find a geometric condition on the distance between points (not conjugate along at least one geodesic), and conditions on the geometry of the manifold and on the right-hand side of the equation, under which the solution, joining the points, exists at least on a small enough time interval. In fact the existence of the required trajectory in the case of less than quadratic growth follows from the fact that the latter case always satisfies the above-mentioned condition. We also find a subclass of systems with quadratic growth for which, under the above condition, the solution exists on every finite time interval.

Besides its mechanical interpretation, the problem where the force is quadratic in velocity is important since it is a generalization of the well-known classical question which asks whether it is possible to join two given points in a manifold by a geodesic curve of a certain connection (see, e.g., [161]). Recall (see Theorem 2.36) that if ∇ and $\bar{\nabla}$ are covariant derivatives of two different connections on a manifold M, there exists a $(1,2)$-tensor field $S(\cdot,\cdot)$ on M such that for any two vector fields X and Y on M the equality $\bar{\nabla}_X Y = \nabla_X Y + S(X,Y)$ holds. From this it follows that in terms of the covariant derivative ∇ the geodesics of the connection $\bar{\nabla}$ are always described by an equation of the form

$$\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) = \alpha(m(t), \dot{m}(t)), \tag{12.1}$$

where $\alpha(m,X) = S_m(X,X)$ is a vector field on M that is quadratic in $X \in T_m M$ at any point $m \in M$.

For the Levi-Civitá connection on a complete Riemannian manifold the existence of a geodesic joining any two points m_0 and m_1 follows from the Hopf-Rinow theorem (Theorem 3.68). However this is not the case in general, even for a Riemannian connection with non-zero torsion: in [20, 85, 140] examples of Riemannian connections are presented for which this problem is not solvable (in particular, on a compact manifold, the two-dimensional torus).

We consider the general case of systems with discontinuous forces or forces with control, i.e., whose Newton law is described by differential inclusion (11.17) (see Section 11.7). The results for continuous single-valued forces (i.e., for equations of the form (11.2)) follow as simple corollaries. The interpretation of set-valued forces as forces with control allows us to investigate the so-called *controllability problem*, i.e., the problem of whether there exists a time-dependent control under which a trajectory of a mechanical system starting at a given point m_0 can reach another point m_1 of the configuration space.

We also consider the problem for constrained systems. In this case we look for a solution that connects a given point and a certain submanifold.

The method of investigation is based on the use of integral operators with parallel translation and the velocity hodograph equation (see Section 11.8).

Below, in Section 13.2, we apply the machinery developed in this Chapter to the investigation of geodesics of some connections on Lorentz manifolds.

The two-point boundary value problem for (11.17) and (11.2) with non-conjugate points has been investigated under various conditions more restrictive than ours. For equation (11.2) (i.e., for single-valued force fields) its solvability was shown by the author for continuous force fields in [88] (bounded case) and in [101] (for linear growth in X), by E. Yakovlev, e.g., in [232], for smooth force fields under some complicated conditions and by V. Ginzburg in [85] for smooth force fields with less than quadratic growth in X. The solvability of this problem for inclusion (11.17) has been demonstrated for set-valued force fields of several types (B. Gel'man and Y. Gliklikh [80], Y. Gliklikh and A. Obukhovskiĭ [124, 125], M. Kisielewicz [158], etc.) but only in the uniformly bounded case.

Here we follow our joint work with P. Zykov [131, 133]. Note that in [241] P. Zykov found some conditions for the solvability of the problem in cases where the right-hand sides have greater than quadratic growth in velocity (see Remarks 12.22 and 12.26).

12.2 Examples of Points that Cannot be Connected by a Trajectory

Example 12.1. Let $X = (x, y) \in \mathbb{R}^2$ and let $a > 0$ be a real number; denote the norm in $\mathbb{R}^2$ by $\|\cdot\|$. In $\mathbb{R}^2$ consider the following system of type (2):

$$\begin{cases} \ddot{x}(t) = -a\|\dot{X}\|\dot{y} \\ \ddot{y}(t) = a\|\dot{X}\|\dot{x} \end{cases}$$

with initial conditions $X(0) = 0$, $\dot{X}(0) = X_0$. Since here the vectors $\dot{X}$ and $\ddot{X}$ are orthogonal to each other along the solution, $\|\dot{X}\|$ is constant. Let $\|X_0\| = C$ and represent the vector X_0 in the form $X_0 = C(-\sin\varphi_0, \cos\varphi_0)$. Then the solution of the above-mentioned Cauchy problem takes the form $x(t) = \frac{1}{a}\cos(Cat + \varphi_0) - \frac{1}{a}\cos\varphi_0$, $y(t) = \frac{1}{a}\sin(Cat + \varphi_0) - \frac{1}{a}\sin\varphi_0$. Hence any solution is a circle of radius $\frac{1}{a}$ and it does not leave the disc of radius $\frac{2}{a}$ centered at the initial point. We should emphasize that the radius decreases as a increases.

Example 12.2. [90, 94]. Consider the mechanical system on the unit sphere S^2 in $\mathbb{R}^3$ with the force field $\bar{\alpha}(\bar{r}) = (-y, x, 0)$, where $\bar{r} = (x, y, z) \in S^2$. The motion of the system is given by the following equations in $\mathbb{R}^3$:

$$\ddot{\bar{r}} = \bar{\alpha}(\bar{r}) - 2\mathcal{K} \cdot \bar{r}$$

or, equivalently,

$$\begin{aligned} \ddot{x} &= -y - 2\mathcal{K} \cdot x, \\ \ddot{y} &= x - 2\mathcal{K} \cdot y, \\ \ddot{z} &= -2\mathcal{K} \cdot z. \end{aligned} \tag{12.2}$$

where

$$\mathcal{K} = \|\dot{\bar{r}}\|^2/2 = (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)/2$$

is the kinetic energy. To obtain these equations, one applies d'Alembert's principle (see, e.g., [190]) to the holonomic constraint $F(\bar{r}) = x^2 + y^2 + z^2$.

Denote the North and South Pole of the sphere by $N = (0,0,1)$ and $S = (0,0,-1)$, respectively. Let $\bar{r}(t) = \big(x(t), y(t), z(t)\big)$ be the trajectory of the system such that $\bar{r}(t_0) = S$ for some t_0 and $\dot{\bar{r}}(t_0) = V \neq 0$. Note that if $V = 0$, then $\bar{r}(t) \equiv S$. It is clear that $V \in T_S S^2$ must have the form $(X, Y, 0)$. We claim that the kinetic energy increases along $\bar{r}(t)$ until $\bar{r}$ hits the North or South Pole. By (12.2) we have

$$\dot{\mathcal{K}}\left(\bar{r}(t)\right) = \dot{x}y + \dot{y}x \quad \text{and} \quad \ddot{\mathcal{K}}\left(\bar{r}(t)\right) = x^2 + y^2.$$

Note also that $\dot{\mathcal{K}}(\bar{r}(t_0)) = 0$. This means that $\dot{\mathcal{K}}(\bar{r}(t)) > 0$, unless $\bar{r}(t) = S$ or $\bar{r}(t) = N$. In fact, the derivative $\dot{\mathcal{K}}(\bar{r}(t))$ is also increasing. Since $\dot{\mathcal{K}}(N) = \dot{\mathcal{K}}(S) = 0$, we have $\dot{\mathcal{K}}(\bar{r}(t)) \neq N$ for any $t > t_0$.

To clarify the geometric picture, consider the function $z(t) = z(\bar{r}(t))$. Let $t_1 > t_0$ be such that $\dot{z}(t_1) = 0$ and $z(t)$ is increasing on $[t_0, t_1]$. The last equation in (12.2) implies that $z(t_1) > 0$ and, as a consequence, $\ddot{z}(t_1) < 0$, i.e., $z(t_1)$ is a local maximum of $z(t)$. Since $\mathcal{K}$ is increasing along $\bar{r}(t)$, we see that $z(t_1) < 1$. In the same way, one may show that

$$\mathrm{sign} z(t_i) = (-1)^{i+1} \quad \text{and} \quad |z(t_i)| > |z(t_{i+1})|$$

for all points $t_1 < t_2 < \ldots$ such that $\dot{z}(t_i) = 0$.

Therefore the trajectory $\bar{r}(t)$ heads to the equator of S^2 and oscillates near it. In particular, the trajectory never reaches the point $N = (0,0,1)$.

Example 12.3. Let us replace the field $\bar{\alpha}$ in the system of Example 12.2 by the force field

$$\Omega\left(\bar{r}, \dot{\bar{r}}\right) = \frac{[\dot{\bar{r}}, \bar{r}]}{1 + \|\dot{\bar{r}}\|},$$

where $[\cdot,\cdot]$ is the vector product in $\mathbb{R}^3$. The equation of motion of the mechanical system is as follows:

$$\ddot{\bar{r}} = \Omega\left(\bar{r}, \dot{\bar{r}}\right) - 2\mathcal{K} \cdot \bar{r}. \tag{12.3}$$

A straightforward calculation shows that

$$\dot{\mathcal{K}} = \left(\Omega\left(\bar{r}, \dot{\bar{r}}\right), \dot{\bar{r}}\right) - 2\mathcal{K} \cdot \left(\bar{r}, \dot{\bar{r}}\right) = 0$$

along a solution of (12.3) (i.e., the force is always orthogonal to the velocity) and $\dot{\bar{b}} = 0$, where

$$\bar{b} = [\dot{\bar{r}}, \ddot{\bar{r}}] = -\frac{\|\dot{\bar{r}}\|^2 \cdot \bar{r}}{1 + \|\dot{\bar{r}}\|} - \|\dot{\bar{r}}\|^2 \cdot [\dot{\bar{r}}, \bar{r}] \, .$$

Therefore, the kinetic energy $\mathcal{K} = \|\dot{\bar{r}}\|^2 /2$ is constant along the trajectory $r(t)$ and $r(t)$ lies in the plane orthogonal to the constant vector $\bar{b}$. In other words, the trajectory is the circle $\left(\bar{r}(t), \bar{b}\right) = \text{const}$ on S^2. Assume that there is a trajectory passing through two antipodal points. Then it must be a great circle on S^2 and, therefore, $\left(\bar{r}(t), \bar{b}\right) = 0$.

Let α be the angle between $\bar{r}(t)$ and $\bar{b}$. A straightforward calculation (based on the explicit formulas for $\|\bar{b}\|$ and $(\bar{r}(t), \bar{b})$ and on the equality $\|\bar{r}(t)\| \equiv 1$) shows that

$$\cos\alpha = \frac{\phi\left(\|\dot{\bar{r}}\|\right)}{\|\dot{\bar{r}}\|^2},$$

where $\phi\colon [0, \infty) \to \mathbb{R}$ is a bounded function. Hence, $(\bar{r}(t), \bar{b})$ tends to zero as $\mathcal{K} \to \infty$, assuming non-zero values only. This means that there is no trajectory in the system passing through two antipodal points. Note also that any two points which are not antipodal can be connected by a trajectory with sufficiently high kinetic energy.

Example 12.4. Replace the force $\mathit{\Omega}\,(\bar{r}, \dot{\bar{r}})$ of the preceding example by the gyroscopic force $A(\bar{r}, \dot{\bar{r}}) = [\dot{\bar{r}}, \bar{r}]$. The equation of motion of the new system is

$$\ddot{\bar{r}} = [\dot{\bar{r}}, \bar{r}] - 2\mathcal{K}\dot{\bar{r}}.$$

The analysis of this example is quite similar to that of Example 2. First, we prove that $\dot{\mathcal{K}} = 0$ and $\dot{\bar{b}} = 0$, where $\bar{b} = [\dot{\bar{r}}, \ddot{\bar{r}}]$. This implies that the trajectory lies in the plane orthogonal to $\bar{b}$. If the trajectory were a great circle, so that $(\bar{r}, \bar{b}) = 0$, then this would give us the equality $[\bar{r}, \dot{\bar{r}}] = 0$, which is impossible.

The author is grateful to Evgenii I. Yakovlev for pointing out Example 12.4.

Example 12.5. Replace the force $\mathit{\Omega}\,(\bar{r}, \dot{\bar{r}})$ of the preceding example by

$$\alpha\,(\bar{r}, \dot{\bar{r}}) = [\bar{r}\,, \dot{\bar{r}}]\,\|\dot{\bar{r}}\|\,.$$

By d'Alembert's principle, as above, the equation of motion with a constraint takes the form: $\ddot{\bar{r}} = [\bar{r}\,, \dot{\bar{r}}]\,\|\dot{\bar{r}}\| - 2T\bar{r}$ where the kinetic energy $T = \frac{1}{2}\dot{\bar{r}}^2$. Since the acceleration is everywhere orthogonal to the velocity, it is obvious that $\dot{T} = 0$. Then, as above, we prove that $\dot{\bar{b}} = 0$ for $\bar{b} = [\dot{\bar{r}}, \ddot{\bar{r}}]$. Direct calculations yield $\dot{\bar{b}} = 0$. This means that any trajectory satisfies the relation $(\bar{b}, \bar{r}) = \text{const}$ (the parentheses denote the inner product in $\mathbb{R}^3$), i.e., it is a circle on the sphere that also lies in a plane orthogonal to the constant vector $\bar{b}$. Antipodal points are joined by a great circle, i.e., $(\bar{b}, \bar{r}) = 0$. From this we get the equality for the mixed product $(\bar{r}, \dot{\bar{r}}, \ddot{\bar{r}}) = 0$, which is impossible. Thus antipodal points on the sphere cannot be connected by a trajectory.

12.3 Existence of Solutions

In the examples of Section 12.2, the points which could not be connected by a trajectory were conjugate along all geodesics of the Levi-Civitá connection. In this section, we prove that if two points are not conjugate along a geodesic, then there exists a trajectory joining the points if the force has less than quadratic growth in velocity, or quadratic growth under some additional assumption. We consider the general case where the force field $\bar{\alpha}(t, m, X)$ is discontinuous or includes a control parameter (see Section 11.7.) Thus the trajectories are, in fact, solutions of the differential inclusion (11.17) with a set-valued vector force field $\mathfrak{a}(t, m, X)$ which can be either upper or lower semicontinuous and has convex closed values. This general result yields, as a simple corollary, the existence of such a trajectory for a mechanical system with continuous $\bar{\alpha}$ [88, 90]. The case of bounded force, which is a special case of force with sub-quadratic growth in velocity, has many specific consequences.

The requisite definitions from set-valued analysis can be found in Section 4.1.

We start with the following technical statement.

Lemma 12.6 *Let a real number δ satisfy the inequality $0 < \delta < \frac{\varepsilon}{(\varepsilon+C)^2}$. Then there exists a sufficiently small positive number φ such that $(\varepsilon t_1^{-1} - \varphi) > 0$ and the inequality $\delta((\varepsilon t_1^{-1} - \varphi) + C t_1^{-1})^2 < \varepsilon t_1^{-2} - \varphi t_1^{-1}$ holds.*

Proof. For δ satisfying the hypothesis of the Lemma we get $\delta(\varepsilon t_1^{-1} + C t_1^{-1})^2 < \varepsilon t_1^{-2}$. From the continuity of both sides of this inequality it follows that there exists a sufficiently small $\varphi > 0$ such that $(\varepsilon t_1^{-1} - \varphi) > 0$ and the inequality $\delta((\varepsilon t_1^{-1} - \varphi) + C t_1^{-1})^2 < (\varepsilon t_1^{-1} - \varphi) t_1^{-1} = \varepsilon t_1^{-2} - \varphi t_1^{-1}$ holds. □

For the remainder of this section, M is a complete Riemannian manifold and by $\|\cdot\|$ we denote the norm in a tangent space generated by the Riemannian metric. Introduce a norm on the set $\mathfrak{a}(t, m, X) \in T_m M$ by the usual formula $\|\mathfrak{a}(t, m, X)\| = \sup_{y \in \mathfrak{a}(t,m,X)} \|y\|$.

Definition 12.7. We say that $\mathfrak{a}(t, m, X)$ has *less than quadratic growth in* X if for any compact $\Theta \subset M$ and any finite interval $[0, l]$ the relation

$$\lim_{\|X\| \to \infty} \frac{\|\mathfrak{a}(t, m, X)\|}{\|X\|^2} = 0$$

holds uniformly in $t \in [0, l]$ and $m \in \Theta$.

Definition 12.8. We say that $\mathfrak{a}(t, m, X)$ has a *quadratic bound* in X if for any compact $\Theta \subset M$ and any finite interval $[0, l]$ the relation

$$\lim_{\|X\| \to \infty} \frac{\|\mathfrak{a}(t, m, X)\|}{\|X\|^2} = a(t, m)$$

holds uniformly in $t \in [0, l]$ and $m \in \Theta$, where $a(t, m) \geq 0$ is a real bounded function on $[0, l] \times \Theta$ that is not identical zero.

For the case of set-valued vector force fields we modify the definitions from Section 4.1 as follows:

Definition 12.9. We say that $\mathfrak{a}(t, m, X)$ satisfies the *upper Carathéodory condition* if:

1) for every $(m, X) \in TM$ the map $F(\cdot, m, X) : I \multimap T_m M$ is measurable;
2) for almost all $t \in I$ the map $F(t, \cdot, \cdot) : TM \multimap TM$ is upper semicontinuous.

Definition 12.10. Let $I = [0, l] \subset R$. The set-valued force field $\mathfrak{a} : I \times TM \multimap TM$ is said to be *almost lower semicontinuous* if there exists a countable sequence of disjoint compact sets $\{I_n\}$, $I_n \subset I$, such that:

(i) $I \setminus \cup_n I_n$ has measure zero;
(ii) the restriction of $\mathfrak{a}$ on each $I_n \times TM$ is lower semicontinuous.

Theorem 12.11 *Let $\mathfrak{a}(t, m, X)$ satisfy the upper Carathéodory condition, have convex closed bounded values and have less than quadratic growth in X. Let the points m_1 and m_0 be non-conjugate along a geodesic g of the Levi-Civitá connection. Then there exists a positive number $L(m_0, m_1, g)$ such that if $0 < t_1 < L(m_0, m_1, g)$ there exists a solution $m(t)$ of* (11.17) *for which $m(0) = m_0$ and $m(t_1) = m_1$.*

Proof. For a C^1-curve $\gamma(t) = \mathcal{S}v(t)$, $v(\cdot) \in C^0(I, T_{m_0}M)$, consider the set-valued vector field $\mathfrak{a}(t, \gamma(t), \dot{\gamma}(t))$. Denote by Γ the operator of parallel translation of vectors along $\gamma(\cdot)$ at the point $\gamma(0) = m_0$. Apply the operator Γ to all sets $\mathfrak{a}(t, \gamma(t), \dot{\gamma}(t))$ along $\gamma(\cdot)$. As a result, for any $v \in C^0(I, T_{m_0}M)$ we obtain a set-valued map $\Gamma \mathfrak{a} \mathcal{S} v : [0, l] \multimap T_{m_0}M$ that has convex values. It is shown in [125] that the map $\Gamma \mathfrak{a} \mathcal{S} : C^0([0, l], T_{m_0}M) \times [0, l] \multimap T_{m_0}M$ satisfies the upper Carathéodory condition. Denote by $\mathcal{P}\Gamma \mathfrak{a} \mathcal{S} v$ the set of all measurable selectors of $\Gamma \mathfrak{a} \mathcal{S} v : [0, l] \multimap T_{m_0}M$ (such selectors exist by [31]). Define on $C^0([0, t_1], T_{m_0}M)$ the set-valued operator $\int \mathcal{P}\Gamma \mathfrak{a} \mathcal{S}$ by the formula

$$\int \mathcal{P}\Gamma \mathfrak{a} \mathcal{S} v = \left\{ \int_0^t f(\tau) \mathrm{d}\tau \,\middle|\, f(\cdot) \in \mathcal{P}\Gamma \mathfrak{a} \mathcal{S} v \right\}.$$

Lemma 12.12 *The map $\int \mathcal{P}\Gamma \mathfrak{a} \mathcal{S}$ sends bounded subsets of $C^0(I, T_{m_0}M)$ to compact sets.*

Proof. Since the metric $\langle\cdot,\cdot\rangle$ is complete, for any ball $\mathcal{U}_K$ in $C^0(I, T_{m_0}M)$ the union of curves $\{(\gamma,\dot\gamma)|\gamma\in\mathcal{U}_K\}$ lies, by Lemma 3.53, in a compact subset of TM. Then for those curves $\|\dot\gamma(t)\|$ is uniformly bounded. Hence, since Definition 12.7 is fulfilled for $\mathfrak{a}$, all sets $\mathfrak{a}(t,\gamma,\dot\gamma)$, where $\gamma\in\mathcal{S}\mathcal{U}_K$, are uniformly bounded. As a consequence, since parallel translations preserve the norm, the sets $(\Gamma\mathfrak{a}\mathcal{S}v)(t)$ for $v\in\mathcal{U}_K$ are also uniformly bounded, and so are all their measurable selectors $\mathcal{P}\Gamma\mathfrak{a}\mathcal{S}v$. Thus, the continuous curves $u\in\bigcup_{v\in\mathcal{U}_K}(\int\mathcal{P}\Gamma\mathfrak{a}\mathcal{S})v$ are uniformly bounded and equicontinuous. The lemma follows. $\square$

Lemma 12.13 *The map $\int\mathcal{P}\Gamma\mathfrak{a}\mathcal{S}$ is upper semicontinuous and has convex values.*

Proof. It suffices to prove that the set-valued map $\int\mathcal{P}\Gamma_A\circ\mathcal{S}$ has a closed graph. In other words, that $v_k\to v_0$ and $u_k\to u_0$, where $u_k\in(\int\mathcal{P}\Gamma\mathfrak{a}\mathcal{S})v_k$ implies that $u_0\in(\int\mathcal{P}\Gamma\mathfrak{a}\mathcal{S})v_0$, i.e., $\dot u_0\in(\Gamma\mathfrak{a}\mathcal{S}v_0)(t)$ for almost all t. Since the map $\int\mathcal{P}\Gamma\mathfrak{a}\mathcal{S}$ sends bounded sets to compact sets, the map is upper semicontinuous provided that it has a closed graph (see [31]). Recall that the sets $(\Gamma\mathfrak{a}\mathcal{S}v_0)(t)$ are convex and the map $(\Gamma\mathfrak{a}\mathcal{S}v)(t)$ is upper semicontinuous in v and t. As a result, we have $\dot u_0\in(\Gamma\mathfrak{a}\mathcal{S}v_0)(t)$. $\square$

Consider the numbers ε and C from Lemma 3.48 constructed for the points m_0 and m_1 and the geodesic g. Let Ξ be the compact set from Lemma 3.51, and let $[0,l]$ be some interval. Choose a positive number $\delta<\frac{\varepsilon}{(\varepsilon+C)^2}$. Since $\mathfrak{a}$ satisfies Definition 12.7, one can easily see that there exists a $Q>0$ such that for $\|X\|\geq Q$ the inequality

$$\max_{(t,m)\in I\times\Xi}\|\mathfrak{a}(t,m,Y)\|<\delta\|X\|^2 \tag{12.4}$$

holds for all $\|Y\|<\|X\|$. For sufficiently small $t_1>0$ we get $t_1\in[0,l]$ and $t_1^{-1}\varepsilon-\varphi>Q$ where φ is as in Lemma 12.6. Let $L(m_0,m_1,g)$ be the upper bound of t_1 such that the above relations hold and let $0<t_1<L(m_0,m_1,g)$. For this t_1 denote by K the corresponding number $t_1^{-1}\varepsilon-\varphi$.

By construction, $t_1^{-1}\varepsilon>K$ and so by Lemma 3.50 the operator $\mathcal{B}(v)=\int\mathcal{P}\Gamma\mathfrak{a}\mathcal{S}(v+C_v)$ that sends $\mathcal{U}_K$ to $C^0([0,t_1],T_{m_0}M)$ is well-defined. Like $\int\mathcal{P}\Gamma\mathfrak{a}\mathcal{S}$, this operator is upper semicontinuous, has convex values and maps bounded sets from $C^0([0,t_1],T_{m_0}M)$ to compact sets.

For $v\in\mathcal{U}_K\subset C^0([0,t_1],T_{m_0}M)$, since parallel translation preserves the norm of a vector, from the construction of the operator $\mathcal{S}$, from (12.4) and from Lemma 12.6 it follows that

$$\begin{aligned}\left\|\mathfrak{a}\left(t,\mathcal{S}(v(t)+C_v),\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(v(t)+C_v)\right)\right\| &< \delta(t_1^{-1}\varepsilon-\varphi+Ct_1^{-1})^2\\ &< (t_1^{-2}\varepsilon-t_1^{-1}\varphi).\end{aligned}$$

Since parallel translation is norm-preserving, from the last inequality it follows that

$$\|\mathcal{Z}(v+C_v)\| = \left\| \int \mathcal{P}\Gamma\mathfrak{a}\mathcal{S}(v(\tau)+C_v) \right\|_{C^0([0,t_1],T_{m_0}M)} < (t_1^{-1}\varepsilon - \varphi) = K.$$

Thus $\mathcal{B}$ sends the ball $\mathfrak{U}_K$ into itself and from the Glicksberg-Ky Fan Theorem (see, e.g., [31]) it follows that $\mathcal{B}$ has a fixed point $u^* \in \mathfrak{U}_K$, i.e. $u^* \in \mathcal{B}u^*$. Let us show that $m(t) = \mathcal{S}(u^*(t) + C_{u^*})$ is the desired solution. By construction we have $m(0) = m_0$ and $m(t_1) = m_1$, $m(t)$ is a C^1-curve and $\dot{m}(t)$ is absolutely continuous. Note that $\dot{u}^*$ is a selector of $\Gamma\mathfrak{a}(t, \mathcal{S}(u^* + C_{u^*}), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(u^* + C_{u^*}))$ since u^* is a fixed point of $\mathcal{Z}$. In other words, the inclusion $\dot{u}^*(t) \in \Gamma\mathfrak{a}(t, \mathcal{S}(u^* + C_{u^*}), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(u^* + C_{u^*}))$ holds for all points t at which the derivative exists. Using the properties of the covariant derivative and the definition of u^*, one can show that $\dot{u}^*(t)$ is parallel to $\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t)$ along $m(\cdot)$ and $\Gamma\mathfrak{a}(t, \mathcal{S}(u^* + C_{u^*}), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(u^* + C_{u^*}))$ is parallel to $\mathfrak{a}(t, m(t), \dot{m}(t))$. Hence, $\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) \in \mathfrak{a}(t, m(t), \dot{m}(t))$. □

Theorem 12.14 *Let $\mathfrak{a}(t, m, X)$ be almost lower semicontinuous, have closed bounded values and have less than quadratic growth in X. Let the points m_1 and m_0 be non-conjugate along a geodesic g of the Levi-civitá connection. Then there exists a positive number $L(m_0, m_1, g)$ such that if $0 < t_1 < L(m_0, m_1, g)$ there exists a solution $m(t)$ of (11.17) for which $m(0) = m_0$ and $m(t_1) = m_1$.*

Proof. Here we use the same notation as in the proof of Theorem 12.11. Notice that from the condition of less than quadratic growth for $\mathfrak{a}$ it follows that for all $v \in C^0([0,l], T_{m_0}M)$ the curves from $\mathcal{P}\Gamma\mathfrak{a}\mathcal{S}v$ are integrable. Hence, the set-valued map $\mathcal{P}\Gamma\mathfrak{a}\mathcal{S}$ sends $C^0([0,l], T_{m_0}M)$ into $L^1(([0,l], \mathcal{A}, \mu), T_{m_0}M)$, where $\mathcal{A}$ is the Borel σ-algebra and μ is the normalized Lebesgue measure. Since $\mathfrak{a}$ is almost lower semicontinuous, by analogy with [155] one can easily show that $\mathcal{P}\Gamma\mathfrak{a}\mathcal{S} : C^0([0,l], T_{m_0}M) \to L^1(([0,l], \mathfrak{a}, \mu), T_{m_0}M)$ is lower semicontinuous and has decomposable values. Then by the Bressan-Colombo Theorem (Theorem 4.9) it has a continuous selector, which we denote by $p\Gamma\mathfrak{a}\mathcal{S}$.

Let Q, $L(m_0, m_1, g)$, $0 < t_1 < L(m_0, m_1, g)$ and K be as in the proof of Theorem 12.11. Then on the ball $\mathfrak{U}_K \subset C^0([0,t_1], T_{m_0}M)$ the operator

$$Bv = \int_0^t p\Gamma\mathfrak{a}\mathcal{S}(v(s) + C_v), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(v(s) + C_v))\mathrm{d}s : \mathfrak{U}_K \to C^0([0,t_1], T_{m_0}M)$$

is well-defined.

Lemma 12.15 *The mapping $B : C^0(I, T_{m_0}M) \to C^0(I, T_{m_0}M)$ is completely continuous.*

Proof. The fact that B sends bounded sets to compact sets is proved by analogy with the argument in the proof of Lemma 12.13.

From the properties of the operators $\mathcal{S}$ and Γ (see Section 3.2) it follows that the operator $B : C^0(I, T_{m_0}M) \to L^1((I, \mathcal{A}, \mu), T_{m_0}M)$ is continuous. Since C_v continuously depends on v (see Theorem 3.47), this means that the vector $\int_0^l p\Gamma\mathfrak{a}(s, \mathcal{S}(v(\cdot)+C_v), \frac{\mathrm{d}}{\mathrm{d}s}\mathcal{S}(v(\cdot)+C_v))\mathrm{d}s \in T_{m_0}M$ is continuous in $v \in C^0(I, T_{m_0}M)$. A very simple modification of the above argument shows that for any $t^* \in I$ the map sending $v(\cdot) \in C^0(I, T_{m_0}M)$ to the restriction of $p\Gamma\mathfrak{a}(t, \mathcal{S}(v(\cdot)+C_v), \frac{\mathrm{d}}{\mathrm{d}t}S(v(\cdot)+C_v))$ on $[0, t^*]$ is continuous as a map from $C^0(I, T_{m_0}M)$ to $L^1(([0,t^*], \mathcal{A}, \mu), T_{m_0}M)$, hence we obtain that the vector $\int_0^{t^*} p\Gamma\mathfrak{a}(s, \mathcal{S}(v(\cdot)+C_v), \frac{\mathrm{d}}{\mathrm{d}s}\mathcal{S}(v(\cdot)+C_v))\mathrm{d}s$ is jointly continuous in t and v for any $t^* \in I$. Thus for any $\varepsilon > 0$, $v \in C^0(I, T_{m_0}M)$ and $t^* \in I$ there exists a $\delta = \delta(\varepsilon, v, t^*) > 0$ such that if $\|v(\cdot) - v_1(\cdot)\|_{C^0(I,T_{m_0}M)} < \frac{1}{2}\delta$ and $|t - t'| < \frac{1}{2}\delta$, then

$$\left\| \int_0^t p\Gamma\mathfrak{a}(s, \mathcal{S}(v(\cdot)), \frac{\mathrm{d}}{\mathrm{d}s}S(v(\cdot)))\mathrm{d}s \right.$$
$$\left. - \int_0^{t'} p\Gamma\mathfrak{a}(s, \mathcal{S}(v_1(\cdot)), \frac{\mathrm{d}}{\mathrm{d}s}\mathcal{S}(v_1(\cdot)))\mathrm{d}s \right\|_{T_{m_0}M} < \varepsilon.$$

Since I is compact, for a given v we can find a unique $\delta = \delta(\varepsilon, v)$ for all $t \in I$. This completes the proof of continuity of B. □

Since parallel translation preserves the norm of a vector, from the construction of $\mathcal{S}$ for any $u \in \mathfrak{U}_K$ with given $\mathfrak{a}$ we get

$$\begin{aligned} \|Bv\| &= \left\| \int_0^t p\Gamma\mathfrak{a}(s, \mathcal{S}(v(s)+C_v), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(v(s)+C_v))\mathrm{d}s \right\|_{C^0([0,t_1],T_{m_0}M)} \\ &< (t_1^{-1}\varepsilon - \varphi) = K. \end{aligned}$$

Hence B sends $\mathfrak{U}_K$ into itself and hence, by the classical Schauder principle, it has a fixed point $u^* \in \mathfrak{U}_K$. Using the same argument as in the proof of Theorem 12.11, one can easily prove that $m(t) = \mathcal{S}(u^* + C_u^*)(t)$ is a solution of (11.17) such that $m(0) = m_0$ and $m(t_1) = m_1$. □

Theorem 12.16 *Let $\mathfrak{a}(t, m, X)$ either satisfy the upper Carathéodory condition and have convex closed bounded values or be almost lower semicontinuous and have closed bounded values. Let also $\mathfrak{a}(t, m, X)$ have a quadratic bound in X and the points m_1 and m_0 be non-conjugate along a certain geodesic g of the Levi-civitá connection. Assume in addition that for $t \in [0, l]$ and $m \in \Xi$, where $[0, l]$ is some interval and Ξ is the compact set from Lemma* 3.51*, for the function $a(t, m)$ from Definition 12.8 there exists a real number δ such that the estimate $a(t, m) < \delta < \frac{\varepsilon}{(\varepsilon+C)^2}$ holds. Then there exists a positive number $L(m_0, m_1, g)$ such that if $0 < t_1 < L(m_0, m_1, g)$ there exists a solution $m(t)$ of* (11.17) *for which $m(0) = m_0$ and $m(t_1) = m_1$.*

The proof of Theorem 12.16 follows the same argument as that for Theorems 12.11 and 12.14. The only modification here is that for $\mathfrak{a}$ with a quadratic bound in X we assume the existence of a δ such that $a(t, m) < \delta < \frac{\varepsilon}{(\varepsilon+C)^2}$

while in the proof of Theorems 12.11 and 12.14 an analogous δ is shown to exist for any $\mathfrak{a}$ with less than quadratic growth in X.

It is worth noting that if there is more than one geodesic along which m_0 and m_1 are not conjugate, then any of these geodesics can be used in the proof. Naturally, different geodesics can give rise to different solutions and constants L.

If a geodesic, along which m_0 and m_1 are not conjugate, is length minimizing, the constant C characterizes the Riemannian distance between these points. C and ε together provide certain characteristics of the Riemannian geometry on M in a neighborhood of m_0. Theorem 12.16 establishes an interrelation between C, ε and the quadratic bounds of (11.17), under which the two-point boundary value problem for non-conjugate points m_0 and m_1 is always solvable.

Note that the case of uniformly bounded force is a particular case of force with less than quadratic growth in velocity. Assume that the configuration space M is compact, the metric $\langle\cdot,\cdot\rangle$ has a non-negative sectional curvature and the force is uniformly bounded. Then there are no conjugate points on M. Recall that in the conditions of Theorems 12.11 and 12.14 there exists a constant $L > 0$ such that any two points can be connected by a trajectory $m(t)$ with $t \in [0, t_0]$ for any $t_0 < L$. If the force is bounded and M is flat and possibly non-compact, one may take $L = \infty$ (see Remark 3.52) In particular, one may take $L = \infty$ if M is a Euclidean space. This means that, in such M, the corresponding two-point boundary-value problem has a solution on any time interval.

Unlike the case of bounded forces, for force fields of less than quadratic growth in velocity (and consequently for those with a quadratic bound) even on flat configuration spaces a trajectory joining the points generally exists only on small time intervals. Nevertheless there is a subclass of forces with quadratic bound that has the following property: if a field and a pair of points satisfy the conditions of Theorem 12.16, there exists a trajectory joining the points on every finite time interval. This is the class satisfying the estimate

$$\|\mathfrak{a}(t, m, X)\| < a(t, m)\|X\|^2, \tag{12.5}$$

where $a(t, m) > 0$ is a continuous real-valued function on $I \times M$. Evidently the force satisfying (12.5) also satisfies the Definition 12.8 of a quadratic bound.

It should be pointed out that the existence of the above-mentioned solution on an arbitrary finite time interval was previously known for single-valued quadratic fields $\mathfrak{a}$ on manifolds that correspond to vector fields of geodesic sprays of connections on tangent bundles. In the latter case, applying a linear change of time along the solution on a given time interval, one obtains a solution on another time interval and by this method a solution on an arbitrary finite interval can be constructed. This approach cannot be extended to the general set-valued case with estimate (12.5).

We begin the proof of the above-mentioned property for inclusions satisfying (12.5) with two technical statements.

Lemma 12.17 *For* $0 < \delta < \frac{\varepsilon}{(\varepsilon+C)^2}$ *and for any* $t > 0$ *both roots* $K_{1,2}$ *of the equation* $\delta(Kt + Ct^{-1})^2 = K$ *are positive.*

Proof. Transform the equation $\delta(Kt + Ct^{-1})^2 = K$ into the form $(\delta t^2)K + (2C\delta - 1)K + C^2t^{-2}\delta = 0$. Its discriminant is equal to $D = 1 - 4C\delta$. This means that for $\delta < \frac{1}{4C}$ the roots are real and take the form $K_{1,2} = \frac{1-2C\delta \pm \sqrt{1-4C\delta}}{2\delta t^2}$. Since $(1 - 2C\delta) > \sqrt{1 - 4C\delta 2\delta t^2}$, we have $K_{1,2} > 0$. But, as pointed out in Remark 3.49, $\varepsilon < C$ and so $\frac{\varepsilon}{(\varepsilon+C)^2} < \frac{1}{4C}$. □

Lemma 12.18 *For* $0 < \delta < \frac{\varepsilon}{(\varepsilon+C)^2}$ *and for all* $t > 0$ *the inequality* $t^{-1}\varepsilon > \frac{1-2C\delta-\sqrt{1-4C\delta}}{2\delta t}$ *holds.*

Proof. In order to prove this statement, consider the following system

$$\begin{cases} \delta < \frac{1}{4C} \\ \frac{1-2C\delta-\sqrt{1-4C\delta}}{2\delta} < \varepsilon. \end{cases}$$

By means of elementary transformations, taking into account Remark 3.49, this system can be transformed into the following form

$$\begin{cases} \left[\begin{array}{l} \delta < \frac{\varepsilon}{\varepsilon^2+2C\varepsilon+C^2} \\ \delta \geq \frac{1}{2(\varepsilon+C)} \end{array}\right. \\ \delta < \frac{1}{4C}. \end{cases}$$

Since by Remark 3.49 $\varepsilon < C$, we obtain that $\delta < \frac{\varepsilon}{(\varepsilon+C)^2}$. □

Theorem 12.19 *Let* $\mathfrak{a}(t, m, X)$ *have convex closed bounded values satisfy the upper Carathéodory condition and the estimate* (12.5) *for some continuous function* $a(t, m) > 0$*. Let the points* m_1 *and* m_0 *be non-conjugate along a geodesic* $g(\cdot)$ *of the Levi-Civitá connection and let the estimate* $a(t, m) < \delta$ *hold on* $I \times \Xi$*, where the compact set* Ξ *is as in Lemma* 3.51 *and* $\delta > 0$ *satisfies the inequality* $\delta < \frac{\varepsilon}{(\varepsilon+C)^2}$*. Then for any* $t_1 > 0$*,* $t_1 \in I$ *there exists a solution* $m(t)$ *of the inclusion* (11.17) *for which* $m(0) = m_0$ *and* $m(t_1) = m_1$*.*

Proof. For a C^1-curve $\gamma(t) = \mathcal{S}v(t)$, $v(\cdot) \in C^0(I, T_{m_0}M)$, consider the set-valued vector field $\mathfrak{a}(t, \gamma(t), \dot{\gamma}(t))$. Denote by Γ the operator of parallel translation of vectors along $\gamma(\cdot)$ at the point $\gamma(0) = m_0$. Apply the operator Γ to all sets $\mathfrak{a}(t, \gamma(t), \dot{\gamma}(t))$ along $\gamma(\cdot)$. As a result for any $v \in C^0(I, T_{m_0}M)$ we obtain a set-valued map $\Gamma\mathfrak{a}\mathcal{S}v : [0, l] \multimap T_{m_0}M$ that has convex values. As in the proof of Theorem 12.11 the map $\Gamma\mathfrak{a}\mathcal{S} : C^0([0, l], T_{m_0}M) \times [0, l] \multimap T_{m_0}M$ satisfies the upper Carathéodory condition. Consider the operator $\int \mathcal{P}\Gamma\mathfrak{a}\mathcal{S}$ from the proof of Theorem 12.11. This operator is upper semicontinuous, has convex values and sends bounded sets from $C^0([0, t_1], T_{m_0}M)$ to compact sets.

Let $0 < t_1 < l$. Taking into account the hypotheses of the Theorem, we obtain from Lemma 12.17 that $K = \frac{1-2C\delta-\sqrt{1-4C\delta}}{2\delta t^2}$ is positive and from Lemma 12.18 that $t_1^{-1}\varepsilon > Kt_1$. Consider the ball $\mathcal{U}_{Kt_1}$ of radius Kt_1 centered at the origin in the Banach space $C^0([0,t_1], T_{m_0}M)$. Since $t_1^{-1}\varepsilon > Kt_1$, by Lemma 6.27 for any $v(\cdot) \in \mathcal{U}_{Kt_1}$ the vector C_v is well-defined. Thus we can introduce the operator $\mathcal{B}$ from the proof of Theorem 12.11. This operator is also upper semi-continuous, convex-valued and sends bounded sets from $C^0([0,t_1], T_{m_0}M)$ to compacts sets.

Since parallel translation preserves the norms of vectors, from the construction of $\mathcal{S}$ and from the hypothesis we derive that for any $v \in \mathcal{U}_{Kt_1}$ and $t \in [0,t_1]$ the estimate

$$\left\| \mathfrak{a}(t, S(v(t)+C_v), \frac{\mathrm{d}}{\mathrm{d}\tau} S(v(t)+C_v)) \right\| < \delta \|v(t)+C_v\|^2$$

holds. By construction $\delta\|v(t)+C_u\|^2 \le \delta(Kt_1 + Ct_1^{-1})^2 = K$. Since parallel translation is norm-preserving, for any curve $u(t) \in \mathcal{Z}v(t)$ and for any $t \in [0,t_1]$ the inequality $\|u(t)\| \le Kt \le Kt_1$ holds. Thus $\mathcal{B}$ sends the ball $\mathcal{U}_{Kt_1}$ into itself and from the Bohnenblust-Karlin fixed point theorem it follows that $\mathcal{B}$ has a fixed point $u^* \in \mathcal{U}_{Kt_1}$, i.e. $u^* \in \mathcal{B}u^*$. The fact that that $m(t) = \mathcal{S}(u^*(t) + C_{u^*})$ is the desired solution is proved by analogy with the proof of Theorem 12.11. $\square$

Theorem 12.20 *Let $\mathfrak{a}(t,m,X)$ be almost lower semicontinuous, have closed bounded values and satisfy (12.5) with a continuous function $a(t,m) > 0$. Let the points m_1 and m_0 be non-conjugate along a geodesic $g(\cdot)$ of the Levi-Civitá connection and let the estimate $a(t,m) < \delta$ hold on $I \times \Xi$, where the compact set Ξ is as in Lemma 3.51 and $\delta > 0$ satisfies the inequality $\delta < \frac{\varepsilon}{(\varepsilon+C)^2}$. Then for any $t_1 > 0$, $t_1 \in I$ there exists a solution $m(t)$ of the inclusion (11.17) for which $m(0) = m_0$ and $m(t_1) = m_1$.*

Proof. Here we use the same notation as in the proof of Theorem 12.19. From the hypothesis it follows that for all $v \in C^0([0,l], T_{m_0}M)$ the curves from $\mathcal{P}\Gamma\mathfrak{a}\mathcal{S}v$ are integrable. Hence the set-valued map $\mathcal{P}\Gamma\mathfrak{a}\mathcal{S}$ sends $C^0([0,l], T_{m_0}M)$ into $L^1(([0,l],\mathcal{A},\mu), T_{m_0}M)$, as in the proof of Theorem 12.14, and the operator

$$\mathcal{P}\Gamma\mathfrak{a}\mathcal{S} : C^0([0,l], T_{m_0}M) \multimap L^1(([0,l],\mathcal{A},\mu), T_{m_0}M)$$

is lower semicontinuous and has decomposable values. Then by the Bressan-Colombo Theorem (Theorem 4.9) it has a continuous selector, which we denote by $p\Gamma\mathfrak{a}\mathcal{S}$.

Let t_1 and K be as in the proof of Theorem 12.19. Then on the ball $\mathcal{U}_{Kt_1} \subset C^0([0,t_1], T_{m_0}M)$ the operator B from the proof of Theorem 12.14 is well-defined and is completely continuous. Since parallel translation preserves the norm of a vector, from the construction of $\mathcal{S}$, from Lemma 12.17 and from Lemma 12.18, for any $u \in \mathcal{U}_{Kt_1}$ with given $\mathfrak{a}$ we get

$$\begin{aligned}\|Bv\| &= \left\| \int_0^t p\Gamma\mathfrak{a}(s, \mathcal{S}(v(s)+C_v), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(v(s)+C_v))\mathrm{d}s \right\|_{C^0([0,t_1],T_{m_0}M)} \\ &\le \delta(Kt_1 + Ct_1^{-1})^2 t_1 = Kt_1.\end{aligned}$$

Hence the completely continuous operator B sends $\mathfrak{U}_{Kt_1}$ into itself and hence, by the classical Schauder principle, it has a fixed point $u^* \in \mathfrak{U}_{Kt_1}$. Using the same argument as in the proof of Theorem 12.19, one can easily prove that $m(t) = \mathcal{S}(u^* + C_u^*)(t)$ is a solution of (7.46) with $m(0) = m_0$ and $m(t_1) = m_1$. □

Consider a bounded Hausdorff continuous force field $\mathfrak{a}(t, m, X)$ with convex closed values on M, as above. Following Definition 11.50 we say that a trajectory $m(t)$ of the mechanical system with force $\mathfrak{a}(t, m, X)$ is *governed by extreme values of controlling force* if a.e. $\dot{m}(t)$ belongs to $\operatorname{Ext}\mathfrak{a}(t, m(t), \dot{m}(t))$ (see Definition 11.48), i.e. (11.18) is satisfied.

Theorem 12.21 *Assume that $\mathfrak{a}$ is Hausdorff continuous and has convex closed values, and let m_1 be non-conjugate to m_0 along at least one geodesic $g(\cdot)$ joining them.*

(i) *If $\mathfrak{a}$ has lower than quadratic growth in velocity, then there exists a positive number $L(m_0, m_1, g)$ such that if $0 < t_1 < L(m_0, m_1, g)$, there exists a trajectory, governed by extreme values of controlling force, for which $m(0) = m_0$ and $m(t_1) = m_1$.*

(ii) *If $\mathfrak{a}$ has a quadratic bound and $a(t, m)$ satisfies the hypothesis of Theorem 12.16, then the conclusion of* (i) *above holds.*

(iii) *If $\|\mathfrak{a}(t, m, X)\| < a(t, m)\|X\|$ and $a(t, m)$ satisfies the hypothesis of Theorem 12.20, the conclusion of* (i) *above holds for every $t_1 > 0$.*

Recall that the trajectories governed by extreme values of controlling force are described by the inclusion (11.18) and, in the case under consideration, the right-hand side of (11.18) is lower semicontinuous (see Lemma 11.49). Thus the Theorem follows from Theorems 12.14, 12.16 and 12.20.

Remark 12.22. We refer the reader to [241, Theorems 3.3 and 3.4] where the following generalization of Theorem 12.16 for right-hand sides with greater than quadratic growth in velocity is obtained. Let $\mathfrak{a}(t, m, X)$ either satisfy the upper Carathéodory condition and have convex closed bounded values or be almost lower semicontinuous and have closed bounded values. Let m_0 and m_1 be non-conjugate along some geodesic of the Levi-Civitá connection. Then ε and C, defined as above, and the compact set Ξ from Lemma 3.51 are well-defined. Suppose that for $t \in [0, t_1]$, for some $t_1 > 0$, and for $m \in \Xi$ the estimate $\|\mathfrak{a}(t, m, X)\| < f(\|X\|)$ holds where $f : [0, \infty) \to [0, \infty)$ is a function increasing on $[0, \infty]$. If $f(\varepsilon t_1^{-1} + Ct_1^{-1}) \le \varepsilon t_1^{-2}$, then there exists a solution $m(t)$ of (11.17) for which $m(0) = m_0$ and $m(t_1) = m_1$. The method of proof is a modification of the one given above in this Section.

12.4 Generalizations to Systems with Constraints

In this section, we show how to generalize the existence theorems of Section 12.3 to systems with constraints (see Section 11.6). In the framework of mechanics with constraints, it is more natural to consider the question of whether or not a submanifold transversal to the union of the least constrained geodesics leaving a specified point is accessible from that point. The author is grateful to Boris D. Gel'man for pointing out this problem.

The main technical trick here is the replacement of the operators $\mathcal{S}$ and Γ by their constraint analogs $\mathcal{S}^{\boldsymbol{\beta}}$ and $\Gamma^{\boldsymbol{\beta}}$ introduced in Section 11.8.2.

Let M be a complete Riemannian manifold equipped with a constraint $\boldsymbol{\beta}$. Let $m_0 \in M$. The exponential map $\exp^{\boldsymbol{\beta}}_{m_0} : \boldsymbol{\beta}_{m_0} \to M$ can be defined in the same manner as that for a manifold without constraint. Explicitly, for $X \in \boldsymbol{\beta}_{m_0}$, we set $\exp^{\boldsymbol{\beta}}_{m_0}(X) = \gamma_X(1)$, where $\gamma_X(t)$ is the least constrained geodesic with $\gamma_X(0) = m_0$ and $\dot\gamma_X(0) = X$. It is clear that $\exp^{\boldsymbol{\beta}}_{m_0}$ is a C^∞-smooth map.

Definition 12.23. A point $m_1 \in \exp^{\boldsymbol{\beta}}_{m_0}(\boldsymbol{\beta}_{m_0})$ is *not conjugate* to m_0 along the geodesic γ_X (where $\gamma_X(1) = m_1$) if the differential $\mathrm{d}\exp^{\boldsymbol{\beta}}_{m_0}$ has maximum rank at $X \in \boldsymbol{\beta}_{m_0}$.

In particular, this means that the image of $\exp^{\boldsymbol{\beta}}_{m_0}$ is a smooth submanifold in a neighborhood of m_1 if m_1 is not conjugate to m_0. Moreover, $\exp^{\boldsymbol{\beta}}_{m_0}$ is a diffeomorphism of a neighborhood of $X \in \beta_{m_0}$ onto a neighborhood of m_1 in the submanifold.

Assume that m_0 is not conjugate to m_1 along a least constrained geodesic γ_X. Let $N \subset M$, $m_1 \in N$, be a submanifold which is transversal to the image of $\exp^{\boldsymbol{\beta}}_{m_0}$. (In other words, the sum of the spaces $T_{m_0}N$ and $T_{m_0}\exp^{\beta}_{m_0}(\beta_{m_0})$ coincides with $T_{m_0}M$.) An example of such a manifold is an open neighborhood of m_1 in M.

Theorem 12.24 *Under the above-mentioned hypothesis for any $K > 0$ there exists a constant $\bar L(m_0, N, K, \gamma_X) > 0$ such that for $0 < t_1 < \bar L(m_0, N, K, \gamma_X)$ and for any continuous curve $u(t) \in \mathfrak{U}_K \subset C^0([0, t_1], \beta_{m_0})$, there exists a vector $C_u \in \beta_{m_0}$ satisfying the condition $\mathcal{S}^{\boldsymbol{\beta}}(u + C_u)(t_1) \in N$. Furthermore, C_u is unique in a neighborhood of $t_1^{-1}X \in \boldsymbol{\beta}_{m_0}$ and is continuous in u.*

The proof is quite similar to that of Theorem 3.47. The only extra argument needed is that the manifold N stays transversal to a C^1-small perturbation of the image of $\exp^{\beta}_{m_0} = \mathcal{S}^{\beta}(\cdot)(1)$. Below, in this section, we use $\hat\varepsilon$ and $\hat C$ in analogy with ε and C from Lemma 3.48.

Let $\mathfrak{a}$ be a set-valued vector field on M and $P_m : T_mM \to \boldsymbol{\beta}_m$ be the field of orthogonal projections. Consider the following set-valued analog of the constrained Newton law (11.13)

$$\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\,\dot m(t) \in P\mathfrak{a}\left(t, m(t), \dot m(t)\right), \tag{12.6}$$

where the constraint covariant derivative $\frac{\bar{\mathrm{D}}}{\mathrm{d}t}$ is defined in Section 11.6. The inclusion (12.6) arises in constraint analogs of the problems considered in Sections 11.7 and 12.3, for example, as a discontinuous force acting on the system, or where the image of $P\mathfrak{a}$ is formed by all possible values of the control force. It is easy to see that if $\mathfrak{a}$ has convex values, the sets $P\mathfrak{a}(t, m, X)$ are also convex and the set-valued vector field $P\mathfrak{a}$ is upper (lower) semicontinuous if $\mathfrak{a}$ is upper (lower) semicontinuous, respectively.

Since the norm of the orthogonal projector P equals 1, by replacing $\mathcal{S}$ and $\varGamma$ by $\mathcal{S}^{\boldsymbol{\beta}}$ and $\varGamma^{\boldsymbol{\beta}}$, respectively, by using the space $C^0(I, \boldsymbol{\beta}_{m_0})$ in place of $C^0(I, T_{m_0}M)$ and by using Theorem 12.24 rather than Theorem 3.47, one can easily prove the following analogs of the non-constrained theorems from Section 12.3.

Theorem 12.25 *Let $\mathfrak{a}(t, m, X)$ either satisfy the upper Carathéodory condition and have convex closed bounded values or be almost lower semicontinuous and have closed bounded values. Let the points m_1 and m_0 be non-conjugate along some least constraint geodesic g.*

(i) *If $\mathfrak{a}$ has less than quadratic growth in X, for any submanifold $N \ni m_1$ transversal to the image of $\exp^{\boldsymbol{\beta}}_{m_0}$, there exists a positive number $L(m_0, N, g)$ such that if $0 < t_1 < L(m_0, N, g)$, there exists an admissible solution $m(t)$ of (12.6) for which $m(0) = m_0$ and $m(t_1) \in N$.*

(ii) *Suppose that $\mathfrak{a}$ has a quadratic bound in X and in addition, for $t \in [0, l] = I$ and $m \in \varXi$, where $[0, l] = I$ is some interval and $\varXi$ is the compact set from Lemma 3.51, for the function $a(t, m)$ from Definition 12.8 there exists a real number δ such that the estimate $a(t, m) < \delta < \frac{\bar{\varepsilon}}{(\bar{\varepsilon}+\bar{C})^2}$ holds. Then for any submanifold $N \ni m_1$ transversal to to the image of $\exp^{\boldsymbol{\beta}}_{m_0}$, there exists a positive number $L(m_0, N, g)$ such that if $0 < t_1 < L(m_0, N, g)$, there exists an admissible solution $m(t)$ of (12.6) for which $m(0) = m_0$ and $m(t_1) \in N$.*

(iii) *Suppose that $\mathfrak{a}$ satisfies the estimate (12.5) with a continuous function $a(t, m) > 0$, that the estimate $a(t, m) < \delta$ holds on $I \times \varXi$, where the compact set $\varXi$ is as in Lemma 3.51 and that $\delta > 0$ satisfies the inequality $\delta < \frac{\bar{\varepsilon}}{(\bar{\varepsilon}+\bar{C})^2}$. Then for any submanifold $N \ni m_1$ transversal to the image of $\exp^{\boldsymbol{\beta}}_{m_0}$, and for any $t_1 > 0$, $t_1 \in I$, there exists an admissible solution $m(t)$ of the inclusion (12.6) for which $m(0) = m_0$ and $m(t_1) \in N$.*

Remark 12.26. In [241, Theorems 3.7 and 3.8] a generalization of Theorem 12.25 is obtained which is analogous to that of Theorem 12.16, mentioned in Remark 12.22.

Chapter 13
Some Problems on Lorentz Manifolds

13.1 Introduction to Relativity Theory

In this Section we give a brief introduction to some core notions of relativity theory. This material suffices to understand the language of relativity and to describe the relativistic problems discussed below. We are mainly interested in the constructions of general relativity, the formulae of special relativity arising as consequences of the latter. Since the exposition is intended for mathematicians, we present it axiomatically, starting from a basic set of postulates. Such an approach allows us to focus on the mathematical background of general relativity and on the physical interpretation of the mathematical developments. We do not touch on the physical background, however it is used to motivate the above-mentioned postulates. We refer the reader to [182, 200] for details.

13.1.1 Space-times

Recall that in Definition 1.50 we introduced the notion of a semi-Riemannian metric on a manifold M as a family of symmetric non-degenerate but not necessarily positive definite bilinear forms $\langle \cdot, \cdot \rangle_m$ on the tangent spaces T_mM, i.e., it is a symmetric non-degenerate $(0, 2)$-tensor field that, like its inverse, is called a *metric tensor* (see Remark 1.52). Recall also that in the definition of physical equivalence (see Sections 1.4 and 1.5), and in the construction of the Levi-Civitá connection, we used only the non-degeneracy of the metric tensor and did not use its positive definiteness. Thus for all semi-Riemannian metrics the notion of physical equivalence and the construction of the Levi-Civitá connection are well-defined. In particular, this means that the notions of covariant derivative, parallel translation and geodesic are well-defined on

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_13,

semi-Riemannian manifolds, and their properties are analogous to those on Riemannian manifolds.

Among the semi-Riemannian metrics, some play a special role in the mathematics of general relativity. Recall that for any non-degenerate symmetric bilinear form on a vector space there exists a basis (an orthonormal basis) such that the matrix (g_{ij}) of the metric tensor with respect to this basis is a diagonal matrix with diagonal entries equal to $+1$ or -1. For positive definite forms, of course, the diagonal entries are all equal to $+1$. For general non-degenerate forms the row of signs $+$ and $-$ corresponding to the values $+1$ and -1 appearing on the diagonal is called the *signature* of the form.

Definition 13.1. A semi-Riemannian metric is said to be a *Lorentz metric* if its signature is either $(-+\cdots+)$ or $(+-\cdots-)$, i.e., the above-mentioned diagonal form of (g_{ij}) includes only one -1 and all other elements of the diagonal are $+1$ (or only one diagonal element is $+1$ and all other elements are -1). A manifold on which a Lorentz metric is specified is called a *Lorentz manifold*.

Notice that the case where only one diagonal element is equal to -1 (and all others $+1$) can be transformed into the other case simply by multiplication by -1, and so both cases are equivalent. For a Riemannian metric, multiplication by -1 results in a negative definite (and hence, not Riemannian) metric, but for a semi-Riemannian metric such multiplication does not lead us out of the class of semi-Riemannian metrics. For the sake of simplicity we choose one of the above equivalent cases, namely the case with signature $(-+\cdots+)$, i.e., where the diagonal form has only one -1 (and all others $+1$). The other case leads to the same theory.

In contemporary physics the notions of space and time, considered separately in classical physics, are united into a common continuum called space-time.

Postulate 1 *The physical space-time of our universe is described mathematically as a* 4*-dimensional Lorentz manifold.*

Among all 4-dimensional Lorentz manifolds, those whose Levi-Civitá connection satisfies the so-called Einstein equation (derived by Einstein and Hilbert) expressing the connection via the distribution of matter in the universe, are called *space-times*. The Einstein equation is a complicated partial differential equation. We discuss it briefly in Section 13.1.6 below. In common with many other PDEs, the Einstein equation has many different solutions depending on the initial data, boundary conditions and other constraints. Such solutions describe the metric in different cases: locally (e.g., in a neighborhood of a certain star), globally, in special cases when certain negligible influences can be omitted, and so on.

Notation 13.2 *Everywhere in this section, by M^4 we denote a space-time under consideration, i.e., a* 4*-dimensional Lorentz manifold whose metric satisfies the Einstein equation.*

In all Sections dealing with general relativity below we use the Levi-Civitá connection of a Lorentz metric to determine the covariant derivative, parallel translation, geodesics, etc.

In a space-time we shall denote local coordinates by the symbols q^0, q^1, q^2 and q^3. The index 0 denotes the direction corresponding to the -1 entry in the signature (i.e., vectors tangent to this axis have negative squares) while the others correspond to $+1$ and so their tangent vectors have positive squares.

Notation 13.3 *In order to avoid confusion we shall use Greek letters for indices which take values from* 0 *to* 3 *and Latin indices for indices which range from* 1 *to* 3.

Thus, $g_{\alpha\beta}$ indicates that all coefficients of the Lorentz metric are under consideration while g_{ij} indicates that only those in the so-called "space directions" are under consideration (i.e., $i, j = 1, 2, 3$; here the word "space" has no physical meaning).

Let us present three examples of space-times. At present the global topological structure (i.e., the actual view) of physical space-time is not known. We can deal only with a neighborhood of our galaxy that resembles a part of a vector space (a chart). In the examples below we obtain various 4-dimensional manifolds that can be considered as models of physical reality. We are mainly interested in the metrics on space-times. These metrics will (according to the general notation of Section 1.5 and of this section) be denoted by $g_{\alpha\beta}\mathrm{d}q^\alpha \otimes \mathrm{d}q^\beta$.

Example 13.4. **Minkowski space.** M^4 is the standard vector space $\mathbb{R}^4$ and the coefficients of the Lorentz metric are of the form $g_{00} = -1$, $g_{ii} = 1$ for $i = 1, 2, 3$ and all others are zero. i.e., $\langle \cdot, \cdot \rangle^M = -\mathrm{d}q^0 \otimes \mathrm{d}q^0 + \sum_{i=1}^{3} \mathrm{d}q^i \otimes \mathrm{d}q^i$. This is the simplest Lorentz manifold playing the same role as the Euclidean space among Riemannian manifolds. It corresponds to the case when the gravitational influence is so small that it can be omitted and so it is used in special relativity where only the electromagnetic field is under consideration. Notice that for any Lorentz manifold M any tangent space T_mM has the structure of Minkowski space. This is an analog of the fact that on a Riemannian manifold every tangent space is equipped with the structure of Euclidean space.

Example 13.5. **Einstein–de Sitter space-time.** M^4 is the "upper" half-space of $\mathbb{R}^4$, i.e., $M^4 = \mathbb{R}^4_+ = \{(q^0, q^1, q^2, q^3) \in \mathbb{R}^4 | q^0 > 0\}$. The coefficients of the Lorentz metric are as follows: $g_{00} = -1$, $g_{ii} = (q^0)^{\frac{4}{3}}$, $i = 1, 2, 3$, and all others are zero, i.e., $\langle \cdot, \cdot \rangle^{ES} = -\mathrm{d}q^0 \otimes \mathrm{d}q^0 + (q^0)^{\frac{4}{3}} \sum_{i=1}^{3} \mathrm{d}q^i \otimes \mathrm{d}q^i$. The Einstein–de Sitter space-time is an example of a so-called *Friedman universe*. The level surfaces $q^0 = \text{const}$ have the structure of Euclidean spaces with an inner product $(q^0)^{\frac{4}{3}} \sum_{i=1}^{3} \mathrm{d}q^i \otimes \mathrm{d}q^i$ depending on the value of q^0 (which can be considered as an analog of the time variable).

Specify a value q^0 and consider two points $x_1 = (q^0, q_1^1, q_1^2, q_1^3)$ and $x_2 = (q^0, q_2^1, q_2^2, q_2^3)$ on its level surface. The Euclidean distance between x_1 and x_2 is

$$(q^0)^{\frac{2}{3}} \sqrt{(q_1^1 - q_2^1)^2 + (q_1^2 - q_2^2)^2 + (q_1^3 - q_2^3)^2}.$$

Consider another value $q_*^0 > q^0$. On its level surface the distance between "the same" points $x_1^* = (q_*^0, q_1^1, q_1^2, q_1^3)$ and $x_2^* = (q_*^0, q_2^1, q_2^2, q_2^3)$ is equal to

$$(q_*^0)^{\frac{2}{3}} \sqrt{(q_1^1 - q_2^1)^2 + (q_1^2 - q_2^2)^2 + (q_1^3 - q_2^3)^2}.$$

Since $q_*^0 > q^0$, the latter distance is greater than the former. This is a model for the so-called redshift, the experimental fact that all distances are increasing in time. Below in Remark 13.13 it will be shown that along the lines $q^1 = c^1$, $q^2 = c^2$, $q^3 = c^3$ where c^1, c^2, c^3 are constants, the variable q^0 has the physical interpretation of time. Notice that the points do not move in the level surfaces and that the growth of distance is a consequence of the variation of the metric.

By contemporary physical theory our Universe was born 15–20 billion years ago from a single point in an event called the Big Bang. Observe that all distances in level surfaces in Einstein-de Sitter space-time tend to zero as $q^0 \to 0$. This is a model of the Big Bang. In this model the Universe (i.e., the level surface) does not shrink into a single point as $q^0 \to 0$ but the metric becomes degenerate.

The Einstein-de Sitter space-time describes our Universe at an early stage after the Big Bang. There are other models, with a different metric, for the Universe at the present time (Robertson-Walker space-times).

Example 13.6. **Schwarzschild space-time.** Let $\mu > 0$. Consider two 2-dimensional manifolds: S^2, the 2-dimensional sphere, and the manifold $A^2 = \{(r,t) \in \mathbb{R}^2 \mid r \neq 2\mu\}$. Let $M^4 = A^2 \times S^2$. Since at any $m = (a,s) \in M^4$ the tangent space $T_m M^4 = T_a A^2 \oplus T_s S^2$, where $a \in A^2$ and $s \in S^2$, we can determine the following Lorentzian metric on M^4: $\langle \cdot,\cdot \rangle_{(a,s)}^{\mathrm{Sch}} = -(1 - \frac{2\mu}{r(a)})\mathrm{d}t \otimes \mathrm{d}t + (1 - \frac{2\mu}{r(a)})\mathrm{d}r \otimes \mathrm{d}r + \langle \cdot,\cdot \rangle_s$ where $r(a)$ is the value of coordinate r at $a \in A^2$ and $\langle \cdot,\cdot \rangle_s$ is the first fundamental form of S^2 at s.

Schwarzschild space-time is M^4 with the metric $\langle \cdot,\cdot \rangle^{\mathrm{Sch}}$. It describes physical space-time in a neighborhood of a black hole with mass $8\pi\mu$. Notice that M^4 is divided into two components: the points with $r(a) > 2\mu$ are said to be "outside the black hole", and the points with $r(a) < 2\mu$ are said to be "inside the black hole". We point out that t plays the role of q^0 outside the black hole (i.e., vectors in its direction have negative square) and r plays the role of q^0 inside the black hole (for some details, see Remark 13.13 below). In the popular scientific literature this fact is usually expressed by the words: *inside a black hole, space and time exchange places.*

13.1.2 World lines. The light cone. Proper time

As a space-time unites space and time into a common continuum, we need to change the notation usually found in the old physics. For example, in the old physics a 'point' typically means an element of space, a 'trajectory' means a line describing how the position in space depends on time, etc.

Points of a space-time are called *events*. The interpretation is "something held at a certain point of space at a certain moment of time". Instead of the word "trajectory" we shall use the term "world line". This is a certain smooth curve in the space-time that is interpreted as a 1-dimensional continuum of events. We do not specify a parameter in the world line from the very beginning, this will be done in a special way a little later. Thus a world line is a certain 1-dimensional manifold embedded (or immersed) into M^4.

In a tangent space T_mM^4 at any event m to the space-time M^4 one can find non-zero vectors whose squares are negative, positive and equal to zero.

Definition 13.7. The set of vectors in T_mM^4 whose square is equal to zero is called the *light cone*. Vectors lying on the light cone are also called *isotropic* or *light-like*.

A vector $X \in T_mM^4$ such that $X^2 < 0$ is called *time-like*. The set of time-like vectors is called the *interior* of the light cone. It is also said that a time-like vector lies inside the light cone.

A vector $X \in T_mM^4$ such that $X^2 > 0$ is called *space-like* or is said to lie outside the light cone.

The interior of the light cone has two non-intersecting components while the light cone itself, and the set of space-like vectors, are connected. The physical interpretation of the two components of the interior of the light cone is that one of them is directed into the future and the other one into the past. Of course, in order to obtain a consistent theory these directions must be coordinated at different points of space-time in a similar way to the notion of an orientation on a manifold (see Section 1.6). We shall not consider here the general construction of such orientation of directions, instead defining only a certain special case of it that is, in any case, general enough for our purposes.

Definition 13.8. A space-time M is called *time-oriented* if there exists a smooth vector field X on M such that $X^2 < 0$ at all events in M. The part of the light cone in the tangent space at any event that contains X is said to be "directed into the future".

Everywhere below we consider time-oriented space-times.

Definition 13.9. A world line for which all tangent vectors are time-like is called a *time-like world line*; a world line for which all tangent vectors are light-like is called a *light-like* or *isotropic world line*; a world line for which all tangent vectors are space-like is called a *space-like world line*.

Postulate 2 *Only time-like and light-like (isotropic) world lines have a sensible physical interpretation. The time-like world lines describe the "life" of objects moving slower than the speed of light while light-like world lines describe the "life" of objects traveling at light speed.*

Postulate 2 may be considered as an experimental fact. One may consider also space-like world lines whose tangent vectors are space-like, but they play an auxiliary role since they describe the motion of objects moving faster than light, which is forbidden in the present theory. From Postulate 2 it follows in particular that the type of world line of a material object cannot change. The physical interpretation of this fact will be clarified later.

Postulate 3 *Along any world line, corresponding to a material object, there is a vector field P, called the* 4-momentum, *that is tangent to the world line, directed into the future and which has constant square.*

For light-like world lines the above constant square is evidently zero while for time-like ones it is a negative number (if it were zero, P would be a zero vector, i.e. directionless, contradicting the fact that it should be directed into the future).

Definition 13.10. The positive real number $\mathbf{m}$ such that $\mathbf{m}^2 = -P^2$ is called the *mass* of the object corresponding to the time-like world line under consideration.

Thus along any time-like or light-like world line $m(\cdot)$ a unique (up to an additive constant) parameter η can be introduced such that $\frac{\mathrm{d}}{\mathrm{d}\eta}m(\eta) = P$.

Postulate 4 *If there is no influence of forces other than gravity, the time-like (or light-like) world line parametrized by η is a geodesic.*

This postulate means that the equation of the above world line is

$$\frac{\mathrm{D}}{\mathrm{d}\eta}\frac{\mathrm{d}}{\mathrm{d}\eta}m(\eta) = 0 \tag{13.1}$$

where $\frac{\mathrm{D}}{\mathrm{d}\eta}$ is the covariant derivative of the Levi-Civitá connection, and that $\frac{\mathrm{d}}{\mathrm{d}\eta}m(\eta) = P$ is parallel along $m(\eta)$.

Along a time-like world line we can construct another important vector field.

Definition 13.11. The vector field V which, at any event of a time-like world line $m(\cdot)$, is tangent to $m(\cdot)$, directed into the future and has square equal to -1, is called the 4-*velocity* of the world line $m(\cdot)$.

V exists along any time-like world line, i.e., its existence need not be postulated, unlike the existence of P. Immediately from the definition it follows that $P = \mathbf{m}V$.

As above, in the case of P there exists a unique (up to an additive constant) parameter τ along a time-like world line $m(\cdot)$ such that $\frac{\mathrm{d}}{\mathrm{d}\tau}m(\tau) = V$.

Definition 13.12. The above parameter τ is called the *proper time* of the time-like world line.

The physical interpretation of proper time is that it is the time that is shown by the watch of an observer whose world line is under consideration. Every observer has its own proper time, i.e., we all live according to our own time.

One can easily show that a time-like world line that is geodesic with respect to the parameter η described above (i.e., such that no force besides gravitation has an influence on it) also satisfies the geodesic equation

$$\frac{\mathrm{D}}{\mathrm{d}\tau}V = 0. \tag{13.2}$$

Remark 13.13. In Examples 13.4 and 13.5, along all lines of the form $q^1 = c^1$, $q^2 = c^2$, $q^3 = c^3$ where c^1, c^2, c^3 are constants, the variable q^0 is a proper time. Indeed, such a curve is described in coordinates as (q^0, c^1, c^2, c^3) so that its derivative in q^0 takes the form $(1, 0, 0, 0)$. Substituting this vector into the metrics of the above-mentioned examples one easily finds that its square is -1.

In Example 13.6 the situation is not so simple. Let $s^* \in S^2$. It is obvious that the curve $(r = \mathrm{const}, t, s^*) \in M^4$ outside the black hole and the curve $(r, t = \mathrm{const}, s^*) \in M^4$ inside black hole are time-like: the squares of their derivatives in t and r, respectively, are negative, but generally speaking are not equal to -1. So, changing the parameters, we can find the proper times along those world lines. Thus, the proper time flows along the former lines outside and along the latter lines inside the black hole. Hence, the space and time variables are exchanged inside the black hole, as was claimed in the example above.

13.1.3 Reference frames and 3-dimensional notions

In this section we explain how the above-mentioned 4-dimensional objects correspond to the 3-dimensional reality around us. Without this interpretation, the content of the previous sections would be nothing more than an abstract mathematical construction, while in fact it is a mathematical model of physics.

Let $m(\tau)$ be a time-like world line describing the evolution in the space-time M of an observer where τ is the proper time. Let $m = m(0)$ be an event in the world line and construct in T_mM a basis e_0, e_1, e_2, e_3 such that e_0 is the 4-velocity of our observer and e_1, e_2, e_3 are space-like vectors with square $+1$, orthogonal on M to each other and to e_0 with respect to the Lorentzian metric (such a basis is called an orthonormal basis).

Definition 13.14. The above basis is called a *reference frame* of the observer at m.

Remark 13.15. The reference frame described in Definition 13.14 is a very particular case of a general notion used in contemporary physics, but it is convenient for our exposition. We refer the reader, e.g., to [182, 199, 200] for a detailed discussion. Below in Section 13.2 we deal with a special type of reference frame suggested by A. Poltorak [196, 197, 198].

The linear span of e_1, e_2, e_3 is a 3-dimensional subspace in T_mM^4, orthogonal to the 4-velocity of the observer, that is interpreted as the space of 3-dimensional velocities of physical objects around the observer at m. Vectors from this subspace of T_mM^4 are called 3-*vectors*. For any 3-vector we can find a unique geodesic starting from m in its direction. The surface filled by such geodesics is interpreted as the set of events perceived by the observer as synchronous with the event m in his (or her) world line.

Notice that for two different observers at the same event m their synchronous surfaces are different since they depend on (i.e., are orthogonal to) the 4-velocity of the observer.

Remark 13.16. (Past and Future Domains) Unlike the synchronous surface, which depends on the 4-velocity (and hence on the reference frame) of the observer, the notions of "past" and "future" are "absolute", i.e., they depend only on the event and so they are the same for all observers located at the same event. These notions are introduced as follows.

Specify an event $m \in M^4$. We say that an event m_1 belongs to the *future* for m if there exists a time-like or light-like world line that starts from m and as τ (or η) increases, eventually reaches m_1 for some value of the parameter. If such a world line is time-like, we say that m_1 belongs to the *proper future* of m.

The notions of "past" and "proper past" are introduced analogously. Note that "proper future" and "proper past" are open domains in M^4 while "past" and "future" are closed, i.e., they are the closures of the "proper future" and the "proper past", respectively.

There are simple examples describing the following phenomenon: If m_2 belongs neither to "future" nor to "past" for m_1 (and so vice versa), there exists an observer for whom those events are synchronous, an observer for whom m_1 happens earlier than m_2 and an observer for whom m_2 happens earlier than m_1.

Let $V \in T_mM$ be the 4-velocity of some object. Represent V as a pair $V = (\dot{q}^0, \bar{V})$ where $\dot{q}^0$ is collinear with e_0 and $\bar{V}$ belongs to the linear span of e_1, e_2, e_3. $\bar{V}$ is interpreted as an infinitesimal motion in space and $\dot{q}^0$ is interpreted as an infinitesimal increment of time. Thus from the usual physical ideology we get the following:

Definition 13.17. $v = \frac{\bar{V}}{\dot{q}^0}$ is called the 3-*velocity corresponding to the* 4-*velocity* V with respect to the above reference frame.

In the same manner we can define the 3-velocity of light. This means that for any light-like vector $X \in T_m M$ we consider its decomposition $X = (X^0, \bar{X})$, where X^0 is collinear with e_0 and $\bar{X}$ belong to the linear span of e_1, e_2, e_3, and then define the 3-velocity of light as $\frac{\bar{X}}{X^0}$.

Proposition 13.18 *For every observer the norm of the* 3-*velocity of light (i.e., light speed) is equal to* 1.

Proof. Without loss of generality we can suppose that $\bar{X}$ is collinear with e_1 (if this is not the case we simply rotate the triple e_1, e_2, e_3 around e_0). Thus the coordinate presentation of X in this basis is $X = (X^0, X^1, 0, 0)$, i.e., $\bar{X} = (X^1, 0, 0)$. On the one hand $X^2 = 0$ and on the other hand, in our case, $X^2 = \langle X, X \rangle = -(X^0)^2 + (X^1)^2$. Thus $X^0 = X^1$ and $\|\frac{\bar{X}}{X^0}\| = \frac{X^1}{X^0} = 1$. $\square$

Following Proposition 13.18, we shall be working in a system of units where the speed of light c is equal to 1.

Now let us turn back to time-like vectors. Of course the observer can see only 3-velocities. If a v with norm less than 1 is given, then it is possible to recover V since it is the unique time-like vector with square -1 such that $v = \frac{\bar{V}}{\dot{q}^0}$. To find V notice that the square of the vector $(1, v)$ is equal to $-1+v^2$ so that the square of $(\frac{1}{\sqrt{1-v^2}}, \frac{v}{\sqrt{1-v^2}})$ is -1. Since $\frac{v}{\sqrt{1-v^2}}$ divided by $\frac{1}{\sqrt{1-v^2}}$ gives v, we have $V = (\frac{1}{\sqrt{1-v^2}}, \frac{v}{\sqrt{1-v^2}})$.

Since $P = \mathbf{m}V$ we obtain $P = (\frac{\mathbf{m}}{\sqrt{1-v^2}}, \frac{\mathbf{m}v}{\sqrt{1-v^2}})$. The 3-vector $p = \frac{\mathbf{m}v}{\sqrt{1-v^2}}$ is interpreted as 3-momentum. Since v is much less that c (i.e., less than 1 in our system of units), the denominator is very close to 1 and, by ignoring this negligible difference, the above formula turns into the usual definition of momentum one finds in high school physics.

In order to understand the physical interpretation of $\frac{\mathbf{m}}{\sqrt{1-v^2}}$ let us find its Taylor expansion in v at a neighborhood of $v = 0$. This expansion has no odd terms and we get:

$$\frac{\mathbf{m}}{\sqrt{1-v^2}} = \mathbf{m} + \frac{\mathbf{m}v^2}{2} + \dots$$

Further terms are negligible, so they can be omitted. $\frac{\mathbf{m}v^2}{2}$ is the kinetic energy. The quantity $\mathbf{m}$ is interpreted as the internal energy of the object with mass $\mathbf{m}$. It is well-known as $E = \mathbf{m}c^2$ (Einstein's famous formula of internal energy) and takes the above form only because $c = 1$ in our system of units.

Thus $\frac{\mathbf{m}}{\sqrt{1-v^2}}$ has the physical interpretation of total energy, a type of energy previously unknown in classical physics.

Notice that as $v \to 1$ (i.e., as the speed of an object approaches light speed) both the 3-momentum and the total energy tend to infinity.

Remark 13.19. Sometimes $\mathbf{m}$ is called the *rest mass* and another sort of mass $\mathbf{m}_r = \frac{\mathbf{m}}{\sqrt{1-v^2}}$, called *relativistic mass*, is introduced. In this notation $p = \mathbf{m}_r v$ and one says that the relativistic mass (and the 3-momentum) tends to infinity as $v \to 1$. From the mathematical point of view (a view shared by many physicists) since this approach is equivalent to that presented above, there is no reason to introduce the notion of relativistic mass.

If a world line $m(\tau)$ is not a geodesic, then it describes the "life" of an observer under the action of the so-called 4-force F and the equation of this world line takes the form

$$\frac{\mathrm{D}}{\mathrm{d}\tau} P = F(m(\tau), V). \tag{13.3}$$

Note that since $\langle V, V\rangle = -1$ (constant), $\frac{\mathrm{d}}{\mathrm{d}\tau}\langle V, V\rangle = 2\langle \frac{\mathrm{D}}{\mathrm{d}\tau} V, V\rangle = 0$. From this it follows that $\frac{\mathrm{D}}{\mathrm{d}\tau} P = \mathbf{m}\frac{\mathrm{D}}{\mathrm{d}\tau} V$ is orthogonal to V.

Remark 13.20. The 4-force is always orthogonal to the 4-velocity and so, in particular, it necessarily depends on the 4-velocity.

The 4-force F can be represented in the same manner as V and P, i.e., as the pair $(f^0, \bar{f})$. The 3-vector $\bar{f}$ is interpreted as 3-force, the physical force that can be measured by the observer. In order to understand the meaning of f^0 recall that $\langle F, V\rangle = 0$ (see above). Hence $-f^0\frac{\mathbf{m}}{\sqrt{1-v^2}} + \bar{f}\cdot\frac{\mathbf{m}v}{\sqrt{1-v^2}} = 0$ where the dot denotes the inner product in the linear span of e_1, e_2, e_3 (recall that it is positive definite since e_1, e_2 and e_3 are space-like). Thus $f^0 = \bar{f}\cdot v$ is the power of the 3-force $\bar{f}$.

13.1.4 Some consequences

The parameter of velocity and hyperbolic trigonometry

The coordinate expansion of a vector $X \in T_m M$ at some $m \in M$ with respect to e_0, e_1, e_2, e_3 takes the form (X^0, X^1, X^2, X^3). Consider the subspace consisting of vectors with $X^2 = X^3 = 0$. This is a 2-dimensional vector space with (Minkowski) inner product $\langle X, Y\rangle = -X^0Y^0 + X^1Y^1$. An analog of the Euclidean unit circle is the set of vectors in this subspace whose square is -1. Of course this set is not a circle but the hyperbola $-(X^0)^2 + (X^1)^2 = -1$ (equivalently, $(X^0)^2 - (X^1)^2 = 1$).

The length of an arc in the above hyperbola, starting at $(1, 0)$, is an analog of the angle, and the abscissa and ordinate of the end point of the arc are analogs of cosine and sine, respectively.

In order to investigate these analogs, represent the hyperbola in parametric form $X(\theta) = (X^0(\theta), X^1(\theta))$ and consider its derivative $\dot{X}(\theta) = (\dot{X}^0(\theta), \dot{X}^1(\theta))$. Since $\langle X(\theta), X(\theta)\rangle = -1$, we have

$$\langle X(\theta), \dot{X}(\theta)\rangle = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\theta}\langle X(\theta), X(\theta)\rangle = 0$$

and so $\dot{X}(\theta)$ is orthogonal to $X(\theta)$, i.e., $\dot{X}(\theta)$ is collinear to $(X^1(\theta), X^0(\theta))$: indeed, $\langle (X^0(\theta), X^1(\theta)), (X^1(\theta), X^0(\theta))\rangle = -X^0(\theta)X^1(\theta) + X^1(\theta)X^0(\theta) = 0$.

Without loss of generality we may suppose that $\dot{X}(\theta) = (X^1(\theta), X^0(\theta))$. Thus, in particular, $\dot{X}(\theta)^2 = -(X^1(\theta))^2 + (X^0(\theta))^2 = -X(\theta)^2 = 1$ so that, on the one hand, the vector $\dot{X}(\theta)$ is space-like and on the other hand the parameter θ is the length of the arc since the norm of the derivative with respect to this parameter is 1.

Definition 13.21. The parameter θ introduced above is called the *parameter of velocity* on the hyperbola $X^2 = -1$.

Now let us try to find the coordinates $X^0(\theta)$ and $X^1(\theta)$ of a point in the hyperbola, i.e., the analogs of cosine and sine. Introduce the variables $Y(\theta) = X^0(\theta) + X^1(\theta)$ and $Z(\theta) = X^0(\theta) - X^1(\theta)$. Since $\dot{X}(\theta) = (X^1(\theta), X^0(\theta))$, we have $\dot{X}^0(\theta) = X^1(\theta)$ and $\dot{X}^1(\theta) = X^0(\theta)$. Thus $\dot{Y}(\theta) = Y(\theta)$ and $\dot{Z}(\theta) = -Z(\theta)$. These are linear differential equations whose solutions with initial conditions $Y(0) = 1$ and $Z(0) = 1$ (corresponding to the initial point $(1, 0)$ of the hyperbola) are $Y(\theta) = e^{\theta}$ and $Z(\theta) = e^{-\theta}$, respectively. Then $X^0(\theta) = \frac{1}{2}(Y(\theta) + Z(\theta)) = \frac{e^{\theta}+e^{-\theta}}{2} = \cosh\theta$, the hyperbolic cosine of θ, and $X^1(\theta) = \frac{1}{2}(Y(\theta) - Z(\theta)) = \frac{e^{\theta}-e^{-\theta}}{2} = \sinh\theta$, the hyperbolic sine of θ.

Now the 3-velocity v, corresponding to $V = (X^0(\theta), X^1(\theta))$ belonging to the unit hyperbola (i.e., it is some 4-velocity), is represented in the form $v = \frac{X^1(\theta)}{X^0(\theta)} = \frac{\sinh\theta}{\cosh\theta} = \tanh\theta$, the hyperbolic tangent of θ. This is why we call θ the parameter of velocity (see Definition 13.21).

Composition of velocities

Consider the reference frame of an observer as described in Section 13.1.3. Suppose that, for example, a brick with 4-velocity V_b is in flight near the observer. Then the 3-velocity v_b of the brick with respect to the observer can be calculated as the hyperbolic tangent of the parameter of velocity θ_1 between e^0 and V_b (see above) in the reference frame of the observer.

Suppose in addition that an ant with 4-velocity V_a is creeping along the brick in the same direction that the brick is passing the observer. The 3-velocity v_{ab} of the ant with respect to the brick is the hyperbolic tangent of the parameter of velocity θ_2 between V_b and V_a in the reference frame of the brick.

Since the brick and the ant are moving in the same direction, in classical Newtonian physics the vector of the ant, with respect to the observer, would be given by the vector in this common direction with norm equal to the sum of the norms of the brick vector, relative to the observer, and the ant vector, relative to the brick. This is not the case in Relativity Theory.

Let us suppose that the direction of e_1 coincides with the common direction of the motion of the brick and the ant. The parameter of velocity for V_a in the reference frame of the observer is $\theta_1 + \theta_2$. It then follows from the above constructions that $v_a = \tanh(\theta_1 + \theta_2)$. One can easily find the following formula for the hyperbolic tangent of a sum:

$$\tanh(\theta_1 + \theta_2) = \frac{\tanh\theta_1 + \tanh\theta_2}{1 + \tanh\theta_1 \tanh\theta_2}$$

(compare this with the formula for the usual tangent of a sum). Thus we get that

$$v_a = \frac{v_b + v_{ab}}{1 + v_b v_{ab}}.$$

This is the well-known relativistic rule for the composition of velocities. Notice that if both v_b and v_{ab} are negligible in comparison with 1 (the light speed in our system of units), this formula turns into the familiar formula for the composition of velocities of classical physics.

Proposition 13.22 *The speed of light does not depend on the speed of the light source.*

Proof. Let us replace the ant of the above example by a beam of light. This means that V_a is replaced by a light-like vector X and that its 3-velocity in the reference frame of the brick is 1. By the above formula for the composition of velocities we get that the 3-velocity of light with respect to the observer is $\frac{v_b+1}{1+v_b} = 1$. □

The assertion of Proposition 13.22 was the starting point of Einstein's relativity theory.

Lorentz transformations

In Euclidean space there are linear operators A which satisfy $Ax \cdot Ay = x \cdot y$ for any vectors x, y, where $\cdot$ is the inner product, i.e. the operator does not change the inner product of any pair of vectors. Such operators are said to be *orthogonal*. By physical reasons we suppose those operators to be orientation preserving. In 2-dimensional Euclidean space the form of the matrix of an orthogonal operator (a rotation) with respect to an orthonormal basis is well-known:

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$$

where φ is the angle of rotation.

In Minkowski space there are analogous operators which leave the Lorentzian inner product unchanged.

Definition 13.23. An operator L in Minkowski space such that for any vectors X, Y the relation $\langle AX, AY\rangle = \langle X, Y\rangle$ holds, where $\langle\cdot,\cdot\rangle$ is the Lorentzian inner product, is called a *Lorentz transformation*.

In particular, if we have two observers at an event of a Lorentzian manifold, a Lorentz transformation sends the reference frame of one observer into that of another. In this case we can restrict ourselves to the 2-dimensional subspace spanned by the 4-velocities of the observers. The basis (a "part" of the reference frame) corresponding to each observer in this subspace consists of the 4-velocity and the space-like vector orthogonal to the 4-velocity and having square $+1$. This basis is orthonormal with respect to the Lorentzian inner product.

By physical reasons we consider Lorentz transformations preserving both standard and time orientations. They form the so-called proper orthochronous Lorentz group (see [72]).

Denote by θ the parameter of velocity between e_0 in some reference frame and its image Le_0 under the Lorentz transformation L. It is well-known that the matrix of L in the corresponding 2-dimensional subspace with respect to an orthonormal basis takes the form:

$$\begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix}.$$

Denote by A the reference frame of an observer. Thus the coordinates $(X^{0'}, X^{1'})$ of the vector X with respect to the reference frame $L(A)$ are expressed via the coordinates (X^0, X^1) of the same vector with respect to A by the formulae $X^{0'} = X^0\cosh\theta + X^1\sinh\theta$ and $X^{1'} = X^0\sinh\theta + X^1\cosh\theta$. Recall that in hyperbolic trigonometry we have the relation $\cosh^2\theta - \sinh^2\theta = 1$ (easily derived from the definitions). So we can divide the right-hand sides of the above expressions by $\sqrt{\cosh^2\theta - \sinh^2\theta} = 1$. Canceling $\cosh\theta$ and taking into account that the 3-velocity v of $L(A)$ with respect to A is the hyperbolic tangent $v = \tanh\theta$, we get

$$X^{0'} = \frac{X^0 + vX^1}{\sqrt{1-v^2}}, \qquad X^{1'} = \frac{vX^0 + X^1}{\sqrt{1-v^2}}.$$

This is the *standard form* of a Lorentz transformation. The formulae describe the differences between the time and space components of any 4-vector in different reference frames.

The twin paradox

One of the most well-known consequences of relativity theory is the so-called twin paradox. Suppose that following the birth of twins one twin is placed in a spaceship that leaves Earth moving at a speed very close to the speed of light while the other twin remains on the planet. When the traveling twin returns to Earth he (or she) is much younger than the Earth-bound twin.

The twin paradox is explained as follows. Introduce the functional of proper time on the set of time-like world lines by analogy with the functionals mentioned in Section 11.4. The formula for the variation of proper time with fixed end-points is practically the same as that for the length on Riemannian manifolds and the proof of this formula is similar to the proof of Theorem 11.11. Imitating the proof of Theorem 11.11, one can prove that the geodesics, and only the geodesics, are extremals with fixed end-points of proper time. But unlike the distance in Section 11.4, the geodesics on a Lorentz manifold attain the maximum proper time among all time-like world lines close to the extremal (while, in contrast, in a Riemannian manifold the geodesics attain minimum length).

The remainder of the argument is as follows. Since there are no forces except gravitation acting on Earth, its world line is a geodesic (see Postulate 4), while the world line of the spaceship is not a geodesic since some other forces act on it. Consider the world lines of both twins intersecting at the two events: the departure of the spaceship from Earth and its return. Since the world line of the Earth-bound twin coincides with that of the Earth, it is a geodesic, while the world line of the traveling twin is not a geodesic. Thus the interval of proper time between the above events is shorter along the world line of the traveling twin and so he (or she) is younger.

13.1.5 The electromagnetic field

Let M be a Riemannian or semi-Riemannian manifold and $\bar{J}$ be a vector field on M.

Definition 13.24. The system of equations

$$\begin{aligned} \mathrm{d}F &= 0 \\ \delta F &= \tilde{J} \end{aligned} \tag{13.4}$$

where F is a 2-form and $\tilde{J}$ is the 1-form physically equivalent to $\bar{J}$, is known as *Maxwell's equations*. F is called the *electromagnetic field* corresponding to the current density $\bar{J}$.

The world line of a charged particle with charge $\mathfrak{e}$ that is moving in the presence of an electromagnetic field F is described by the so-called *Lorentz equation*

$$\frac{\mathrm{D}}{\mathrm{d}\tau} P = \mathfrak{e}\, \overline{(V \rfloor F)}. \tag{13.5}$$

On the right-hand side of (13.5) there is a 4-force (see Section 13.1.3) in which: $V \rfloor F$ is the interior product of the 4-velocity V and the 2-form of the electromagnetic field F (thus the result is a 1-form) and $\mathfrak{e}\,\overline{(V \rfloor F)}$ is the vector physically equivalent to the 1-form multiplied by the scalar $\mathfrak{e}$ (with the corresponding sign).

If $M = \mathbb{R}^n$ then, as stated above, any closed form is exact. Hence, from the first equation of (13.4) it follows that $F = \mathrm{d}A$ where A is some 1-form. On an arbitrary manifold M such A may not exist and usually one assumes that F is exact (i.e., $F = \mathrm{d}A$ for some A). The 1-form A is called the *vector potential* of the electromagnetic field F. Such A is not unique; both A and $A + \mathrm{d}\lambda$, where λ is an arbitrary smooth function, lead to the same electromagnetic field $F = \mathrm{d}A$ since by Theorem 1.67 $\mathrm{d}\mathrm{d}\lambda = 0$.

In the following, we work in Minkowski space (see Example 13.4). In this space equations (13.4) take on a familiar special form.

To make our notation consistent with the usual one for electrodynamics, we replace the symbol q^0 by t. The coordinate system (t, q^1, q^2, q^3) on Minkowski space determines a certain observer for whom the "zero" coordinate axis is the world line and so t is the proper time. Thus all our constructions are made in the reference frame of this observer.

Recall that physical equivalence with respect to an orthonormal frame (see formulae (1.20) and (1.21)) leaves invariant the absolute value of the components. But since in the Lorentz case $g_{00} = g^{00} = -1$ and $g_{ii} = g^{ii} = 1$, the sign of the "zero" coordinate under a transition to a physically equivalent object is changed while the other signs remain unchanged.

The 1-form of the 4-current density $\tilde{J}$ in the above reference frame has the coordinate decomposition $\tilde{J} = \rho \mathrm{d}t + \bar{j}_i \mathrm{d}q^i$. The physically equivalent 4-vector of the current density has the form $\bar{J} = -\rho \frac{\partial}{\partial t} + j^i \frac{\partial}{\partial q^i}$ where $j_j = j^i$. The function ρ is called the "density of electric charge" and the 3-vector $\bar{j} = (j^1, j^2, j^3)$ is called "three-dimensional current density" (where everything is in the reference frame of our observer).

In our reference frame the 2-form F has the coordinate decomposition $F = E_i \mathrm{d}t \wedge \mathrm{d}q^i + B_1 \mathrm{d}q^2 \wedge \mathrm{d}q^3 - B_2 \mathrm{d}q^1 \wedge \mathrm{d}q^3 + B_3 \mathrm{d}q^1 \wedge \mathrm{d}q^2$. Consider the vectors $E = (E_1, E_2, E_3)$ and $B = (B_1, B_2, B_3)$ composed from the components of F. Under a Lorentz transformation of Minkowski space these vectors transform as ordinary 3-dimensional vectors. E is called the *electric field strength* and B is called the *magnetic field strength*.

The vector-potential A in our reference frame takes the form $A = \varphi \mathrm{d}q^0 + A_j \mathrm{d}q^j$ where φ is called the *potential* of the electric field and $\tilde{A} = (A_1, A_2, A_3)$ is called the *vector potential* of the magnetic field. Then

$$
\begin{aligned}
F &= \mathrm{d}A \\
&= \frac{\partial \varphi}{\partial q^\alpha} \mathrm{d}q^\alpha \wedge \mathrm{d}t + \frac{\partial A_j}{\partial q^\alpha} \mathrm{d}q^\alpha \wedge \mathrm{d}q^j \qquad (13.6) \\
&= -\frac{\partial \varphi}{\partial q^i} \mathrm{d}t \wedge \mathrm{d}q^i + \frac{\partial A_i}{\partial t} \mathrm{d}t \wedge \mathrm{d}q^i + \left(\frac{\partial A_1}{\partial q^2} - \frac{\partial A_2}{\partial q^1} \right) \mathrm{d}q^1 \wedge \mathrm{d}q^2 \\
&\quad + \left(\frac{\partial A_2}{\partial q^3} - \frac{\partial A_3}{\partial q^2} \right) \mathrm{d}q^2 \wedge \mathrm{d}q^3 - \left(\frac{\partial A_1}{\partial q^3} - \frac{\partial A_3}{\partial q^1} \right) \mathrm{d}q^1 \wedge \mathrm{d}q^3.
\end{aligned}
$$

From (13.6) we obtain that $E = -\text{grad}\,\varphi + \frac{\partial \tilde{A}}{\partial t}$ and $B = \text{rot}\,\tilde{A}$. These relations between A and F are commonly used in classical electrodynamics.

By formula (1.30) we find that $\mathrm{d}F = \frac{\partial E_i}{\partial q^\alpha}\mathrm{d}q^\alpha \wedge \mathrm{d}t \wedge \mathrm{d}q^i + \frac{\partial B_1}{\partial q^\alpha}\mathrm{d}q^\alpha \wedge \mathrm{d}q^2 \wedge \mathrm{d}q^3 - \frac{\partial B_2}{\partial q^\alpha}\mathrm{d}q^\alpha \wedge \mathrm{d}q^1 \wedge \mathrm{d}q^3 + \frac{\partial B_3}{\partial q^\alpha}\mathrm{d}q^\alpha \wedge \mathrm{d}q^1 \wedge \mathrm{d}q^2$. Recall that since the exterior product is skew-symmetric, it equals zero if at least two factors are equal to each other. Then from the first equation of (13.4) we obtain that $(\frac{\partial B_1}{\partial q^1} + \frac{\partial B_2}{\partial q^2} + \frac{\partial B_3}{\partial q^3})\mathrm{d}q^1 \wedge \mathrm{d}q^2 \wedge \mathrm{d}q^3 = 0$. This means that div $B = 0$.

Another consequence of (13.4) is that $(\frac{\partial E_1}{\partial q^2} - \frac{\partial E_2}{\partial q^1} + \frac{\partial B_3}{\partial t})\mathrm{d}t \wedge \mathrm{d}q^1 \wedge \mathrm{d}q^2 = 0$ and analogous relations hold for the other coordinates. One can easily see that this yields $\text{rot}\, E = -\frac{\partial B}{\partial t}$.

Now let us turn to the calculation of δF. The Riemannian volume form on Minkowski space is $\mathrm{d}t \wedge \mathrm{d}q^1 \wedge \mathrm{d}q^2 \wedge \mathrm{d}q^3$. On the other hand, by the above-mentioned rules of physical equivalence the 2-vector $\bar{F}$, physically equivalent to F, takes the form $\bar{F} = -E^i \frac{\partial}{\partial t} \wedge \frac{\partial}{\partial q^i} + B^1 \frac{\partial}{\partial q^2} \wedge \frac{\partial}{\partial q^3} - B^2 \frac{\partial}{\partial q^1} \wedge \frac{\partial}{\partial q^3} + B^3 \frac{\partial}{\partial q^1} \wedge \frac{\partial}{\partial q^2}$, $E^i = E_i$, $B^i = B_i$. Then $*F = B_i \mathrm{d}t \wedge \mathrm{d}q^i - E_1 \mathrm{d}q^2 \wedge \mathrm{d}q^3 + E_2 \mathrm{d}q^1 \wedge \mathrm{d}q^3 - E_3 \mathrm{d}q^1 \wedge \mathrm{d}q^2$ (cf. the above formula for F in coordinates). Using the definition of $\delta = *^{-1}\mathrm{d}*$ (see Definition 1.71) and the second equation of (13.4), by analogy with the above calculations we obtain that div $E = \rho$ and rot $B = \bar{j} + \frac{\partial E}{\partial t}$.

The four equalities derived above:

$$\begin{array}{ll} \text{div}\, E = \rho, & \qquad \text{rot}\, E = -\frac{\partial B}{\partial t}, \\ \text{div}\, B = 0, & \qquad \text{rot}\, B = \bar{j} + \frac{\partial E}{\partial t} \end{array}$$

yield the form of Maxwell's equations that is usually found in classical electrodynamics.

13.1.6 Gravitational fields

In this section we briefly describe the Einstein equation of a gravitational field and discuss some its features. Details can be found, e.g., in [182, 200].

By Ric we denote the Ricci tensor of the Levi-Civitá connection of a Lorentz metric $\langle \cdot, \cdot \rangle$ and by S the scalar curvature (see Section 2.5).

Definition 13.25. A $(0, 2)$-tensor

$$G(X, Y) = \text{Ric}(X, Y) - \frac{1}{2} S \cdot \langle X, Y \rangle$$

is called an *Einstein tensor*.

Space-time which is outside the influence of large masses and physical fields is said to be "empty". The Einstein equation in "empty" space-time has the form

$$G = 0. \tag{13.7}$$

It is not hard to show that (13.7) is equivalent to $\mathrm{Ric} = 0$ (see [182, 200]).

Thus outside the influence of large masses and physical (non-gravitational) fields space-time has zero Ricci curvature. Taking into account the material of Section 2.5 one can see that "empty" space-time is not necessarily flat. In particular there may exist solutions of (13.7) which are "gravitational waves". We remark that gravitational waves have not been directly observed.

The condition of "emptiness" is fulfilled by the inter-planetary space of the solar system. Specify a reference frame in the solar system in which the planetary speeds are small relative to the speed of light. Let the metric satisfy (13.7), be independent of time in this reference frame and have small enough curvature in M^4. Under these assumptions it is shown that the equation of a time-like geodesic in this reference frame is well-approximated by Newton's classical law of gravitation (for details, see, e.g., [51]).

In the general case the Einstein equation takes the form

$$G = T, \tag{13.8}$$

where the T on the right-hand side is the so-called *stress-energy tensor*.

The equation (13.8) continues to make mathematical sense if T is any $(0, 2)$-tensor with the same features as G.

From a physical point of view the notion of a stress-energy tensor is much more complicated. Such a tensor must be assigned to every type of matter in order to describe the features of the latter (for details, see [182, 200]). We present two examples of stress-energy tensors, one for a beam of particles and another for an electromagnetic field.

Example 13.26. For a beam of particles with density η its stress-energy tensor equals $\eta\tilde{P}\otimes\tilde{P}$ where $\tilde{P}$ is the 1-form physically equivalent to the 4-momentum P.

Example 13.27. Let F be an electromagnetic field (see Section 13.1.5). The components of F, and of the tensors physically equivalent to it, are denoted by $F_{\alpha\beta}$, F^{α}_{β} and $F^{\alpha\beta}$. Then the stress-energy tensor of F has the components

$$T_{\alpha\beta} = \frac{1}{4\pi}\left(F_{\alpha\mu}F^{\mu}_{\beta} - \frac{1}{4}g_{\alpha\beta}F^{\mu\nu}F_{\mu\nu}\right).$$

The behavior of a particle in a gravitational field is described by the geodesic equations (13.1) and (13.2) (in the absence of forces other than gravitation) or by equation (13.3) if other forces are present. Thus for those equations we need to know the Levi-Civitá connection of the Lorentz metric rather than the metric itself and so the connection plays the role of the field strength. It is clear that Einstein equations (13.7) and (13.8) are equations with respect to the Christoffel symbols of the Levi-Civitá connection.

Comparing formulae (2.33) and (2.35) for Christoffel symbols with formula (13.6) for the description of the electromagnetic field via its 4-potential, we see that that the coefficients of the metric play the role of the potential for the field strength.

13.2 A Two-Point Boundary Value Problem on a Lorentz Manifold Arising in A. Poltorak's Concept of Reference Frame

13.2.1 Discussion of the problem

In [196, 197] A. Poltorak suggested a concept in which a reference frame in general relativity is defined as a certain smooth manifold with a connection. In the most simple cases this is Minkowski space (see Example 13.4) with its natural flat connection but in more complicated cases some more general manifolds and connections may also appear.

In the reference frame, the gravitational field is described as a $(1,2)$-tensor G (see Theorem 2.36) that on any pair of vector fields X and Y takes the value

$$G(X,Y) = \nabla_X Y - \bar{\nabla}_X Y,$$

where $\bar{\nabla}$ is the covariant derivative of the Levi-Civitá connection of the Lorentz metric while ∇ is the covariant derivative of the connection in the reference frame. Denote by $\frac{\mathrm{D}}{\mathrm{d}\tau}$ the covariant derivative of the connection in the reference frame along a given world line with respect to some parameter τ. Then the geodesic $m(\tau)$ of the Levi-Civitá connection in M (a world line in the absence of all force fields except gravitation) is described in the reference frame by the equation

$$\frac{\mathrm{D}}{\mathrm{d}\tau} m'(\tau) = G_{m(\tau)}(m'(\tau), m'(\tau)) \tag{13.9}$$

where $G_m(X,Y)$ is the value of $G(X,Y)$ at point m (cf. equation (12.1)). Notice that the right-hand side of (13.9) is quadratic in the velocity $m'(\tau)$.

We refer the reader to [196, 197] for more details on Poltorak's concept and for a physical interpretation of the covariant derivative of a connection in the reference frame, of the tensor G and several other objects associated with it. The subsequent development of this idea can be found in [198].

We suggest a version of the concept where the manifold of the reference frame is the tangent space $T_m M^4$ at an event $m \in M^4$ and where a Lorentz-orthonormal basis e_α, where $\alpha = 0,1,2,3$, is specified (the time-like vector e_0 is the observer's 4-velocity). We assume that this reference frame is valid in a neighborhood $\mathcal{O}$ of the origin in $T_m M^4$, which is identified with a neighborhood $\mathcal{U}$ of the event m by the exponential map of the Levi-Civitá connection of the Lorentz metric (the normal chart).

We deal with two choices of connection on the manifold T_mM^4. In the first one, we consider on T_mM^4 its natural flat connection of Minkowski space (the main case considered by A. Poltorak). In the second case, we involve a Riemannian connection of a certain (positive definite) Riemannian metric on T_mM^4. This case is motivated by a natural development of the idea yielding Euclidean models in quantum field theory. Observe that the above-mentioned Riemannian connection may not be the Levi-Civitá connection, and that a non-zero torsion connection compatible with the metric is also allowed. In principle, this also allows one to consider electromagnetic interactions.

For the two reference frames mentioned above, we investigate the question of whether it is possible to connect two events m_0 and m_1 in M^4 by a time-like geodesic if they are connected in the reference frame by a geodesic of the corresponding connection whose initial vector is time-like, i.e., lies inside the light cone in the space $T_{m_0}M^4$. This question can be interpreted as follows: does the event m_1 belong to the proper future of the event m_0 if this is true in the reference frame? (See Remark 13.16.) The fact that this is not always so is illustrated by the following example:

Example 13.28. For the sake of simplicity we deal with a 2-dimensional space-time. Consider the manifold $\mathbb{R}^2_+ = \{(q^0, q^1) \in \mathbb{R}^2 | q^0 > 0\}$ with the Lorentz metric $g = -\mathrm{d}q^0 \otimes \mathrm{d}q^0 + (q^0)^{4/3}\mathrm{d}q^1 \otimes \mathrm{d}q^1$. This is a 2-dimensional Lorentz submanifold in the Einstein-de Sitter space-time (see Example 13.5). The flat Minkowski connection of the metric $h = -\mathrm{d}q^0 \otimes \mathrm{d}q^0 + \mathrm{d}q^1 \otimes \mathrm{d}q^1$ plays the role of a connection in the reference frame while g is the metric transferred into the reference frame from the space-time by the exponential map as described above.

Let $m_0 = (a_0, b_0) \in \mathbb{R}^2_+$ be an event. The proper future I_h of m_0 with respect to the flat Minkowski connection is the interior of the light cone at m_0, i.e., the events from its future are located between the lines $(a_0 + t, b_0 + (a_0^{-2/3})t)$ and $(a_0 + t, b_0 - (a_0^{-2/3})t)$ where $t > 0$. On the other hand, the proper future I_g of m_0 with respect to the Levi-Civitá connection of g consists of the events located between the curves $(a_0 + t, b_0 + 3\sqrt[3]{a_0 + t})$ and $(a_0 + t, b_0 - 3\sqrt[3]{a_0 + t})$, $t > 0$. One can easily see that the closure $\bar{I}_g$ (simply called "the future"), except the event m_0, is a proper subset of the open set I_h. This means that if an event $m_1 \in I_h$ is "far enough" from m_0 and "close enough" to the boundary of the light cone I_h, it may not lie in I_g, hence it may not be connected to m_0 by a time-like geodesic of the Levi-Civitá connection of g. But by construction m_1 is connected to m_0 by a time-like geodesic of the flat Minkowski connection (its initial vector is time-like).

If m_1 is conjugate with m_0 along all geodesics joining them in a reference frame, it may be impossible to connect m_0 and m_1 by a geodesic of another connection (in particular, of the connection in the space-time; see general examples of this sort in Section 12.2).

For the two reference frames described above we find geometric conditions that in each case guarantee a positive answer to the aforementioned question.

The conditions take the form of a certain interrelation between the tensor G and some geometric characteristics measuring in particular "how far" m_1 is from m_0 and "how close" m_1 is to the boundary of the "proper future" of m_0 and to conjugate points (if they exist) in the reference frame. For this investigation we use the machinery developed in Chapter 12.

13.2.2 The reference frame with flat connection

In this section we investigate the reference frame of the first type mentioned above, i.e., the manifold of the reference frame is $T_m M^4$ with a basis e_α, where $\alpha = 0, 1, 2, 3$, such that the time-like vector e_0 is the 4-velocity of some observer, and the connection in $T_m M^4$ is the flat connection of the Minkowski space.

In this case, it is convenient to regard $\mathcal{O}$ as a domain in a linear space on which there are given the Lorentz metric, the tensor G and other objects described above. Here we can make use of the familiar facts of linear algebra. In particular, the tangent space $T_{\bar{m}} M^4$ at $\bar{m} \in \mathcal{O}$ can be identified with $T_m M^4$ by a translation and for any $\bar{m} \in \mathcal{O}$ we may consider the light cone in $T_{\bar{m}} M^4$ (generated by the Lorentz metric tensor at $\bar{m}$) as lying in $T_m M^4$ but depending on $\bar{m}$.

Geodesics in $T_m M^4$ with respect to the flat connection are straight lines. Thus, in this reference frame, the question under consideration takes the following form: *is it possible to connect the events m_0 and m_1 on M^4 by a time-like geodesic if they are connected by the straight line $a(\tau)$ in $\mathcal{O}$ so that $a(0) = m_0$, $a(T) = m_1$ and which lies inside the light cone in $T_{m_0} M^4$*? Here τ is a parameter that can be, say, the proper time on M^4 or the natural parameter in the reference frame, etc. In this case the fact that the straight line $a(\tau)$ belongs to the light cone in $T_{m_0} M^4$ is equivalent to the fact that the vector of the derivative $a'(0) = \frac{\mathrm{d}}{\mathrm{d}\tau} a(\tau)_{|\tau=0}$ lies inside that cone, as postulated.

Since the covariant derivative with respect to the flat connection coincides in this case with the ordinary derivative in $T_m M^4$, equation (1) takes the form

$$\frac{\mathrm{d}}{\mathrm{d}\tau} m'(\tau) = G_m(m', m'). \tag{13.10}$$

Thus the main problem is reduced to the two-point boundary value problem for (13.10). Since the right-hand side of (13.10) has quadratic growth in velocity, for some pairs of points the two-point problem may have no solutions (see Sections 12.1 and 12.2).

Recall that the tangent space $T_m M^4$ to the Lorentz manifold M^4 has the natural structure of a Minkowski space whose inner product is the metric tensor of M^4 at the event m. Since the specified basis e_α, where $\alpha = 0, 1, 2, 3$, is Lorentz-orthonormal, the Minkowski inner product of $X = X^\alpha e_\alpha$ and $Y = Y^\alpha e_\alpha$ has the form

$$X \cdot Y = -X^0 Y^0 + X_i Y^i,$$

where $X_i = X^i$ for $i = 1, 2, 3$. Introduce a Euclidean inner product in $T_m M^4$ by changing the sign of the time-like summand, i.e., by setting

$$(X, Y) = X^0 Y^0 + X_i Y^i.$$

Hereafter in this section all norms and distances are determined with respect to the latter inner product.

By a linear change of time introduce a parameter s along $a(\cdot)$ such that for the line $\tilde{a}(s)$ obtained from $a(\tau)$ we get $\tilde{a}(0) = m_0$ and $\tilde{a}(1) = m_1$. Consider the Banach space $C^0([0,1], T_m M^4)$ of continuous curves in $T_m M^4$ with the usual supremum norm.

Lemma 13.29 *There exists a sufficiently small real number $\varepsilon > 0$ such that, for any curve $\tilde{v}(s)$ from the ball $\mathfrak{U}_\varepsilon \subset C^0([0,1], T_m M^4)$ of radius ε centered at the origin, there exists a vector $\tilde{C}_{\tilde{v}} \in T_m M^4$ belonging to a bounded neighborhood of the vector $\tilde{a}'(0) = \frac{\mathrm{d}}{\mathrm{d}s}\tilde{a}(s)_{|s=0}$ and such that $\tilde{C}_{\tilde{v}}$ lies inside the light cone of the space $T_{m_0} M^4$ and the curve $m_0 + \int_0^s (\tilde{v}(t) + \tilde{C}_{\tilde{v}})\mathrm{d}t$ takes the value m_1 at $s = 1$. The vector $\tilde{C}_{\tilde{v}}$ continuously depends on $\tilde{v}(\cdot)$ and $\|\tilde{C}_{\tilde{v}}\| < C$ for any curve $\tilde{v} \in \mathfrak{U}_\varepsilon$ for some $C > 0$.*

Proof. By explicit integration one can easily prove that $C_{\tilde{v}}$ such that $m_0 + \int_0^1 (\tilde{v}(t) + \tilde{C}_{\tilde{v}})\mathrm{d}t = m_1$ exists and is continuous in $\tilde{v}$. Then by continuity, from the fact that the vector $\tilde{a}'(0)$ lies inside the light cone of the space $T_{m_0} M^4$, it follows that for a perturbation $\tilde{v}(\cdot)$ sufficiently small with respect to the norm, the vector $\tilde{C}_{\tilde{v}}$ also lies inside the same light cone. Take for C the upper bound of the set of norms of the vectors $\tilde{C}_{\tilde{v}}$ from the above-mentioned bounded neighborhood of $\frac{\mathrm{d}}{\mathrm{d}s}\tilde{a}(s)_{|s=0}$. □

C is an estimate of the Euclidean distance between m_0 and m_1.

We turn back to the parametrization of the line $a(\cdot)$ by the parameter τ. Consider the Banach space $C^0([0,T], T_m M^4)$.

Lemma 13.30 *Let a real number $k > 0$ be such that $T^{-1}\varepsilon > k$, where ε is as in Lemma 13.29. Then for any curve $v(t)$ from the ball $\mathfrak{U}_k \subset C^0([0,T], T_m M^4)$ of radius k centered at the origin, there exists a vector $C_v \in T_m M^4$ from a bounded neighborhood of the vector $a'(0) = \frac{\mathrm{d}}{\mathrm{d}\tau}a(\tau)_{|\tau=0}$ such that the vector C_v lies inside the light cone of the space $T_{m_0} M^4$ and the curve $m_0 + \int_0^\tau (v(t) + C_v)\mathrm{d}t$ takes the value m_1 at $\tau = T$. The vector C_v is continuous in $v(\cdot)$.*

Proof. Changing the time along $a(\tau)$ construct the straight line $\tilde{a}(s) = a(Ts)$ that meets the conditions $\tilde{a}(0) = m_0$ and $\tilde{a}(1) = m_1$ as in Lemma 13.29. For any curve $v(\cdot) \in \mathfrak{U}_k \subset C^0([0,T], T_m M^4)$, construct the curve $\tilde{v}(s) = Tv(Ts)$ that lies in $\mathfrak{U}_\varepsilon \subset C^0([0,1], T_m M^4)$, i.e., which satisfies Lemma 13.29. In particular, for this curve, there exists a vector $\tilde{C}_{\tilde{v}}$ such that $\|\tilde{C}_{\tilde{v}}\| < C$ from Lemma 13.29. By explicit calculation one can easily derive that

$$m_0 + \int_0^1 (\tilde{v}(s) + \tilde{C}_{\tilde{v}})\mathrm{d}s = m_0 + \int_0^T (v(t) + C_v)\mathrm{d}t = m_1,$$

where $C_v = T^{-1}\tilde{C}_{\tilde{v}}$. □

By construction $\|C_v\| < T^{-1}C$ for any $v \in \mathfrak{U}_k$.

For the tensor G, introduced above, define the norm $\|G_m\|$ by the standard formula

$$\|G_m\| = \sup_{X \in T_m M^4, \|X\| \le 1} \|G_m(X,X)\|.$$

The definition immediately implies the estimate

$$\|G_m(X,X)\| \le \|G_m\|\|X\|^2 \text{ for any } X \in T_m M^4. \tag{13.11}$$

Theorem 13.31 *Let m_0 and m_1 be connected in $\mathcal{O}$ by a straight line $a(\tau)$ that lies inside the light cone of the space $T_{m_0}M^4$ and satisfies the conditions $a(0) = m_0$ and $a(T) = m_1$. Let m_0 and m_1 belong to a ball $V \subset T_m M^4$ such that for any $\hat{m} \in V$ the inequality $\|G_{\hat{m}}\| < \frac{\varepsilon}{(\varepsilon + C)^2}$ holds, where ε and C are as in Lemma 13.29. Then on M^4 there exists a time-like geodesic $m_0(\tau)$ of the Levi-Civitá connection of the Lorentz metric such that $m_0(0) = m_0$ and $m_0(T) = m_1$.*

Proof. Consider the ball $\mathfrak{U}_K \subset C^0([0,T], T_m M^4)$ of radius $K = T^{-1}\varepsilon - \varphi$ centered at the origin, where φ is as in Lemma 12.6. Since $K < T^{-1}\varepsilon$, the assertion of Lemma 13.30 is true for $\mathfrak{U}_K$ and the following completely continuous operator

$$Bv = \int_0^\tau G_{m_0 + \int_0^t (v(s)+C_v)\mathrm{d}s}(v(t) + C_v, v(t) + C_v)\mathrm{d}t$$

is well-defined on this ball. Let us show that this operator has a fixed point in $\mathfrak{U}_K$. Recall that, for any curve $v \in \mathfrak{U}_K$, its C^0-norm is not greater than $K = T^{-1}\varepsilon - \varphi$ and that by Lemma 13.30 $\|C_v\| < T^{-1}C$. Then the hypothesis of the Theorem, the estimate (13.11) and Lemmas 12.6, 13.29 and 13.30 imply that

$$\begin{aligned}
&\left\| G_{m_0 + \int_0^t (v(s)+C_v)\mathrm{d}s} \left((v(t) + C_v), (v(t) + C_v)\right) \right\| \\
&\le \left\| G_{m_0 + \int_0^t (v(s)+C_v)\mathrm{d}s} \right\| \left((\varepsilon T^{-1} - \varphi) + CT^{-1}\right)^2 \\
&< (T^{-2}\varepsilon - T^{-1}\varphi).
\end{aligned}$$

From the last inequality we obtain

$$\left\| \int_0^\tau G_{m_0+\int_0^t (v(s)+C_v)ds}((v(t)+C_v),(v(t)+C_v))\mathrm{d}t \right\| < (T^{-1}\varepsilon - \varphi) = K.$$

This means that the operator $\mathcal{B}$ sends the ball $\mathcal{U}_K$ into itself and so, by Schauder's principle, $\mathcal{B}$ has a fixed point $v_0(t)$ in this ball. It is easy to see that $m_0(\tau) = m_0 + \int_0^\tau (v_0(t) + C_{v_0})\mathrm{d}t$ is a solution of the differential equation (13.10) such that $m_0(0) = m_0$ and $m_0(T) = m_1$. Notice that by construction $m_0(\tau)$ is a geodesic of the Levi-Civitá connection of the Lorentz metric on M^4. The equality $\mathcal{B}v_0 = v_0$ and the definition of $\mathcal{B}$ implies that $v_0(0) = 0$, and hence $\frac{\mathrm{d}}{\mathrm{d}\tau}m_0(\tau)_{|\tau=0} = C_{v_0}$, where by Lemma 13.30 the vector C_{v_0} lies in the light cone of the space $T_{m_0}M^4$, i.e., its Lorentz scalar square is negative. Since the scalar square of the derivative is constant along the geodesic of the Levi-Civitá connection of the Lorentz metric, this geodesic is time-like. □

Remark 13.32. Recall (see above) that C estimates the Euclidean distance between m_0 and m_1. By construction, ε estimates in some sense "how close" m_1 is to the boundary of the proper future of m_0 in the reference frame. This clarifies the meaning of the condition $\|G_{\hat{m}}\| < \frac{\varepsilon}{(\varepsilon+C)^2}$.

13.2.3 The reference frame with Riemannian connection

In this section we investigate the reference frame at the event $m \in M^4$ of the second type described in Section 13.2.1. The manifold here is identical to the manifold in the previous section: T_mM^4 with a specified orthonormal basis, while the connection may be not flat but it is assumed to be compatible with a (positive definite) Riemannian metric on T_mM^4 (see Section 13.2.1). We do not assume this connection to be torsionless.

Here, the question of the existence of a geodesic of the Levi-Civitá connection on M^4 that we are looking for is reduced to the solvability of the two-point boundary value problem for equation (13.9) in the reference frame.

An important difference between this case and the case of a flat connection is the fact that for non-flat connections conjugate points may exist. This yields additional difficulties since there are examples (see Section 12.2) showing that, for a pair of points conjugate along all geodesics joining them, the boundary value problem for a second order differential equation may have no solutions at all. We suppose from the very beginning that m_0 and m_1 are connected by a geodesic in the reference frame along which they are not conjugate. Under these conditions, we find conditions under which the problem for (13.9) is solvable.

The case of a non-flat connection requires more complicated machinery than the previous case. In particular, we replace ordinary integral operators (used in the previous section) by the integral operators with parallel translation from Section 3.2. In fact the method we use here is a modification of that elaborated in Chapter 12.

Everywhere in this section the norms in the tangent spaces and the distances on manifolds are induced by the above-mentioned positive definite Riemannian metric.

Recall that the operator $\mathcal{S}$ from Section 3.2 is well-defined on complete Riemannian manifolds. Consider the neighborhood $\mathcal{O}$ in $T_m M^4$ described in Section 13.2.1. Without loss of generality we may assume that the Riemannian metric on $\mathcal{O}$ is a restriction of a complete Riemannian metric on $T_m M^4$. Indeed, take a relatively compact domain $\mathcal{O}_1 \subset \mathcal{O}$ with smooth boundary such that $\mathcal{O}_1$ contains the points $0 \in T_{m_0}M^4$, m_0 and m_1 as well as the geodesic $\gamma(t)$, where $t \in [0,1]$, joining m_0 and m_1 (if $\mathcal{O}$ is relatively compact one can take $\mathcal{O}_1 = \mathcal{O}$). Then it is possible to change the Riemannian metric outside $\mathcal{O}_1$ so that it becomes complete on $T_m M^4$, and to use $\mathcal{O}_1$ in place of $\mathcal{O}$. Thus, the operator $\mathcal{S}$ is well-defined in this case.

Recall that for a constant curve $v(t) \equiv X \in T_{m_0}\mathcal{M}$ we get by construction that $\mathcal{S}v(t) = \exp X$, where exp is the exponential map of the given connection (see Section 3.2).

Let the points $m_0, m_1 \in \mathcal{O}$ be connected in $\mathcal{O}$ by a geodesic $\gamma(t)$ of the Riemannian connection so that $\gamma(0) = m_0$ and $\gamma(1) = m_1$. In particular, we get $m_1 = \exp(\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t)_{|t=0}) = \mathcal{S}(\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t)_{|t=0})$, where exp is the exponential map of the Riemannian connection. Let m_0 and m_1 be non-conjugate along $\gamma(\cdot)$ and the vector $\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t)_{|t=0}$ lie inside the light cone of $T_{m_0}M^4$.

Hereafter we denote by $\mathfrak{U}_k$ the ball of radius k centered at the origin in some Banach space of continuous maps.

Lemma 13.33 *In the conditions and notation of Lemmas* 3.48 *and* 3.50*, the number ε can be chosen so that for the curve $\tilde{u}(\cdot) \in \mathfrak{U}_\varepsilon \subset C^0([0,1], T_{m_0}\mathcal{O})$ the vector $\tilde{C}_{\tilde{u}}$ lies inside the light cone of the space $T_{m_0}M^4$ and, for the curve $u \in \mathfrak{U}_K \subset C^0([0,T], T_{m_0}\mathcal{O})$, the vector C_u also lies inside the light cone of the space $T_{m_0}M^4$.*

Proof. The fact that, for sufficiently small $\varepsilon > 0$, the vector $C_{\tilde{u}}$ belongs to the interior of the light cone is derived from continuity considerations as in Lemma 13.29. For C_u, this statement follows from the fact that $C_v = T^{-1}C_{\tilde{v}}$, where $\tilde{v}(s) = Tv(Ts) \in \mathfrak{U}_\varepsilon \subset C^0([0,1], T_{m_0}\mathcal{O})$ (see above). □

Hereafter we choose ε satisfying the hypotheses of Lemmas 3.48 and 13.33.

Let $\gamma(t)$ be a C^1-curve given for $t \in [0,T]$ and let $X(t,m)$ be a vector field on $\mathcal{O}$. Denote by $\Gamma X(t,\gamma(t))$ the curve in $T_{\gamma(0)}\mathcal{O}$ obtained by parallel translation of vectors $X(t,\gamma(t))$ along $\gamma(\cdot)$ at the point $\gamma(0)$ with respect to the connection of the reference frame.

Imitating the previous section, we introduce the norm $\|G_m\|$ by means of the norms of vectors with respect to the Riemannian metric on $\mathcal{O}$, as mentioned above. Clearly, (13.11) is valid for $\|G_m\|$, as in the previous section.

With the help of $\mathcal{S}$ and Γ we construct the integral operator $B : \mathfrak{U}_k \to C^0([0,T], T_{m_0}M^4)$ of the following form:

$$Bv = \int_0^\tau \Gamma G_{\mathcal{S}(v(t)+C_v)} \left(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(v(t)+C_v), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(v(t)+C_v) \right) \mathrm{d}t, \tag{13.12}$$

where k and T satisfy the conditions of Lemmas 3.50 and 13.33. One can easily see that the operator B is completely continuous.

Theorem 13.34 *Let $m_0, m_1 \in \mathcal{O}$ and let there exist a geodesic $\gamma(\tau)$ of the connection in the reference frame such that $\gamma(0) = m_0$, $\gamma(T) = m_1$, m_0 and m_1 are not conjugate along $\gamma(\cdot)$ and the vector $\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t)_{|t=0}$ lies inside the light cone of the space $T_{m_0}M^4$. If m_0 and m_1 belong to the ball $V \subset \mathcal{O}$ such that at any $m \in V$ the inequality $\|G_m\| < \frac{\varepsilon}{(\varepsilon+C)^2}$ holds, where ε and C are as in Lemmas* 3.48 *and* 13.33*, then there exists a time-like geodesic $m_0(\tau)$ of the Levi-Civitá connection of the Lorentz metric on M^4 such that $m_0(0) = m_0$ and $m_0(T) = m_1$.*

Proof. Let $k := T^{-1}\varepsilon - \varphi$, where φ is as in Lemma 12.6. For this k, the hypothesis of Lemma 3.50 is satisfied. Hence, on the ball

$$\mathfrak{U}_k \subset C^0([0,T_1], T_{m_0}M^4),$$

the operator (13.12) is well-defined. Recall that, for the curve $v(\cdot) \in \mathfrak{U}_k$, its C^0-norm is not greater than k, that $\|C_v\| < T^{-1}C$ and that the parallel translation with respect to the Riemannian connection preserves the norms of vectors. Then taking into account the definition of operator $\mathcal{S}$ we see that

$$\left\| \frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{S}(v(\tau)+C_v) \right\| < (T^{-1}\varepsilon - \varphi) + T^{-1}C.$$

Hence, formula (13.11), the hypothesis of theorem and Lemma 12.6 imply that

$$\begin{aligned} &\left\| G_{S(v(\tau)+C_v)} \left(\frac{\mathrm{d}}{\mathrm{d}\tau}S(v(\tau)+C_v), \frac{\mathrm{d}}{\mathrm{d}\tau}S(v(\tau)+C_v) \right) \right\| \\ &\le \left\| G_{S(v(\tau)+C_v)} \right\| ((T^{-1}\varepsilon - \varphi) + T^{-1}C)^2 \\ &< (T^{-2}\varepsilon - T^{-1}\varphi). \end{aligned} \tag{13.13}$$

Since the operator Γ of parallel translation preserves the norms of the vectors, from the last inequality we obtain:

$$\begin{aligned} &\left\| \int_0^\tau \Gamma G_{S(v(t)+C_v)} \left(\frac{\mathrm{d}}{\mathrm{d}t}S(v(t)+C_v), \frac{\mathrm{d}}{\mathrm{d}t}S(v(t)+C_v) \right) \mathrm{d}t \right\| \\ &\le \int_0^\tau \left\| G_{S(v(t)+C_v)} \left(\frac{\mathrm{d}}{\mathrm{d}t}S(v(t)+C_v), \frac{\mathrm{d}}{\mathrm{d}t}S(v(t)+C_v) \right) \right\| \mathrm{d}t \\ &< (T^{-1}\varepsilon - \varphi) = k. \end{aligned} \tag{13.14}$$

This means that the completely continuous operator B sends the ball $\mathfrak{U}_k$ into itself. Hence, by Schauder's principle, B has a fixed point $v^*(\tau)$ in $\mathfrak{U}_k$.

Then from the definition of the operator $\mathcal{S}$ and the usual properties of the covariant derivative it follows that $m^*(\tau) = \mathcal{S}(v^*(\tau) + C_{v^*})$ is a solution of the differential equation (13.9) (see Chapter 12). By construction the curve $m^*(\tau)$ is a geodesic of the Levi-Civitá connection of the Lorentz metric on M^4 and, for it, $m^*(0) = m_0$ and $m^*(T) = m_1$.

From the equality $Bv^* = v^*$ and (13.12) it follows that $v^*(0) = 0$. Then from the definition of the operator $\mathcal{S}$ it follows that $\frac{\mathrm{d}}{\mathrm{d}\tau} m_0(\tau)_{|\tau=0} = C_{v^*}$, where the vector C_{v^*} belongs to the light cone of the space $T_{m_0}M^4$ by Lemma 13.33. This means that the Lorentz scalar square of the vector C_{v^*} is negative. Since C_{v^*} is the initial vector of the derivative of the geodesic $m^*(\tau)$ of the Levi-Civitá connection on M^4 and since the Lorentz scalar square of the derivative along this geodesic is constant, the geodesic $m^*(\tau)$ is time-like. $\square$

The meaning of condition $\|G_m\| < \frac{\varepsilon}{(\varepsilon + C)^2}$ is analogous to that in the previous section (see Remark 13.32), but here C estimates the Riemannian distance between m_0 and m_1 while ε depends on the geometry of the reference frame and in some sense estimates “how close” m_1 is to the boundary of the “proper future” of m_0 and “how close” it is to conjugate points in the reference frame.

13.3 A Classical Particle in a Classical Gauge Field

In this Section we introduce and investigate a special class of differential equations on the total spaces of fiber bundles equipped with connections. We interpret these equations as equations of motion for a classical particle in a classical gauge field.

It should be pointed out that the term ‘gauge field’, as used in contemporary physics, means a connection on a fiber bundle. Gauge fields are generally used in quantum physics.

The problem of the motion of a classical particle in a classical gauge field was first treated by S.K. Wong (see [231]). Further developments can be found, e.g., in [143, 201, 213, 214, 215]. The symplectic geometry and the Hamiltonian formalism on some special fiber bundles were exploited in [213, 214, 215]. Here we suggest another approach using the connection (i.e., the gauge field itself) on bundles different from those of [143, 213, 214, 215]. This approach seems to be more natural; we show that in the special case of the electromagnetic field our equation is reduced to the ordinary Lorentz equation (13.5).

In what follows we denote fiber bundles and their total spaces by the same symbols.

13.3.1 A brief introduction to gauge fields and some preliminary constructions

Let M^4 be a Lorentz manifold with a metric $g(\cdot,\cdot)$. We retain the standard notation M^4 of this chapter for a Lorentz manifold since our principal example of such a manifold is a space-time. Nevertheless please note that in all constructions below in this section we do not use the fact that the dimension of the manifold is 4 and so all our results are valid in arbitrary finite-dimensional Lorentz manifolds.

Consider a principal bundle $\varPi\colon \mathcal{E} \to M^4$ with structure group G. The Lie algebra of G will be denoted by $\mathfrak{g}$.

Let H be a connection on the bundle $\mathcal{E}$. Recall (see Section 2.7) that H is a distribution on $\mathcal{E}$, i.e., a sub-bundle of $T\mathcal{E}$, that is invariant with respect to the natural right action of G on $\mathcal{E}$ and is complementary to the vertical distribution V consisting of subspaces tangent to the fibers of $\mathcal{E}$. Note that all subspaces $\mathsf{V}_b \subset T_b\mathcal{E}$ are canonically isomorphic to $\mathfrak{g}$ (see Section 2.7). We introduce a semi-Riemannian (Lorentz) metric $g^{\mathcal{E}}$ on the manifold $\mathcal{E}$ by determining an inner product in $\mathfrak{g}$, and hence in all subspaces V, by defining the inner product in the subspaces H as the inverse image of g with respect to $T\varPi$, and by setting the subspaces V_b and H_b in all $T_b\mathcal{E}$ to be orthogonal to each other. Following this we can treat the manifold $\mathcal{E}$ as an ordinary semi-Riemannian manifold, in particular, we can consider all the usual operations with differential forms on $\mathcal{E}$.

In what follows we shall also denote by V and H the projections of $T_b\mathcal{E} = \mathsf{V}_b \oplus \mathsf{H}_b$ onto the vertical (i.e., V_b) and horizontal (i.e., H_b), respectively, subspaces in $T_b\mathcal{E}$ where $b \in \mathcal{E}$ is an arbitrary point. So, for $X \in T_b\mathcal{E}$, $\mathsf{V}X$ is its vertical component and $\mathsf{H}X$ is its horizontal component. This notation will be used when dealing with all types of fiber bundles with connections.

Consider the connection form φ and the curvature form $\varPhi = \mathrm{D}\varphi$ of H where the covariant differential D on the right hand side is defined by the usual formula $\mathrm{D}\varphi(\cdot,\cdot) = \mathrm{d}\varphi(\mathsf{H}\cdot,\mathsf{H}\cdot)$ (see Section 2.7). Recall that φ is a vertical 1-form (i.e., equal to zero on subspaces H) and $\varPhi$ is a horizontal 2-form (i.e., equal to zero on subspaces V), both taking values in $\mathfrak{g}$. We interpret $\varPhi$ as the gauge field strength. The gauge field, as usual, is determined by the equations

$$\mathrm{D}\varPhi = 0, \qquad \mathrm{D} * \varPhi = *J, \tag{13.15}$$

where J is a horizontal 1-form on $\mathcal{E}$ with values in $\mathfrak{g}$ and $*$ is the star operator. The first equation is Bianchi's identity (2.39). Note that (13.15) are analogous to Maxwell's equations (13.4).

Let $\mathcal{F}$ be a (real or complex) linear space with inner product $h(\cdot,\cdot)$. Suppose that the left action of G on $\mathcal{F}$ is given and $h(\cdot,\cdot)$ is invariant under this action. We interpret $\mathcal{F}$ as the space of internal states of the particle and the group G as the group of its internal symmetries. In addition, we suppose that a charge is given, i.e., a map $e\colon \mathcal{F} \to \mathfrak{g}^*$ that is constant on the orbits of G. The space $\mathfrak{g}^*$ is a coalgebra, i.e., $\mathfrak{g}^*$ is dual to $\mathfrak{g}$.

Note that the electromagnetic field is an example of a gauge field with $G = U(1)$ (the group of unitary operators in the complex plane $\mathbb{C}^1$), $\mathcal{F} = \mathbb{C}^1$ and $h(X,Y) = X\bar{Y}$, where the bar indicates complex conjugation. Below we show how in this case Maxwell's equations (13.4) can be derived from (13.15) (see formula (13.22)).

Consider the bundle $\pi\colon \mathcal{Q} \to M^4$ associated to $\mathcal{E}$ with standard fiber $\mathcal{F}$ (see Section 1.3). Recall that $\mathcal{Q}$ is obtained by the factorization $\lambda\colon \mathcal{E}\times\mathcal{F} \to \mathcal{Q}$ with respect to the right action of G on $\mathcal{E} \times \mathcal{F}$ determined as $(\bar{e}, \bar{f})g = (bg, g^{-1}f)$ for $g \in G$, $b \in \mathcal{E}$ and $f \in \mathcal{F}$. In particular, the tangent map $T\lambda$ sends the spaces of the connection H from the tangent spaces to $\mathcal{E}$ into the tangent spaces to $\mathcal{Q}$ (see Section 2.7). The connection on $\mathcal{Q}$ obtained from H (recall that $T\lambda$ is one-to-one on H) is denoted by H^π. Since $\mathcal{Q}$ is a bundle with connection, the notation V and H will be also used to denote the projections onto vertical and horizontal, respectively, subspaces in tangent spaces to $\mathcal{Q}$, as was mentioned above. To avoid any confusion we shall denote the vertical distribution on $\mathcal{Q}$ by V^π.

Consider a point $q = (m, c) \in \mathcal{Q}$, where $m = \pi q$ so that c belongs to the fiber $\mathcal{Q}_m$ of $\mathcal{Q}$ at m. Note that by definition $\mathsf{V}^\pi_{(m,c)}$, the vertical subspace in $T_{(m,c)}\mathcal{Q}$, is the tangent space to $\mathcal{Q}_m$ and so, since the latter is a linear space, these spaces are naturally isomorphic to each other. Denote by $p^\pi : \mathsf{V}_{(m,c)} \to \mathcal{Q}_m$ the natural linear isomorphism as in formulae (1.2) and (2.11). Recall that the connector (connection map) of H^π is the map $K^\pi : T\mathcal{Q} \to \mathcal{Q}$ defined as the composition $K = p^\pi \circ \mathsf{V}$.

The restriction of $T\pi : \mathsf{H}^\pi \to TM^4$ to H_q is a linear isomorphism of H_q onto T_mM^4 for $m = \pi q$. Its inverse map $T\pi^{-1}_{|\mathsf{H}^\pi_q}$ sends the tangent space T_mM^4 into the connection $\mathsf{H}^\pi_q \subset T_q\mathcal{Q}$ at a specified point $q \in \mathcal{Q}$ such that $\pi q = m$.

Let $\mathfrak{U}$ be a chart on M^4. The bundle $\mathcal{Q}$ over $\mathfrak{U}$ is represented as the cartesian product $\mathfrak{U}\times\mathcal{F}$. The subspace ${}^{\mathfrak{U}}\mathsf{H}^\pi_q$ in $T_q\mathcal{Q}$, $q = (m, c)$, $m = \pi q \in \mathfrak{U}$, corresponding to the tangent space to $\mathfrak{U}$ in the above product, is obviously complementary to V_q. It is called the *Euclidean connection* of the coordinate system in $\mathfrak{U}$ (see Sections 2.1 and 2.2). Using the presentation $T_q\mathcal{Q} = {}^{\mathfrak{U}}\mathsf{H}^\pi_q \oplus \mathsf{V}_q$ we obtain the description of any vector $Y \in T_q\mathcal{Q}$ in local coordinates as a quadruple $Y = (m, c, Y_1, Y_2)$, $Y_1 \in {}^{\mathfrak{U}}\mathsf{H}^\pi_q$ and $Y_2 \in \mathsf{V}_q$.

Since ${}^{\mathfrak{U}}\mathsf{H}^\pi_q$ is complementary to $\mathsf{V}_q = \mathrm{Ker}T\pi$, the map $T\pi : {}^{\mathfrak{U}}\mathsf{H}^\pi_q \to T_mM^4$ is one-to-one (as for H^π_q, see above). For a vector $X \in T_mM^4$ consider the vector $T\pi^{-1}_{|\,{}^{\mathfrak{U}}\mathsf{H}^\pi_q}X - T\pi^{-1}_{|\mathsf{H}^\pi_q}X$. Since this vector belongs to V_q, we can apply p^π to it. As in Sections 2.2 and 2.3, the resulting vector $\boldsymbol{\Gamma}^\pi_m(c, X) = p^\pi(T\pi^{-1}_{|\,{}^{\mathfrak{U}}\mathsf{H}^\pi_q}X - T\pi^{-1}_{|\mathsf{H}^\pi_q}X) \in \mathcal{Q}_m$ is called the *local coefficient* of the connection H^π (or *local connector* of H^π) in the chart $\mathfrak{U}$. Recall that the connector K^π has the following description in local coordinates:

$$K^\pi(m, c, Y_1, Y_2) = (m, Y_2 + \boldsymbol{\Gamma}^\pi_m(c, Y_1)).$$

Consider the tangent bundle $\tau\colon TM^4 \to M^4$ (here, to avoid any confusion, we denote the natural projection by τ) and let H^τ be the Levi-Civitá connection of the metric $g(\cdot,\cdot)$. This means that H^τ is a distribution on TM^4 complementary to the vertical distribution V^τ. Special features of H^τ can be found in Section 2.6. For a point $(m, X) \in TM^4$, where $m = \tau(m, X)$ and $X \in T_mM^4$, the subspace $\mathsf{V}^\tau_{(m,X)}$ in the second tangent space $T_{(m,X)}TM^4$ is tangent to the fiber T_mM^4. Denote by $p^\tau : \mathsf{V}^\tau_{(m,X)} \to T_mM^4$ the natural linear isomorphism as in formulae (1.2) and (2.11). Now we can consider the connector $K^\tau\colon T^2M^4 \to TM^4$ of H^τ defined by the natural formula $K^\tau = p^\tau \circ \mathsf{V}^\tau$ (here T^2M^4 is the second tangent bundle TTM^4).

The general construction of the Euclidean connection, the presentation of vectors as quadruples and the definition of the local coefficient of a connection in a chart described above for the bundle $\mathcal{Q}$ are valid for the tangent bundle with H^τ. Denote the local coefficient of H^τ by $\boldsymbol{\Gamma}^\tau_m(Y, X)$, $X, Y \in T_mM^4$. (Here T_mM^4 plays the role of $\mathcal{Q}_m$ and so $c \in \mathcal{Q}_m$ is replaced by $Y \in T_mM^4$.)

We shall use the connection $\mathsf{H}^{\mathcal{Q}}$ on the manifold $\mathcal{Q}$ (the total space of the bundle $\mathcal{Q}$) constructed from the connections H^π and H^τ in Section 2.8. Denote by $K\colon T^2\mathcal{Q} \to T\mathcal{Q}$ its connector.

Recall that $K = K^{\mathsf{H}} \oplus K^{\mathsf{V}}$ where $K^{\mathsf{H}}\colon T^2\mathcal{Q} \to \mathsf{H}^\pi$ and $K^{\mathsf{V}}\colon T^2\mathcal{Q} \to \mathsf{V}^\pi$, K^{H} is determined by the formula $K^{\mathsf{H}} = T\pi^{-1}_{|\mathsf{H}^\pi} \circ K^\tau \circ T^2\pi$, $T^2\pi : T^2\mathcal{Q} \to T^2M^4$ is the tangent map to $T\pi : T\mathcal{Q} \to TM^4$ and the latter is the tangent map to π. Note that the image $K^\tau \circ T^2\pi(T_{(q,C)}T\mathcal{Q})$ belongs to T_mM^4 where $m = \pi q$ so that $T\pi^{-1}_{|\mathsf{H}^\pi}$ sends it into $\mathsf{H}^\pi_q \subset T_q\mathcal{Q}$ as is required in the definition of a connector.

On the other hand $K^{\mathsf{V}} = (p^\pi)^{-1} \circ K^\pi \circ TK^\pi$. Note that the image $K^\pi \circ TK^\pi(T_{(q,C)}T\mathcal{Q})$ belongs to the fiber $\mathcal{Q}_m$ where $m = \pi q$ and $(p^\pi)^{-1}$ sends it into V_q as required.

The covariant derivative of $\mathsf{H}^{\mathcal{Q}}$ on $\mathcal{Q}$ is defined by the usual formula $\frac{\mathrm{D}}{\mathrm{d}t} = K \circ \frac{\mathrm{d}}{\mathrm{d}t}$. By construction we have $\frac{\mathrm{D}}{\mathrm{d}t} = \frac{\mathrm{D}}{\mathrm{d}t}^{\mathsf{H}} + \frac{\mathrm{D}}{\mathrm{d}t}^{\mathsf{V}}$, where $\frac{\mathrm{D}}{\mathrm{d}t}^{\mathsf{H}} = K^{\mathsf{H}} \circ \frac{\mathrm{d}}{\mathrm{d}t}$ and $\frac{\mathrm{D}}{\mathrm{d}t}^{\mathsf{V}} = K^{\mathsf{V}} \circ \frac{\mathrm{d}}{\mathrm{d}t}$.

Define a Riemannian metric $g^{\mathcal{Q}}(\cdot,\cdot)$ on $\mathcal{Q}$ so that it coincides with $h(\cdot,\cdot)$ in vertical subspaces V^π, with the inverse image $T\pi^*g$ in the horizontal subspaces H^π and so that the spaces H^π_q and V^π_q at any point $q \in M^4$ are orthogonal to each other.

In this particular case a vector $\bar{\alpha}$ and a 1-form $\tilde{\alpha}$ on $\mathcal{Q}$ are said to be physically equivalent to each other with respect to the metric $g^{\mathcal{Q}}$ if for any vector Y we have $\tilde{\alpha}(Y) = g^{\mathcal{Q}}(\bar{\alpha}, Y)$ (cf. Section 1.4). We shall denote them by the same symbol, using a bar over vectors and a tilde over 1-forms.

13.3.2 The equation of motion

Since λ is a one-to-one map from the standard fiber $\mathcal{F}$ onto each fiber of $\mathcal{Q}$, the charge e is well-defined on the whole manifold $\mathcal{Q}$. Also $T\lambda$ is a one-to-one map of the connections and $\varPhi$ is equivariant (see [26]). Hence we can set up the form $\tilde{\varPhi}$ with values in $\mathfrak{g}$ on the manifold $\mathcal{Q}$ by the equality $\tilde{\varPhi}_q(X,Y) = \varPhi_b(T\lambda^{-1}\mathsf{H}X, T\lambda^{-1}\mathsf{H}Y)$ for $q = \lambda(b,f)$, where $b \in \mathcal{E}$, $f \in \mathcal{F}$, and for $X, Y \in T_q\mathcal{Q}$. Unlike $\tilde{\varPhi}_q(\cdot,\cdot)$, the 2-form $e(q) \bullet \tilde{\varPhi}_q(\cdot,\cdot)$ on $T_q\mathcal{Q}$ takes values in $\mathbb{R}$ and so it is an ordinary scalar-valued 2-form. The symbol $\bullet$ denotes the coupling of the elements $e(q) \in \mathfrak{g}^*$ and $\tilde{\varPhi}_q(\cdot,\cdot) \in \mathfrak{g}$.

Thus, on the total space $\mathcal{Q}$ we may consider the following equation, an analog of Newton's second law (11.2) and of the Lorentz equation (13.5), in the form

$$\frac{\mathrm{D}}{\mathrm{d}t}\dot{q} = \overline{e(q) \bullet \tilde{\varPhi}_q(\cdot, \dot{q})}. \tag{13.16}$$

We interpret these equations as equations of motion for a classical particle with charge in a classical gauge field.

Proposition 13.35 *The vector $\overline{e(q) \bullet \tilde{\varPhi}_q(\cdot, \dot{q})}$ is horizontal, i.e., at any point $q \in \mathcal{Q}$ it belongs to H^π_q.*

Proof. Since $\varPhi$ is a horizontal form by construction, so too is $\tilde{\varPhi}(\cdot,\cdot)$, i.e., $\tilde{\varPhi}_q(Y,\dot{q}) = 0$ for all $Y \in \mathsf{V}^\pi_q$. Thus $e(q) \bullet \tilde{\varPhi}_q(Y,\dot{q}) = 0$ for the same Y. Hence $g^{\mathcal{Q}}(\overline{e(q) \bullet \tilde{\varPhi}_q(\cdot,\dot{q})}, Y) = 0$. □

From Proposition 13.35 it follows that equation (13.16) is equivalent to the system

$$\frac{\mathrm{D}}{\mathrm{d}t}^{\mathsf{H}}\dot{q} = \overline{e(q) \bullet \tilde{\varPhi}(\cdot, \dot{q})}, \tag{13.17}$$

$$\frac{\mathrm{D}}{\mathrm{d}t}^{\mathsf{V}}\dot{q} = 0. \tag{13.18}$$

Theorem 13.36 *If $q(t)$ is a solution of (13.16) with horizontal initial condition $\dot{q}(0) = \dot{q}_0 \in \mathsf{H}^\pi$, then $q(t)$ is a horizontal curve (i.e., $\dot{q} \in \mathsf{H}^\pi$ for all t, for which it exists) and $e(q(t))$ is constant.*

Proof. To prove this we consider local coordinates so that $q(t) = (m(t), c(t))$, where $m(t) \in M^4$, and $c(t)$ belongs to the fiber $\mathcal{Q}_{m(t)}$. Hence $\dot{q} = (m, c, \dot{m}, \dot{c})$ and $K^\pi\dot{q} = K^\pi(m,c,\dot{m},\dot{c}) = (m, \dot{c} + \boldsymbol{\Gamma}^\pi_m(c,\dot{m}))$. So, the condition $\dot{q}_0 \in \mathsf{H}^\pi$ in the local coordinates means that $\dot{c}_0 + \boldsymbol{\Gamma}^\pi_m(c_0,\dot{m}_0) = 0$. We have,

$$\begin{aligned} TK^\pi\frac{\mathrm{d}}{\mathrm{d}t}\dot{q} = \frac{\mathrm{d}}{\mathrm{d}t}K^\pi\dot{q} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(m, \dot{c} + \boldsymbol{\Gamma}^\pi_m(c,\dot{m})\right) \\ &= \left(m, \dot{c} + \boldsymbol{\Gamma}^\pi_m(c,\dot{m}), \dot{m}, \frac{\mathrm{d}}{\mathrm{d}t}(\dot{c} + \boldsymbol{\Gamma}^\pi_m(c,\dot{m}))\right). \end{aligned}$$

By (13.18) $\frac{\mathrm{D}^{\mathsf{V}}}{\mathrm{d}t}\dot{q} = (p^{\pi})^{-1} \circ K^{\pi} \circ TK^{\pi} \frac{\mathrm{d}}{\mathrm{d}t}\dot{q} = 0$ so that $TK^{\pi} \frac{\mathrm{d}}{\mathrm{d}t}\dot{q}$ belongs to the connection. This means that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(m, \dot{c} + \boldsymbol{\Gamma}^{\pi}_{m}(c, \dot{m})\right) = \left(m, \dot{c} + \boldsymbol{\Gamma}^{\pi}_{m}(c, \dot{m}), \dot{m}, -\boldsymbol{\Gamma}^{\pi}_{m}(\dot{c} + \boldsymbol{\Gamma}^{\pi}_{m}(c, \dot{m}), \dot{m})\right).$$

Thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{c} + \boldsymbol{\Gamma}^{\pi}_{m}(c, \dot{m})\right) = -\boldsymbol{\Gamma}^{\pi}_{m}\left(\dot{c} + \boldsymbol{\Gamma}^{\pi}_{m}(c, \dot{m}), \dot{m}\right). \tag{13.19}$$

Equality (13.19) is an ordinary linear differential equation with respect to $\dot{c} + \boldsymbol{\Gamma}^{\pi}_{m}(c, \dot{m})$. For zero initial conditions (see above) it has a unique solution $\dot{c} + \boldsymbol{\Gamma}^{\pi}_{m}(c, \dot{m}) = 0$. This proves the first statement.

By construction the curve $q(t) = (m(t), c(t))$ is horizontal and so it can be represented as $(m(t), b_t(f))$, where b_t is a horizontal lift of $m(t)$ into $\mathcal{E}$ and f is some vector in $\mathcal{F}$. Hence $b_t(f)$ belongs to an orbit of G in $\mathcal{F}$ and so $e(q(t)) = e(b_t(f))$ is constant. □

Theorem 13.37 *If $m(t) = \pi q(t)$ is the projection of a solution of* (13.16)*, then $m(t)$ satisfies $\frac{\mathrm{D}}{\mathrm{d}t}^{\tau}\dot{m}(t) = T\pi(\overline{e(q) \bullet \tilde{\Phi}(\cdot, \dot{q})})$ where $\frac{\mathrm{D}}{\mathrm{d}t}^{\tau}$ is the covariant derivative of the Levi-Civitá connection on M^4.*

Proof. Denote $q = (m, c)$ as above. Describing vectors in local coordinates by quadruples, we obtain $\dot{q} = (m, c, \dot{m}, \dot{q})$ and $\frac{\mathrm{d}}{\mathrm{d}t}\dot{q} = (m, c, \dot{m}, \dot{c}, \dot{m}, \dot{c}, \ddot{m}, \ddot{c})$.

For $m = \pi q$ the vector $\dot{m}$ clearly belongs to $T_m M^4$ and so we represent it in the form $(m, \dot{m})$ as a point of TM^4. Then $\frac{\mathrm{d}}{\mathrm{d}t}\dot{m} = (m, \dot{m}, \dot{m}, \ddot{m})$.

By the definition of covariant derivative

$$\frac{\mathrm{D}}{\mathrm{d}t}^{\tau}\dot{m} = K^{\tau}\frac{\mathrm{d}}{\mathrm{d}t}\dot{m} = K^{\tau}(m, \dot{m}, \dot{m}, \ddot{m}).$$

On the other hand, from the construction of K^{H}

$$T\pi\frac{\mathrm{D}}{\mathrm{d}t}^{\mathsf{H}}\dot{q} = T\pi K^{\mathsf{H}}\frac{\mathrm{d}}{\mathrm{d}t}\dot{q} = K^{\tau} \circ T^2\pi(m, c, \dot{m}, \dot{c}, \dot{m}, \dot{c}, \ddot{m}, \ddot{c}) = K^{\tau}(m, \dot{m}, \dot{m}, \ddot{m}).$$

Thus, $\frac{\mathrm{D}}{\mathrm{d}t}^{\tau}\dot{m} = T\pi\frac{\mathrm{D}}{\mathrm{d}t}^{\mathsf{H}}\dot{q}$. So, applying $T\pi$ to both sides of (13.17) we obtain the theorem. □

The most important cases are those where G is $U(1)$, or $SU(2)$, or $SU(3)$ and $\mathcal{F}$ is a complex space with appropriate dimension. They correspond to the well-known gauge fields. As an example we consider the particular case of an electromagnetic field with $G = U(1)$, $\mathcal{F} = \mathbb{C}^1$, $h(X, Y) = X\bar{Y}$, where the bar denotes complex conjugation.

Here the algebra $\mathcal{U}(1) = \mathbb{R}^1$ is one dimensional and hence commutative. Thus φ and Φ are ordinary differential forms with values in $\mathbb{R}$. The charge takes values in real numbers and it is constant on the circles in C^1 centered at zero. Recall the structure equation (2.40): $\mathrm{d}\varphi = -\frac{1}{2}[\varphi, \varphi] + \Phi$. From the commutativity of $U(1)$ it follows that $[\varphi, \varphi](X, Y) = \varphi(X)\varphi(Y) - \varphi(Y)\varphi(X) = 0$, for each pair $X, Y \in T_b\mathcal{E}$, $b \in \mathcal{E}$. Thus $\Phi = \mathrm{D}\varphi = \mathrm{d}\varphi$ and

$$\begin{aligned} \mathrm{D}\Phi &= \mathrm{D}\mathrm{d}\varphi = \mathrm{d}\mathrm{d}\varphi(\mathsf{H}\cdot, \mathsf{H}\cdot, \mathsf{H}\cdot) = 0 \\ \mathrm{D} * \Phi &= \mathrm{D} * \mathrm{d}\varphi = \mathrm{d} * \mathrm{d}\varphi(\mathsf{H}\cdot, \mathsf{H}\cdot, \mathsf{H}\cdot). \end{aligned} \tag{13.20}$$

From (12.6) and (13.20) we obtain

$$\mathrm{d}\Phi = 0, \qquad \mathrm{d} * \Phi = *J. \tag{13.21}$$

Recall that Φ is horizontal and equivariant (see [26]). Since the group $U(1)$ is commutative, the latter means that Φ is invariant with respect to the natural right action of G on $\mathcal{E}$: $\Phi_{R_g b}(TR_g X, TR_g Y) = \Phi_b(X, Y)$ where R_g is the right action of $g \in G$ on $\mathcal{E}$, $b \in \mathcal{E}$ and $X, Y \in T_b\mathcal{E}$. Thus the equality

$$\Psi_m(X, Y) = \Phi_b(T\Pi^{-1}_{|\mathsf{H}_b} X, T\Pi^{-1}_{|\mathsf{H}_b} Y),$$

for any $m \in M^4$, $b \in \mathcal{E}$, $\Pi b = m$ and $X, Y \in T_m M^4$, defines a 2-form $\Psi(\cdot, \cdot)$ on M that is well-defined (i.e., does not depend on the choice of b over m).

For Ψ equations (13.21) are reduced to

$$\mathrm{d}\Psi = 0, \qquad \mathrm{d} * \Psi = *\tilde{J}. \tag{13.22}$$

Taking into account Definition 1.71 of the codifferential $\delta = *^{-1}\mathrm{d}*$, one can easily see that (13.22) are the ordinary Maxwell equations (13.4) on the Lorentz manifold M^4 where $\tilde{J}_m(X) = J(T\Pi^{-1}_{H_b} X)$, $m = \Pi b$ and $X \in T_m M^4$ are well-defined. Thus we may consider Ψ as the strength of the electromagnetic field on M^4.

Clearly $T\pi(\overline{e(q) \bullet \tilde{\Phi}(\cdot, \dot{q})}) = e(q)\overline{\Psi(\cdot, \dot{m})}$. Hence, from Theorem 13.37 it follows that a particle with charge $e(q)$ is governed by the Lorentz equation (13.5):

$$\frac{\mathrm{D}^\tau}{\mathrm{d}t}\dot{m} = e(q)\overline{\Psi(\cdot, \dot{m})}$$

on M^4 and the charge e is constant.

Chapter 14
Mechanical Systems with Random Perturbations

14.1 Setting Up the Problem

It is a well-known fact that a second order differential equation $\ddot{x}(t) = \bar{\alpha}(t, x(t), \dot{x}(t))$ expressing Newton's law in $\mathbb{R}^n$ may be represented as a first order system on the space of dimension $2n$:

$$\begin{cases} \dot{x}(t) = v(t) \\ \dot{v}(t) = \bar{\alpha}(t, x(t), v(t)). \end{cases} \tag{14.1}$$

We call the first equation of the above system horizontal and the second one vertical. This is consistent with the terminology in the general case of a mechanical system on a non-linear configuration space M, where Newton's law is presented by means of covariant derivatives in the form (11.2), equivalent to equation (11.3) with a special vector field (second order differential equation) on the phase space TM. Recall that the field (11.3) is the sum of the Levi-Civitá geodesic spray $\mathcal{Z}$ (which is horizontal, i.e. belongs to the connection) and the vertical lift of the vector force field (which belongs to the vertical subspace).

Thus a random perturbation of Newton's law can arise in the horizontal equation and in the vertical equation (possibly both). Note that a vertical perturbation is a perturbation of the force field while a horizontal perturbation is a perturbation of velocity. All the cases are physically reasonable but they require different methods of investigation. It should also be pointed out that under random perturbations, Newton's law becomes a random differential equation, which we shall present in terms of mean derivatives. Taking into account various possible constructions of second order mean derivatives (forward, backward, mixed, etc.), this yields several equations of motion from different parts of physics.

In this chapter we deal with Newton's law in terms of forward mean derivatives. Its physical interpretation is the description of the motion of an ordi-

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_14,

nary mechanical system with random perturbations. We consider both perturbations of forces and of velocities.

The versions of Newton's law in terms of backward and mixed mean derivatives will be considered in forthcoming chapters. Newton's law in special mixed derivatives (the so-called Newton-Nelson equation) describes the motion of a quantum particle (see Chapter 15). Newton's law in backward mean derivatives on groups of diffeomorphisms describes the motion of a viscous incompressible fluid (Chapter 16).

We begin this Chapter with an investigation of the so-called Langevin equation.

The Langevin equation describes mechanical systems with both deterministic and random forces which have comparable magnitudes (i.e., neither the deterministic nor random part can be neglected) where the random force is a transformed white noise. Examples of such processes are well-known in physics. One can easily see that in this case the trajectories of the process are a.s. C^1-smooth. This makes the analysis technically much simpler than that for more general systems. In particular we can apply the machinery of ordinary integral operators with Riemannian parallel translation of Section 3.2 and study the Langevin equations arising on non-linear configuration spaces.

In Section 14.2, we introduce the Langevin equation on a Riemannian manifold and reduce it to the velocity hodograph equation, which is an equation in a single tangent (i.e., vector) space. This enables us to apply some standard results and carry out a detailed analysis. We also study an important particular case of the Langevin equation: the equation describing the so-called Ornstein-Uhlenbeck processes arising, for example, in the mathematical model of physical Brownian motion [188]. Sometimes only the latter equation is called the Langevin equation, whereas the equation applicable in a more general context is said to be the generalized Langevin equation.

In Section 14.3 we study the case where the force field in the Langevin equation is set-valued (i.e., it is constructed from an essentially discontinuous force or where a force with feedback control is under consideration).

Throughout Sections 14.2 and 14.3, all Riemannian manifolds are assumed to be complete, but not necessarily uniformly or stochastically complete.

In Section 14.4 we investigate mechanical systems with random perturbations of velocity motivated by the motion of a particle, subjected to a deterministic force, that in addition moves in an enveloping medium with random influence. First we consider systems in $\mathbb{R}^n$ with single-valued and set-valued forces. The systems on manifolds are investigated under some more restrictive assumptions. In particular, we assume the manifold to be stochastically complete and use the machinery of stochastic parallel translation to reduce the system to the corresponding velocity hodograph equation.

14.2 The Langevin Equation and Ornstein-Uhlenbeck Processes on Manifolds

Everywhere in this section we deal with processes given on a finite time interval $[0, l] \subset \mathbb{R}$.

Consider a mechanical system in the sense of Section 11.1, i.e., a Riemannian manifold M together with a force field on M. As just mentioned, M is assumed to be complete, i.e., a free particle on M cannot escape to infinity in finite time. The Riemannian metric enables us to identify differential forms and vector fields on M, so henceforth we regard the force field as a vector field (see Section 11.1). In addition to the Definition 11.2 of a vector force field we give the following:

Definition 14.1. A map A from $\mathbb{R} \times TM$ to the bundle of $(1,1)$-tensors over M (either single or set-valued), such that $\pi_1 A(t,m,X) = \pi(m,X) = m$ (where π_1 is the projection of the bundle of $(1,1)$-tensors onto M) will be called a *tensor force field*.

Recall that a $(1,1)$-tensor at $m \in M$ is a linear operator in T_mM. Thus for a tensor force field $A(t,m,X)$ and a vector field $Y(m)$ on M the composition $A(t,m,X)Y(m)$ is a vector force field.

Let $\bar{\alpha}(t,m,X)$ be a vector force field and $A(t,m,X)$ a $(1,1)$-tensor field on M. In other words, for every $t \in [0,l]$, $m \in M$, and $X \in T_mM$, we have a vector $\bar{\alpha}(t,m,X) \in T_mM$ and a linear operator $A(t,m,X)\colon T_mM \to T_mM$. Making use of the construction given in Section 7.7.2, specify a Wiener process $w(t)$ on the tangent spaces to M and denote by $\dot{w}$ the white noise of $w(t)$ (see Section 6.2.1).

The Langevin equation describes the evolution of a system with the force field

$$\bar{\alpha}(t,m,X) + A(t,m,X)\dot{w}. \tag{14.2}$$

More formally, the equation of motion must read

$$\frac{\mathrm{D}}{\mathrm{d}t}\,\dot{\xi}(t) = \bar{\alpha}\left(t, \xi(t), \dot{\xi}(t)\right) + A\left(t, \xi(t), \dot{\xi}(t)\right)\dot{w}(t), \tag{14.3}$$

however this expression is meaningful only in the sense of distributions.

Our first goal is to give a rigorous meaning to (14.3) without using distributions. We do this in terms of forward mean derivatives, the construction of which can be simplified in the case under consideration. Then we transform the obtained equation into an equivalent integral form employing the integrals with Riemannian parallel translation from Section 3.2 by analogy with that given in Section 11.8. In particular we use the transition to the corresponding velocity hodograph equation which is much easier to study since it is an equation in a single tangent (i.e., linear) space. Note that previously, in [94, 118, 119], the Langevin equation was considered only in this integral form. A local coordinate version of the equation was independently given in [180].

We assume that $\bar{\alpha}(t,m,X)$ and $A(t,m,X)$ are jointly continuous in all variables and that these fields have linear growth in X. In other words, there exists a constant $K>0$ such that

$$\|\bar{\alpha}(t,m,X)\| + \|A(t,m,X)\| < K\,(1+\|X\|) \tag{14.4}$$

for all $t\in[0,l]$, $m\in M$, and $X\in T_mM$, where the norm is given by the Riemannian metric. The estimate (14.4) is a version of the Itô condition.

One can show, appealing to physical reasoning, that a process subjected to the force (14.2) a.s. has continuous velocities and as a consequence it a.s. has C^1-smooth sample paths. Below we shall show that solutions of the Langevin equation do indeed exist in the class of processes with C^1-smooth sample paths. This is why we start by examining some features of such processes.

Let $\xi(t)$ be a stochastic process on M with a.s. C^1-smooth sample paths given on a probability space $(\Omega,\mathcal{F},\mathsf{P})$, and let a vector field Y be given on M. As above, by $\Gamma_{t.s}$ we denote the operator of parallel translation along a C^1-smooth curve $x(\cdot)$ from $x(s)$ to $x(t)$. Since the sample paths of $\xi(t)$ are C^1-smooth, the parallel translation in the definition of a forward mean derivative, by formula (9.15), is the ordinary parallel translation (i.e., we needn't deal with the general construction of a stochastic parallel translation from Section 7.6). Thus we have obtained:

Lemma 14.2 *The covariant forward mean derivative of the vector field Y along the process $\xi(t)$ on M with a.s. C^1-smooth sample paths at time t is the L^1 random element of the form*

$$\mathbf{D}Y(t,\xi(t)) = \lim_{\triangle t\downarrow 0} E_t^{\xi}\left(\frac{\Gamma_{t,t+\triangle t}Y(t+\triangle t,\xi(t+\triangle t)) - Y(t,\xi(t))}{\triangle t}\right), \tag{14.5}$$

where $\Gamma_{t,s}$ is the ordinary parallel translation along C^1-smooth curves.

Consider a probability space $(\Omega,\mathcal{F},\mathsf{P})$ and a non-decreasing family of complete σ-subalgebras $\mathcal{B}_t$ of $\mathcal{F}$. In a given tangent space $T_{m_0}M$ introduce a Wiener process $w(t)$ adapted to $\mathcal{B}_t$, and an Itô diffusion type process $v(t)$ of the form $v(t)=\int_0^t b(s)\mathrm{d}s+\int_0^t B(s)\mathrm{d}w(s)$ with $b(t)$ and $B(t)$ a.s. having continuous sample paths. In particular this means that $v(t)$ is non-anticipative with respect to $\mathcal{B}_t$ and a.s. has continuous sample paths. Thus we can apply the operator $\mathcal{S}$ introduced in Section 3.2 to the sample paths of $v(t)$. Then we obtain the process $\xi(t)=\mathcal{S}v(t)$ having C^1-smooth sample paths. Recall that $\mathcal{S}\colon C^0([0,l],T_{m_0}M)\to C^1_{m_0}([0,l],M)$ is continuous. Since in addition $v(t)$ is non-anticipative with respect to $\mathcal{B}_t$, this proves the following:

Lemma 14.3 *The process $\mathcal{S}v(t)$ is non-anticipative with respect to $\mathcal{B}_t$.*

Consider a special case of the probability space $(\bar{\Omega},\bar{\mathcal{F}},\bar{\mathsf{P}})$ where $\bar{\Omega}=C^0([0,l],T_{m_0}M)$, $\bar{\mathcal{F}}$ is the σ-algebra generated by cylinder sets and the measure $\bar{\mathsf{P}}$ is the measure generated by a certain stochastic process in $T_{m_0}M$. In this case we shall deal with the family $\bar{\mathcal{B}}_t$ of σ-subalgebras of $\bar{\mathcal{F}}$ where for some t the σ-subalgebra $\bar{\mathcal{B}}_t$ is generated by cylinder sets with bases on $[0,t]$.

Lemma 14.4 *The process $\mathcal{S}v(t)$ is non-anticipative with respect to $\bar{\mathcal{B}}_t$.*

Proof. Indeed, if the curves $v_1(\cdot)$ and $v_2(\cdot)$ from $\bar{\Omega} = C^0([0,l], T_{m_0}M)$ coincide at all $t \in [0,l_0]$ where $0 < l_0 < l$, then $\mathcal{S}v_1(t)$ coincides with $\mathcal{S}v_2(t)$ for $t \in [0,l_0]$ by construction of the operator $\mathcal{S}$. By [83, Chapter III, Section 2] this is the assertion of Lemma 14.4. □

Consider the vector field $\dot{\xi}(t)$ along $\xi(t) = \mathcal{S}v(t)$.

Theorem 14.5

$$\mathbf{D}\dot{\xi}(t) = E_t^{\xi}(\Gamma_{t,0}b(t)).$$

Proof. From the properties of parallel translation and from the construction of $\xi(t)$ it follows that

$$\begin{aligned}&E_t^{\xi}(\Gamma_{t,t+\triangle t}\dot{\xi}(t+\triangle t) - \dot{\xi}(t))\\ &= E_t^{\xi}\left(\Gamma_{t,0}\left(\int_t^{t+\triangle t} b(s)\mathrm{d}s + \int_t^{t+\triangle t} B(s)\mathrm{d}w(s)\right)\right).\end{aligned}$$

Note that $\mathcal{N}_t^{\xi}$ is a σ-subalgebra in $\mathcal{P}_t^v$. Since the Itô integral $\int_t^{t+\triangle t} B(s)\mathrm{d}w(s)$ is a martingale with respect to $\mathcal{P}_t^v$, by the properties of conditional expectation we obtain that

$$E_t^{\xi}\left(\int_t^{t+\triangle t} B(s)\mathrm{d}w(s)\right) = 0.$$

The Theorem follows. □

Along a process $\xi(t) = \mathcal{S}v(t)$ as above we can define the covariant quadratic mean derivative of $\dot{\xi}(t)$ as follows. Introduce the notation $\triangle\dot{\xi}(t) = \Gamma_{t,t+\triangle t}\dot{\xi}(t+\triangle t) - \dot{\xi}(t)$ where (as above in this section) $\Gamma_{t,s}$ is the ordinary parallel translation along C^1-smooth curves.

Definition 14.6. The *quadratic mean derivative* of $\dot{\xi}(t)$ along $\xi(t) = \mathcal{S}v(t)$ on M at time t is an L^1 random element of the form

$$\mathbf{D}_2\dot{\xi}(t) = \lim_{\triangle t \downarrow 0} E_t^{\xi}\left(\frac{\triangle\dot{\xi}(t) \otimes \triangle\dot{\xi}(t)}{\triangle t}\right),$$

where $\otimes$ is the tensor product and $\Gamma_{t,s}$ is the ordinary parallel translation along C^1-smooth curves.

Theorem 14.7 *Consider a process $\xi(t) = \mathcal{S}v(t)$ with $v(t) = \int_0^t b(s)\mathrm{d}s + \int_0^t B(s)\mathrm{d}w(s)$ in $T_{m_0}M$ as above. Then $\mathbf{D}_2\dot{\xi}(t) = E_t^{\xi}(\Gamma_{t,0}(B(t)B^*(t)))$ where B^* is the adjoint operator.*

Proof. As in the proof of Theorem 14.5, from the properties of parallel translation and from the construction of $\xi(t)$ it follows that

$$E_t^{\xi}(\triangle\dot{\xi}(t)\otimes\triangle\dot{\xi}(t)) = E_t^{\xi}(\Gamma_{0,t}(\triangle v(t)\otimes\triangle v(t))),$$

where $\triangle v(t) = \int_t^{t+\triangle t} b(s)\mathrm{d}s + \int_t^{t+\triangle t} B(s)\mathrm{d}w(s)$. In addition from the properties of the Itô integral we obtain that in the expression $\triangle v(t)\otimes\triangle v(t)$ only the summand $(\int_t^{t+\triangle t} B(s)\mathrm{d}w(s))\otimes(\int_t^{t+\triangle t} B(s)\mathrm{d}w(s))$ is infinitesimal of the same order as $\triangle t$ while all other summand are infinitesimals of higher order than $\triangle t$. Now the Theorem follows from Definition 14.6 and from the properties of the Itô integral. □

The fact that $\xi(t)$ a.s. has C^1-smooth paths is equivalent to the equality $D_2\xi(t) = 0$ (see Theorem 8.12 and Lemma 9.16). Thus we are in a position to give the following:

Definition 14.8. The *Langevin equation* with force field (14.2) is the system

$$\begin{cases} \mathbf{D}\dot{\xi}(t) = \bar{\alpha}(t,\xi(t),\dot{\xi}(t)) \\ \mathbf{D}_2\dot{\xi}(t) = A(t,\xi(t),\dot{\xi}(t))\,A^*(t,\xi(t),\dot{\xi}(t)) \\ D_2\xi(t) = 0. \end{cases} \tag{14.6}$$

Let $\xi(t)$ be a stochastic process with values in M which is non-anticipative with respect to $\mathcal{B}_t$ and such that the sample trajectories of ξ are a.s. C^1-smooth and $\xi(0) = m_0 \in M$. Thus, as above, we can use the ordinary Riemannian parallel translation along the sample paths of ξ. Retaining the notation of Sections 3.2 and 11.8, we denote $\Gamma_{0,t}$ by Γ and so $\Gamma\bar{\alpha}(t,\xi(t),\dot{\xi}(t))$ and $\Gamma A(t,\xi(t),\dot{\xi}(t))$ are obtained by the parallel translation of $\bar{\alpha}(t,\xi(t),\dot{\xi}(t))$ and $A(t,\xi(t),\dot{\xi}(t))$, respectively, along $\xi(\cdot)$ from the point $\xi(t)$ to $\xi(0) = m_0$, where $\bar{\alpha}$ and A are the coefficients of force field (14.2). The processes $\Gamma\bar{\alpha}(t,\xi,\dot{\xi})$ and $\Gamma A(t,\xi,\dot{\xi})$ take values in $T_{m_0}M$ and $L(T_{m_0}M, T_{m_0}M)$, respectively, and their trajectories are a.s. continuous, for so are the fields $\bar{\alpha}(t,m,X)$ and $A(t,m,X)$. Since parallel translation preserves the Riemannian norm, it follows from (14.4) that

$$\left\|\Gamma\bar{\alpha}\left(t,\xi(t),\dot{\xi}(t)\right)\right\| + \left\|\Gamma A\left(t,\xi(t),\dot{\xi}(t)\right)\right\| < K\left(1+\left\|\Gamma\dot{\xi}(t)\right\|\right). \tag{14.7}$$

Lemma 14.9 *The processes $\Gamma\bar{\alpha}\big(t,\xi(t),\dot{\xi}(t)\big)$ and $\Gamma A\big(t,\xi(t),\dot{\xi}(t)\big)$ are non-anticipative with respect to $\mathcal{B}_t$.*

The lemma is a consequence of the fact that the parallel translation operator Γ is continuous on the space of C^1-curves equipped with the C^1-topology and of our assumptions that ξ is non-anticipative and the fields $\bar{\alpha}$ and A are both continuous.

By Lemma 14.9, we can define the process $z(t)$ in $T_{m_0}M$ as

$$z(t) = \int_0^t \Gamma\bar{\alpha}\left(\tau, \xi(\tau), \dot{\xi}(\tau)\right)\, \mathrm{d}\tau + \int_0^t \Gamma A\left(\tau, \xi(\tau), \dot{\xi}(\tau)\right)\, \mathrm{d}w(\tau), \tag{14.8}$$

where the second term on the right-hand side is the Itô integral. It is clear that $z(t)$ given by (14.8) is non-anticipative with respect to $\mathcal{B}_t$ and almost surely has continuous trajectories.

Setting $v(t) = z(t)$, we obtain the Langevin equation (14.6) in the integral form:

$$\xi(t) = \mathcal{S}\left(\int_0^t \Gamma\bar{\alpha}\left(\tau, \xi(\tau), \dot{\xi}(\tau)\right)\, \mathrm{d}\tau + \int_0^t \Gamma A\left(\tau, \xi(\tau), \dot{\xi}(\tau)\right)\, \mathrm{d}w(\tau) + \bar{C}\right). \tag{14.9}$$

Indeed, one can easily see that a process satisfying (14.9) also satisfies (14.6). On the other hand, taking into account the relationship between Newton's equation (11.2) and the integral equation (11.19), we see that a solution $\xi(t)$ of (14.9) is a stochastic trajectory of the system with the force field given by (14.2) and with the initial condition $\xi(0) = m_0$ and $\dot{\xi}(0) = \bar{C} \in T_{m_0}M$.

Definition 14.10. We say that (14.9) has a *weak solution* on $[0, l] \subset R$ with initial conditions $\xi(0) = m_0$, $\dot{\xi}(0) = C$ if there exist a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, an M-valued stochastic process $\xi(t)$ with a.s. C^1-smooth sample paths, defined on $(\Omega, \mathcal{F}, \mathsf{P})$ with initial conditions $\xi(0) = m_0$ and $\dot{\xi}(0) = C$ and a Wiener process $w(t)$ in $\mathbb{R}^n$, defined on $(\Omega, \mathcal{F}, \mathsf{P})$ and adapted to $\xi(t)$, such that for all $t \in [0, l]$ P-a.s. (14.9) is fulfilled.

Definition 14.11. We say that (14.9) has a *strong solution* on $[0, l] \subset R$ with initial conditions $\xi(0) = m_0$, $\dot{\xi}(0) = C$ if on every probability space $(\Omega, \mathcal{F}, \mathsf{P})$ which admits a Wiener process, and for any Wiener process $w(t)$ in $\mathbb{R}^n$, defined on $(\Omega, \mathcal{F}, \mathsf{P})$, there exists an M-valued stochastic process $\xi(t)$, non-anticipative with respect to $w(t)$ and having a.s. C^1-smooth sample paths, that is defined on $(\Omega, \mathcal{F}, \mathsf{P})$ with initial condition $\xi(0) = m_0$, such that for all $t \in [0, l]$ P-a.s. (14.9) is fulfilled.

Remark 14.12. Equation (14.9) involves the Itô integral $\int_0^t \Gamma A(\tau, \xi, \dot{\xi})\, \mathrm{d}w$. A similar equation involving the Stratonovich integral $\int_0^t \Gamma A(\tau, \xi, \dot{\xi}) \circ \mathrm{d}w$ is also well-defined.

The velocity hodograph equation corresponding to (14.9) is

$$v(t) = \int_0^t \Gamma\bar{\alpha}\left(t, \mathcal{S}v(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}v(t)\right)\, \mathrm{d}\tau + \int_0^t \Gamma A\left(t, \mathcal{S}v(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}v(t)\right)\, \mathrm{d}w(\tau) + \bar{C} \tag{14.10}$$

(cf. Section 11.8.1). It is clear that the fields $\varGamma\bar{\alpha}\left(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t)\right)$ and $\varGamma A\left(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t)\right)$ are defined along any curve $x(\cdot) \in C^0\big([0,l], T_{m_0}M\big)$ and continuous on the space $\mathbb{R} \times C^0\big([0,l], T_{m_0}M\big)$. By construction (see Section 3.2.1) and by the properties of parallel translation, we have

$$\left\| \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}\left(x(t)\right) \right\| = \|x(t)\|\,,$$

and, therefore, by (14.7),

$$\left\|\varGamma\bar{\alpha}\left(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t)\right)\right\| + \left\|\varGamma A\left(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t)\right)\right\| \le K\left(1 + \|x(t)\|\right). \tag{14.11}$$

Lemma 14.13 *$\varGamma\bar{\alpha}\left(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t)\right)$ and $\varGamma A\left(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t)\right)$ are non-anticipative with respect to the family $\bar{\mathcal{B}}_t$ from Lemma 14.4.*

The assertion of Lemma 14.13 clearly follows from the construction of $\varGamma\bar{\alpha}\left(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t)\right)$ and $\varGamma A\left(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t)\right)$ and from the properties of parallel translation, as well as from Lemma 14.4.

Equation (14.10) is an Itô stochastic differential equation of diffusion type on the linear space $T_{m_0}M$. Since Definitions 6.23 and 6.24 are valid for (14.10), we needn't introduce any special notions of strong and weak solutions for it.

It is clear that $v(t)$ and the Wiener process $w(t)$ in $T_{m_0}M$ satisfy (14.10) if and only if $\mathcal{S}v(t)$ (taking values in M) and $w(t)$ satisfy (14.9). Observe also that $\mathcal{S}v(t)$ is defined on the same probability space and has the same measurability properties with respect to $w(t)$ as $v(t)$. Thus, we have proved:

Theorem 14.14 *The process $v(t)$ is a strong (respectively, weak) solution of (14.10) if and only if $\mathcal{S}v(t)$ is a strong (respectively, weak) solution of (14.9).*

Remark 14.15. Let us specify a realization of $w(t)$ in $T_{m_0}M$. Applying to it the parallel translation along $\mathcal{S}v(\cdot)$, we obtain realizations of $w(t)$ in all spaces $T_{\mathcal{S}v(\cdot)}M$. These realizations give rise to a force field defined along the trajectory. One may also use the realizations to introduce the notion of a solution of (14.9) in a similar way to Definitions 7.92 and 7.93.

Theorem 14.16 *Assume that $\bar{\alpha}(t,m,X)$ and $A(t,m,X)$ are jointly continuous in all variables and satisfy (14.4). Then on $[0,l]$ there exists a weak solution of equation (14.9) for any initial conditions $\xi(0) = m_0$ and $\dot{\xi}(0) = \bar{C} \in T_{m_0}M$.*

Proof. First, we pass to (14.10), which is equivalent to (14.9). Note that (14.10) is a diffusion type equation on a vector space. Recall that this means that the coefficients of (14.10) depend on the past, i.e., on the entire trajectory on the interval $[0,t]$. As has been shown, $\varGamma\bar{\alpha}\left(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t)\right)$ and

$\Gamma A\left(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t)\right)$ are well-defined and continuous on $\mathbb{R}\times C^0([0,l], T_{m_0}M)$ Moreover, they satisfy (14.11), the linear growth condition and by Lemma 14.13 they are not anticipative with respect to the family of σ-subalgebras $\bar{\mathcal{B}}_t$. Thus, the standard existence theorem in linear spaces (see, e.g., [83, Chapter 3, Section 2] or [162, Section 19.3.8]) guarantees that a weak solution of (14.10) exists. To complete the proof we simply apply Theorem 14.14. □

The following results can also be proved by passing to (14.10) and applying the results of the standard theory of stochastic equations on vector spaces [83, 84, 162].

Theorem 14.17 *Let $\bar{\alpha}(t, m, X)$ and $A(t, m, X)$ be as in Theorem 14.16. Assume that the operator $A(t, m, X)$ is invertible for all t, m, and X. If a solution of the equation*

$$\xi(t) = \mathcal{S}\left(\int_0^t \Gamma A\left(\tau, \xi(\tau), \dot{\xi}(\tau)\right)\,\mathrm{d}w(\tau)\right) \tag{14.12}$$

is weakly unique, then so is a solution of (14.9).

Theorem 14.18 *Let $\bar{\alpha}(t, m, X)$ be jointly continuous in all variables, satisfy (14.4), and be such that the solution of the Cauchy problem for (11.2) is unique. In addition, let $A_\varepsilon(t, m, X)$, where $\varepsilon \in (0, \delta)$ and $\delta > 0$, be jointly continuous in ε, t, m and X, and satisfy (14.4) with K independent of ε. Assume also that:*

(i) $A_0 = 0$;
(ii) $\lim\limits_{\varepsilon\to 0} A_\varepsilon \to 0$ *uniformly on every compact subset of* $[0, l] \times TM$;
(iii) *a solution of the equation*

$$\xi(t) = \mathcal{S}\left(\int_0^t \Gamma\bar{\alpha}\left(\tau, \xi(\tau), \dot{\xi}(\tau)\right)\,\mathrm{d}\tau \right.$$
$$\left. + \int_0^t \Gamma A_\varepsilon\left(\tau, \xi(\tau), \dot{\xi}(\tau)\right)\,\mathrm{d}w(\tau) + \bar{C}_\varepsilon\right) \tag{14.13}$$

is weakly unique for some $\bar{C}_\varepsilon$ such that $\lim\limits_{\varepsilon\to 0} \bar{C}_\varepsilon = \bar{C}$.

Then the measures on $C^1_{m_0}([0, l], M)$ corresponding to the solutions of (14.13) weakly converge as $\varepsilon \to 0$ to the measure concentrated on the unique solution of (11.2).

Example 14.19. Let $A = \varepsilon I$, where I is the identity operator. Then it is clear that a solution of (14.12) is unique. Thus, for $\bar{\alpha}$ as before, the equation

$$\xi(t) = \mathcal{S}\left(\int_0^t \Gamma\bar{\alpha}\left(\tau, \xi(\tau), \dot{\xi}(\tau)\right)\,\mathrm{d}\tau + \varepsilon w(t) + C\right)$$

has a unique solution. If, for example, $\bar{\alpha}$ is locally Lipschitz in m and X, then Theorem 14.18 holds true for the latter equation.

Remark 14.20. Let β be a (possibly, non-holonomic) constraint on M. Employing the operators $\mathcal{S}^{\boldsymbol{\beta}}$ and $\Gamma^{\boldsymbol{\beta}}$ defined in Section 11.8.2, one can extend the notion of a Langevin equation to manifolds with constraints. In this case, the velocity hodograph equation in a fiber of β turns out to be very similar to (14.10). As a consequence, analogs of all results of this section hold true for the Langevin equations with constraints.

It is known that equation (14.9) has a strongly unique strong solution provided that the coefficients of the diffusion type equation (14.9) satisfy a Lipschitz type condition (see e.g., [66, 170] and Section 6.2.3). However, the existence of a strong solution is rather hard to prove in the general case where the coefficients involve the operators Γ and $\mathcal{S}$. The reason is that Γ and $\mathcal{S}$ are defined by means of parallel translation and, as a consequence, we have a condition imposed on the entire mechanical system, rather than just on the force field.

On the other hand, the existence can easily be verified for certain particular force fields. Here we consider three examples of such fields:

(i) The *drag force*:

$$\begin{aligned}\bar{\alpha}(t,m,X) &= \phi\,(t,\|X\|)\cdot\hat{a}_m(X),\\ A(t,m,X) &= \Psi\big(t,\|X\|\big)\cdot\hat{A}_m(X),\end{aligned}$$

where ϕ and Ψ are scalar functions, $\hat{a}$ is a $(1,1)$-tensor field with $\nabla\hat{a}=0$, and $\hat{A}$ is a field of operators $\hat{A}_m\colon T_mM\to L(T_mM)$ parallel along every curve in M. (Note that the equation $\nabla\hat{a}=0$ is a restriction of the same kind as that imposed on $\hat{A}$: the operators $\hat{a}_m\colon T_mM\to T_mM$ are parallel along every curve.) For example, one may take $\hat{a}=\pm I$ or, if M is an oriented two-dimensional manifold, then $\hat{a}_m$ may be a rotation by a fixed angle. The same operators can be taken as examples of $\hat{A}$ if we assume in addition that $\hat{A}_m(X)$ is independent of X (i.e., $\hat{A}_m$, regarded as a function of X, is constant).

(ii) A particular case of (i) involving friction and constant diffusion:

$$\bar{\alpha}(t,m,X) = -b(t)\cdot X, \qquad A(t,m,X)=\phi(t)\cdot\hat{A}_m,$$

where the friction coefficient $b\geq 0$ is a real-valued function of time and $\hat{A}$ is a $(1,1)$-tensor field with $\nabla\hat{A}=0$.

(iii) A force given in a “stationary coordinate system”. Let the mappings $\bar{\alpha}_{m_0}(t)\colon T_{m_0}M\to T_{m_0}M$ and $A_{m_0}(t)\colon T_{m_0}M\to L(T_{m_0}M)$, $t\in[0,l]$ be given. The operators $\bar{\alpha}$ and A at other points of the trajectory $\xi(t)$ are obtained by the parallel translation of $\bar{\alpha}_{m_0}$ and A_{m_0} along $\xi(\cdot)$. (See Section 11.9 for a mechanical interpretation of parallelism.)

Theorem 14.21 *Let $\bar{\alpha}$ and A be as in* (i)-(iii)*. Assume also that $\bar{\alpha}_{m_0}$ and A_{m_0} are Lipschitz in $X \in T_{m_0}M$ and satisfy* (14.4)*, the linear growth condition. Then* (14.9) *has a strongly unique strong solution on* $[0, l]$*.*

Proof. Under the hypotheses of the theorem, equation (14.10) on $T_{m_0}M$ is equivalent to the following:

$$v(t) = \int_0^t \bar{\alpha}\left(\tau, m_0, v(\tau)\right) \mathrm{d}\tau + \int_0^t A\left(\tau, m_0, v(\tau)\right) \mathrm{d}w(\tau) + \bar{C}. \qquad (14.14)$$

This equation has a strongly unique strong solution defined on $[0, l]$. The initial velocity $\bar{C}$ can be viewed as a random vector measurable with respect to the σ-algebra $\mathcal{B}_0$ [170, 66]. To finish the proof it suffices to apply Theorem 14.14. □

If $\bar{\alpha}$ and A are as in (ii), then the hypotheses of Theorem 14.21 are automatically satisfied, provided that b and ϕ are bounded. In this case, the solution $v(t)$ of (14.14) and the solution $\mathcal{S}v(t)$ of the Langevin equation are called the *Ornstein-Uhlenbeck velocity process* and *the Ornstein-Uhlenbeck coordinate process*, respectively.

The assumption that b and ϕ are bounded can be omitted in the hodograph equation for Ornstein-Uhlenbeck processes so that the velocity process exists on a random interval up to the explosion time (see Definition 6.32).

Recall that Ornstein-Uhlenbeck processes describe Brownian motion in a medium with a drag force. A detailed discussion can be found in [188]. Ornstein-Uhlenbeck processes on manifolds are also discussed in [154].

Let $v(t)$ be a solution of (14.14). Denote by $Ev(t)$ the mathematical expectation of $v(t)$ in $T_{m_0}M$.

Definition 14.22. The curve $\mathcal{S}(Ev(t))$ on M is said to be the *mathematical expectation* of the process $\mathcal{S}v(t)$. The function $E(Ev(t) - v(t))^2$ is called the *dispersion* of $\mathcal{S}v(t)$.

It is easy to see that for the system defined in (i) and, in particular, for (ii) the mathematical expectation of a solution of (14.9) satisfies (11.2).

Passing to the hodograph equation and applying standard results on equations in a vector space, we obtain the following theorem.

Theorem 14.23 *Under the assumptions of Theorem* 14.21*, the solutions of*

$$\begin{aligned}\xi(t) = \mathcal{S}\Bigg(&\int_0^t \varGamma\bar{\alpha}\left(\tau, \xi(\tau), \dot{\xi}(\tau)\right) \mathrm{d}\tau \\ &+ \varepsilon \int_0^t \varGamma A\left(\tau, \xi(\tau), \dot{\xi}(\tau)\right) \mathrm{d}w(\tau) + \varepsilon\bar{C}\Bigg)\end{aligned} \qquad (14.15)$$

converge as $\varepsilon \to 0$ to the solution of (11.19) *in the topology of the space*

$$\mathcal{S}\left(C^0([0, l], L^2(\varOmega, T_{m_0}M))\right).$$

The mathematical expectation of the solution of (14.15) *uniformly converges to the solution of* (11.19).

Here $L^2(\Omega, T_{m_0}M)$ is the space of square integrable maps from Ω to $T_{m_0}M$ and $\mathcal{S}(C^0([0,l], L^2(\Omega, T_{m_0}M)))$ is the image of the space of continuous curves in $L^2(\Omega, T_{m_0}M)$ given on $[0,l]$, under the mapping $\mathcal{S}$. Note that the convergence means that the dispersion of ξ converges uniformly to zero. Recall also that equations (11.19) and (11.2) are equivalent.

Remark 14.24. For the Langevin equation with constraint mentioned in Remark 14.20, the Ornstein-Uhlenbeck processes with constraint can be defined analogously as strong solutions of the corresponding hodograph equations.

14.3 Set-Valued Forces. Langevin Type Inclusions

In this section we investigate second order stochastic differential inclusions on Riemannian manifolds which are set-valued analogs of the Langevin equations from Section 14.2. As mentioned above, the set-valued force arises in a system with control or may be obtained from a discontinuous force (for instance, if dry friction is considered or if the motion takes place in a complicated medium). If the force is discontinuous there are well-known methods of transition to a set-valued force (for stochastic differential equations the pioneering paper was [41]). Examples of systems having discontinuous forces with random components of the above-mentioned sort are rather common in physics, for example, they describe the motion of a physical Brownian particle in a complicated medium. The use of Riemannian manifolds allows one to cover mechanics on non-linear configuration spaces.

In this Section we use the set-valued vector force fields and set-valued tensor force fields introduced in Definition 11.43 and Definition 14.1, respectively.

Let $\boldsymbol{\alpha}$ and $\mathbf{A}$ be a set-valued vector force field and a set-valued tensor field, respectively. For a stochastic process $\xi(t)$ with a.s. C^1-smooth sample paths consider the set-valued maps $\Gamma\boldsymbol{\alpha}(\tau, \xi(\tau), \dot{\xi}(\tau))$ and $\Gamma\mathbf{A}(\tau, \xi(\tau), \dot{\xi}(\tau))$ sending $[0,l]$ into $T_{m_0}M$ and into the space of linear operators on $T_{m_0}M$ and denote by $\mathcal{P}\Gamma\boldsymbol{\alpha}(\tau, \xi(\tau), \dot{\xi}(\tau))$ and $\mathcal{P}\Gamma\mathbf{A}(\tau, \xi(\tau), \dot{\xi}(\tau))$ the sets of their Borel measurable selectors.

A *Langevin inclusion* is a system of the form

$$\begin{cases} \mathbf{D}\dot{\xi}(t) \in \boldsymbol{\alpha}(t, \xi(t), \dot{\xi}(t)) \\ \mathbf{D}_2\dot{\xi}(t) \in \mathbf{A}(t, \xi(t), \dot{\xi}(t))\, \mathbf{A}^*(t, \xi(t), \dot{\xi}(t)) \\ D_2\xi(t) = 0 \end{cases} \tag{14.16}$$

where $\mathbf{D}$ and $\mathbf{D}_2$ are defined in Section 14.2 by means of ordinary parallel translation along C^1-smooth curves. In integral form (14.16) is expressed as

$$\xi(t) \in \mathcal{S}\left(\int_0^t \mathcal{P}\Gamma\boldsymbol{\alpha}\left(\tau,\xi(\tau),\dot{\xi}(\tau)\right)\,\mathrm{d}\tau + \int_0^t \mathcal{P}\Gamma\mathbf{A}\left(\tau,\xi(\tau),\dot{\xi}(\tau)\right)\,\mathrm{d}w(\tau) + C\right). \tag{14.17}$$

Definition 14.25. We say that (14.16) has a *weak solution* on $[0,l] \subset \mathbb{R}$ with initial conditions $\xi(0) = m_0$, $\dot{\xi}(0) = C$ if there exist a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, an M-valued stochastic process $\xi(t)$ with a.s. C^1-smooth sample paths, defined on $(\Omega, \mathcal{F}, \mathsf{P})$ with initial conditions $\xi(0) = m_0$ and $\dot{\xi}(0) = C$, a Wiener process $w(t)$ in $\mathbb{R}^n$, defined on $(\Omega, \mathcal{F}, \mathsf{P})$ and adapted to $\xi(t)$, a single-valued vector field $\bar{\alpha}(t,m,X)$ on M and a single-valued $(1,1)$-tensor field $A(t,m,X)$ such that:

(i) for all t the random vector $\bar{\alpha}(t,\xi(t),\dot{\xi}(t))$ belongs to $\boldsymbol{\alpha}(t,\xi(t),\dot{\xi}(t))$ P-a.s.;
(ii) for all t the random tensor $A(t,\xi(t),\dot{\xi}(t))$ belongs to $\mathbf{A}(t,\xi(t),\dot{\xi}(t))$ P-a.s.;
(iii) the integrals $\int_0^t \Gamma\bar{\alpha}(\tau,\xi(\tau),\dot{\xi}(\tau))\mathrm{d}\tau$ and $\int_0^t \Gamma A(\tau,\xi(\tau),\dot{\xi}(\tau))\mathrm{d}w(\tau)$ are well-defined for $\xi(t)$, $w(t)$, $\bar{\alpha}$ and A,

and for all $t \in [0,l]$ P-a.s.

$$\xi(t) = \mathcal{S}\left(\int_0^t \Gamma\bar{\alpha}\left(\tau,\xi(\tau),\dot{\xi}(\tau)\right)\,\mathrm{d}\tau + \int_0^t \Gamma A\left(\tau,\xi(\tau),\dot{\xi}(\tau)\right)\,\mathrm{d}w(\tau) + C\right). \tag{14.18}$$

Definition 14.26. We say that (14.16) has a *strong solution* on $[0,l] \subset \mathbb{R}$ with initial conditions $\xi(0) = m_0$, $\dot{\xi}(0) = C$ if on any probability space $(\Omega, \mathcal{F}, \mathsf{P})$ which admits a Wiener process, and for any Wiener process $w(t)$ in $\mathbb{R}^n$, defined on $(\Omega, \mathcal{F}, \mathsf{P})$, there exist: a stochastic process $\xi(t)$ with a.s. C^1-smooth sample paths in M, defined on $(\Omega, \mathcal{F}, \mathsf{P})$ and non-anticipating with respect to $w(t)$ with initial condition $\xi(0) = m_0$ and $\dot{\xi}(0) = C$, a single-valued vector field $\bar{\alpha}(t,m,X)$ on M and a single-valued $(1,1)$-tensor field $A(t,m,X)$ such that:

(i) for all t the random vector $\bar{\alpha}(t,\xi(t),\dot{\xi}(t))$ belongs to $\boldsymbol{\alpha}(t,\xi(t),\dot{\xi}(t))$ P-a.s.;
(ii) for all t the random tensor $A(t,\xi(t),\dot{\xi}(t))$ belongs to $\mathbf{A}(t,\xi(t),\dot{\xi}(t))$ P-a.s.;
(iii) the integrals $\int_0^t \Gamma\bar{\alpha}(\tau,\xi(\tau),\dot{\xi}(\tau))\mathrm{d}\tau$ and $\int_0^t \Gamma A(\tau,\xi(\tau),\dot{\xi}(\tau))\mathrm{d}w(\tau)$ are well-defined for $\xi(t)$, $w(t)$, $\bar{\alpha}$ and A and P-a.s. (14.18) holds for all $t \in [0,l]$.

As in Section 14.2 one can easily prove that $\xi(t)$ as above satisfies (14.18) if and only if its velocity hodograph $v(t)$ (i.e., $v(t)$ given by the relation $\xi(t) = \mathcal{S}v(t)$) satisfies the velocity hodograph equation of the form

$$v(t) = \int\limits_0^t \Gamma \bar{\alpha}\left(\tau, \mathcal{S}v(\tau), \frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{S}v(\tau)\right) \mathrm{d}\tau + \int\limits_0^t \Gamma A\left(\tau, \mathcal{S}v(\tau), \frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{S}v(\tau)\right) \mathrm{d}w(\tau) + C \tag{14.19}$$

which is an equation of diffusion type in the tangent (i.e., linear) space at m_0 and hence is easier to study. Below, we shall find $\bar{\alpha}$ and A as in Definitions 14.25 and 14.26 and a corresponding $v(t)$, being a solution of (14.19) in the weak or strong sense, and then obtain $\xi(t) = \mathcal{S}v(t)$ satisfying (14.18).

If both $\boldsymbol{\alpha}$ and $\mathbf{A}$ have continuous selectors satisfying the Itô condition (see (14.20) below), the existence of a weak solution trivially follows from that for the Langevin equation obtained in Section 14.2. If this is not the case, the existence problem for Langevin inclusions requires special constructions.

For vector and tensor set-valued force fields on manifolds we use the following modification of Definition 3.49:

Definition 14.27. A continuous single-valued force field $\bar{\alpha}_\varepsilon(t, m, X)$ is called an ε-approximation of the set-valued force field $\boldsymbol{\alpha}(t, m, X)$ on M if its graph $(t, m, X, \bar{\alpha}_\varepsilon(t, m, X))$ lies in an ε-neighborhood of $(t, m, X, \boldsymbol{\alpha}(t, m, X))$ (the graph of $\boldsymbol{\alpha}$) in $[0, l] \times TM \oplus TM$, where $\oplus$ denotes the Whitney sum. For $(1, 1)$-tensor fields the definition is analogous.

One can easily see that the natural analog of Theorem 4.11 holds for both vector and $(1, 1)$-tensor force fields.

We say that $\boldsymbol{\alpha}$ and $\mathbf{A}$ satisfy the Itô condition if they have linear growth in velocities, i.e., there exists a $\Theta > 0$ for which the following inequality holds:

$$\|\boldsymbol{\alpha}(t, m, X)\| + \|\mathbf{A}(t, m, X)\| < \Theta(1 + \|X\|). \tag{14.20}$$

Theorem 14.28 *Let the set-valued force field $\boldsymbol{\alpha}(t, m, X)$ and set-valued $(1, 1)$-tensor field $\mathbf{A}(t, m, X)$ be upper semi-continuous with convex bounded closed values and satisfy the Itô condition (14.20) for some Θ.*

Then for any $m_0 \in M$ and $C \in T_{m_0}M$ the Langevin inclusion (14.16) has a weak solution with initial conditions $\xi(0) = m_0$, $\dot{\xi}(0) = C$, well-defined for all $t \in [0, \infty)$.

Proof. Let $l > 0$. Denote by $\mathcal{B}$ the Borel σ-algebra on $[0, l]$ and by λ the normalized Lebesgue measure on it. Here we use the following notation: $\tilde{\Omega} = C^0([0, l], T_{m_0}M)$ is the Banach space of continuous curves $x : [0, l] \to T_{m_0}M$ with the usual uniform norm and $\tilde{\mathcal{F}}$ is the σ-algebra generated by cylindrical sets on $\tilde{\Omega}$. By $\tilde{\mathcal{P}}_t$ we denote the σ-algebra generated by cylinder sets with bases over $[0, t]$ (cf. Section 6.1.1).

We shall use several measures on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ and on the product space $[0, l] \times \tilde{\Omega}$ we shall introduce the corresponding product measures.

Take a sequence $\varepsilon_i \to 0$ and construct sequences $f_i(t, m, X)$ and $a_i(t, m, X)$ of continuous ε_i-approximations of $F(t, m, X)$ and $A(t, m, X)$, respectively, as in Theorem 4.11. In particular, denote by $\Psi_i(t,m,X)$ a continuous set-valued force field with convex closed values whose graph belongs to the ε_i-neighborhood of the graph of $F(t,m,X)$ and such that for all (t,m,X) the inclusion $F(t,m,X) \subset \Psi_i(t,m,X)$ holds (the existence of such $\Psi_i(t,m,X)$ follows from [79]). Then as in the proof of Theorem 4.11 the minimal selectors $f_i(t,m,X)$ of $\Psi_i(t,m,X)$ point-wise converge to the minimal selector $f(t,m,X)$ of $F(t,m,X)$ as $i \to \infty$ and $f(t,m,X)$ is Borel measurable. By an analogous argument we introduce a continuous $(1,1)$-tensor field $\hat{\Psi}_i(t,m,X)$ whose graph belongs to the ε_i-neighborhood of the graph of $A(t,m,X)$ and such that for all (t,m,X) the inclusion $A(t,m,X) \subset \hat{\Psi}_i(t,m,X)$ holds. The minimal selectors $a_i(t,m,X)$ of $\hat{\Psi}_i(t,m,X)$ point-wise converge to the minimal selector $a(t,m,X)$ of $A(t,m,X)$ as $i \to \infty$ and $a(t,m,X)$ is Borel measurable.

Taking into account Definition 14.27 and inequality (14.20) we have

$$\|f_i(t,m,X)\| + \|a_i(t,m,X)\| < Q(1 + \|X\|)$$

for a certain $Q > \Theta$ and for all i.

Pass from the sequences $f_i(t,m,X)$ and $a_i(t,m,X)$ to the sequences $\tilde{f}_i : [0,l] \times \tilde{\Omega} \to TM$ and $\tilde{a}_i : [0,l] \times \tilde{\Omega} \to TM$, where $\tilde{f}_i(t,x(\cdot)) = f_i(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t))$ and $\tilde{a}_i(t,x(\cdot)) = a_i(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t))$. In addition, introduce $\tilde{f}(t,x(\cdot)) = f(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t))$ and $\tilde{a}(t,x(\cdot)) = a(t, \mathcal{S}x(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t))$.

Consider the maps $\Gamma \tilde{f}_i(t,x(\cdot))$ from $[0,l] \times \tilde{\Omega}$ into $T_{m_0}M$ and $\Gamma \tilde{a}_i(t,x(\cdot))$ from $[0,l] \times \tilde{\Omega}$ into the set of linear endomorphisms on $T_{m_0}M$. One can easily see that $\Gamma \tilde{f}_k(t,x(\cdot))$ point-wise converges to $\Gamma \tilde{f}(t,x(\cdot))$ and $\Gamma \tilde{a}_k(t,x(\cdot))$ point-wise converges to $\Gamma \tilde{a}(t,x(\cdot))$ as $k \to \infty$.

Since $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}x(t)$ is by construction parallel to $x(t)$ along $\mathcal{S}x(\cdot)$ and the parallel translation preserves the norms, we get

$$\|\Gamma \tilde{f}_i(t,x(\cdot))\| + \|\Gamma \tilde{a}_i(t,x(\cdot))\| < Q(1 + \|x(\cdot)\|). \tag{14.21}$$

By construction, $\Gamma \tilde{f}_i(t,x(\cdot))$ and $\Gamma \tilde{a}_i(t,x(\cdot))$ are continuous on $[0,l] \times \tilde{\Omega}$ (this follows from the continuity of Γ, see [99, 106, 107, 115]) and measurable with respect to the σ-subalgebra $\mathcal{P}_t$ in $\mathcal{F}$ generated by cylindrical sets with bases over $[0,t]$. Since it also satisfies (14.21), by Theorem 6.26 there exists a weak solution $v_i(t)$ of the equation

$$v_i(t) = \int_0^t \Gamma \tilde{f}_i(\tau, v_i(\cdot))\mathrm{d}\tau + \int_0^t \Gamma \tilde{a}_i(\tau, v_i(\cdot))\mathrm{d}w(t) + C \tag{14.22}$$

Denote by μ_i the measure on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ corresponding to v_i. Recall that $v_i(t)$ is represented as the coordinate process $v_i(t,x(\cdot)) = x(t)$ on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu_i)$.

By Lemma 6.28, since all $\Gamma\tilde{f}_i(t,x(\cdot))$ and $\Gamma\tilde{a}_i(t,x(\cdot))$ satisfy (14.21) with the same Q, one can show that the set of measures $\{\mu_i\}$ is weakly compact and so there exists a subsequence converging weakly to some probability measure μ on $(\tilde{\Omega},\tilde{\mathcal{F}})$. For the sake of convenience we use the same notation and say that μ_i itself is the converging subsequence. Denote by $v(t)$ the coordinate process on $(\tilde{\Omega},\tilde{\mathcal{F}},\mu)$.

The fact that the process $v(t) - \int_0^t \Gamma\tilde{f}(s,v(s))\mathrm{d}s$ is a martingale on $(\tilde{\Omega},\tilde{\mathcal{F}},\mu)$ with respect to $\tilde{\mathcal{P}}_t$ is proved by analogy with Lemma 8.47.

Choose an orthonormal basis in $T_{m_0}M$. Then the vectors in $T_{m_0}M$ are considered as coordinate columns. If X is such a vector, the transposed row vector is denoted by X^*. Note that for a column X and a row Y^* the product XY^* with respect to matrix multiplication is a matrix. Linear operators from $T_{m_0}M$ to $T_{m_0}M$ are represented in coordinates as $n\times n$ matrices, where the symbol $*$ denotes conjugate transposition.

Consider the sequence $a_i(t,m,X)$ of ε_i-approximations of $A(t,m,X)$ that point-wise converges to the Borel-measurable selector $a(t,m,X)$ (see the beginning of this proof). One can easily see that $a_i(t,m,X)(a_i(t,m,X))^*$ point-wise converges to $a(t,m,X)(a(t,m,X))^*$. Then by analogy with the proof of Lemma 8.48 one can show that

$$v(t)v(t)^* - \int_0^t \Gamma\tilde{a}(s,x(\cdot))(\Gamma\tilde{a}(s,x(\cdot)))^*\mathrm{d}s$$

is a martingale on $(\tilde{\Omega},\tilde{\mathcal{F}},\mu)$ with respect to $\tilde{\mathcal{P}}$ and so

$$E\Bigg((x(t+\Delta t)-x(t))(x(t+\Delta t)-x(t))^* - \\ - \int_t^{t+\Delta t} \Gamma\tilde{a}(s,x(\cdot))(\Gamma\tilde{a}(s,x(\cdot)))^*\mathrm{d}s \mid \tilde{\mathcal{P}}_t\Bigg) = 0. \tag{14.23}$$

Using the standard Girsanov technique one can derive from the above arguments that on $(\tilde{\Omega},\tilde{\mathcal{F}},\mu)$ there exists a Wiener process $w(t)$, adapted to $\mathcal{P}_t$, such that $v(t)$ on $(\tilde{\Omega},\mathcal{F},\mu)$ satisfies the equality

$$v(t) = C + \int_0^t \Gamma\tilde{f}(s,v(\cdot))\mathrm{d}s + \int_0^t \Gamma\tilde{a}(s,v(\cdot))\mathrm{d}w(s) \tag{14.24}$$

(see [83]). Then, taking into account the construction of $\tilde{f}$ and the operators $\mathcal{S}$ and Γ, one can easily see that the process $\xi(t) = \mathcal{S}v(t)$ satisfies the equation

$$\xi(t) = \mathcal{S}\left(\int\limits_0^t \Gamma f\left(s, \xi(s), \frac{\mathrm{d}}{\mathrm{d}s}\xi(s)\right)\mathrm{d}s + \int\limits_0^t \Gamma a\left(s, \xi(s), \frac{\mathrm{d}}{\mathrm{d}s}\xi(s)\right)\mathrm{d}w(s) + C\right) \tag{14.25}$$

Since $f(t, m, X) \in F(t, m, X)$ and $a(t, m, X) \in A(t, m, X)$ and $l > 0$ is an arbitrary number, this completes the proof. □

In some cases we can prove the existence of a strong solution of the Langevin inclusion (14.16). Let us present an example of such an existence theorem.

In what follows we use $[0, l]$, $\mathcal{B}$, $\tilde{\Omega}$, $\mathcal{F}$ and $\mathcal{P}_t$ as introduced in the proof of Theorem 14.28. By $\mathcal{B}_t$ we denote the Borel σ-algebra on $[0, t]$ for $t \in [0, l]$.

We introduce the notation $\mathrm{comp}Z$ for the space of compact subsets in the metric space Z. Thus, we say that the set-valued force vector field $B(t, m, X)$ sends $[0, l] \times TM$ into $\mathrm{comp}\, TM$ if for any $(t, m, X) \in [0, l] \times TM$ the image $B(t, m, X) \subset T_m M$ is compact.

We recall several definitions.

Definition 14.29. A single-valued map $\beta : [0, l] \times \tilde{\Omega} \to \mathbb{R}^n$ is called $\{\mathcal{P}_t\}$-*progressively measurable* if for every t it is measurable with respect to $\mathcal{B}_t \times \mathcal{P}_t$.

Definition 14.30. A set-valued map $B : [0, l] \times \tilde{\Omega} \to \mathrm{comp}\mathbb{R}^n$ is called $\{\mathcal{P}_t\}$-*progressively measurable* if $\{(t, \omega) \in [0, l] \times \tilde{\Omega} \mid B(t, \omega) \cap C \neq \emptyset\} \in \mathcal{B}_t \times \mathcal{P}_t$ for every closed set $\mathbf{K} \subset \mathbb{R}^n$.

Definition 14.31. We say that a set-valued vector force field $B : [0, l] \times TM \to \mathrm{comp}\, TM$:

(i) is *dissipative* if for all $t \in [0, l]$, $m \in M$, $X, Y \in T_m M$ and $U \in B(t, m, X)$, $V \in B(t, m, Y)$ the inequality $\langle X - Y, U - V\rangle \leq 0$ holds.
(ii) is *maximal* if for t, m, X, Y and V as in (i) the inequality $\langle X - Y, U - V\rangle \leq 0$ is equivalent to the assumption that $U \in B(t, m, X)$.

Denote by $w(t)$ a one-dimensional Wiener process. Let $F(t, m, X)$ and $G(t, m, X)$ be set-valued vector force fields on M as above. Then we can consider the stochastic differential inclusion of Langevin type

$$\xi(t) \in \mathcal{S}\left(\int_0^t \mathcal{P}\Gamma F(\tau, \xi(\tau), \dot{\xi}(\tau))\mathrm{d}\tau + \int_0^t \mathcal{P}\Gamma G(\tau, \xi(\tau), \dot{\xi}(\tau))\mathrm{d}w(\tau) + C\right). \tag{14.26}$$

Inclusion (14.26) is a particular case of (14.17) since $\Gamma G(\tau, \xi(\tau), \dot{\xi}(\tau))\mathrm{d}w(\tau)$ can be represented as $\Gamma G(\tau, \xi(\tau), \dot{\xi}(\tau))(P\mathrm{d}W(\tau))$ (P is the orthogonal projection onto the linear span of vectors $\Gamma G(\tau, \xi(\tau), \dot{\xi}(\tau))$).

Theorem 14.32 *Let the set-valued vector fields $F(t, m, X)$ and $G(t, m, X)$, $F, G : [0, l] \times TM \to \mathrm{comp}\, TM$ be Borel measurable, uniformly bounded, dissipative and maximal. Then there exists a strong solution of (14.26), well-defined for $t \in [0, l]$, with initial conditions $\xi(0) = m_0$ and $\dot{\xi}(0) = C$ for any $m_0 \in M$ and $C \in T_{m_0} M$.*

Proof. Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space admitting a one-dimensional Wiener process $w(t)$. Denote by $\mathcal{P}_t^w$ the σ-subalgebra of $\mathcal{F}$ generated by all $w(s)$ for $0 \leq s \leq t$ and completed by all sets of zero probability. Let $Y : \Omega \to \Omega$ be a measurable map. From the properties of parallel translation and the assumed hypothesis one can easily derive that the coefficients

$$\Gamma F(t, \omega, Y) = \Gamma F \left(t, \mathcal{S}Y(\omega)(t), \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}Y(\omega)(t) \right)$$

and

$$\Gamma G(t, \omega, Y) = \Gamma G \left(t, \mathcal{S}Y(\omega)(t), \frac{d}{\mathrm{d}}\mathrm{d}t\mathcal{S}Y(\omega)(t) \right)$$

for $\omega \in \Omega$ satisfy all the conditions of [183, Theorem 1] and so on $(\Omega, \mathcal{F}, \mathsf{P})$ there exists a continuous $\mathcal{P}_t^w$-progressively measurable process $v(t)$ $(v(0) = 0)$ in $T_{m_0}M$ and L^2-selectors $f(t, \omega)$ of $\Gamma F(t, \omega, v)$ and $g(t, \omega)$ of $\Gamma G(t, \omega, v)$ such that a.s.

$$v(t) = \int_0^t f(\tau, \omega)\mathrm{d}\tau + \int_0^t g(\tau, \omega)\mathrm{d}w(\tau) + C. \tag{14.27}$$

Consider the M-valued process $\xi(t) = \mathcal{S}v(t)$ with $v(t)$ satisfying (14.27). In the same manner as in the proof of Theorem 14.28 we can construct Borel measurable selectors $f(t, m, X)$ of $F(t, m, X)$ and $g(t, m, X)$ of $G(t, m, X)$ such that a.s.

$$\xi(t) = \mathcal{S} \left(\int_0^t \Gamma f(\tau, \xi(\tau), \dot{\xi}(\tau))\mathrm{d}\tau + \int_0^t \Gamma g(\tau, \xi(\tau), \dot{\xi}(\tau))\mathrm{d}w(\tau) + C \right).$$

□

14.4 Systems with Random Perturbation of Velocity

In the previous two sections we dealt with equations obtained from the ordinary version of Newton's law by a stochastic perturbation of the vertical component on the right-hand side, i.e., of the force field (see Section 14.1). Here we investigate the systems in which the horizontal part is subjected to stochastic influence. This means that a random perturbation of velocity arises. Such a situation can appear, for example, if a particle, subjected to a deterministic force, in addition moves within a random media. Note that in this model the perturbation is independent of the particle velocity.

We investigate the system (14.1) with $\bar{\alpha}$ independent of velocities, i.e., it turns into

$$\begin{cases} \dot{x}(t) = v(t) \\ \dot{v}(t) = \bar{\alpha}(t, x(t)). \end{cases} \tag{14.28}$$

A particular example of such a force is $-\mathrm{grad}\,\mathcal{U}$ in a conservative mechanical system.

Now suppose that in (14.28) the right-hand side of the horizontal (i.e. first) equation is subjected to a random perturbation of the form $A(t, x(t))\dot{w}(t)$ where $\dot{w}(t)$ is white noise. Note that this perturbation is independent of the velocity of the particle. In appropriate terms this means that the process $\xi(t)$ describing the motion of the particle satisfies the equality $\xi(t) = \xi_0 + \int_0^t v(s, \xi(s))\mathrm{d}s + \int_0^t A(s, \xi(s))\mathrm{d}w(s)$ where the vector field $v(t, x)$ satisfies the relation $Dv(t, \xi(t)) = \bar{\alpha}(t, \xi(t))$. The formal equation of motion in terms of forward mean derivatives then takes the form

$$\begin{cases} D\xi(t) = v(t, \xi(t)) \\ D_2\xi(t) = A(t, \xi(t))A^*(t, \xi(t)) \\ Dv(t, \xi(t)) = \bar{\alpha}(t, \xi(t)). \end{cases} \tag{14.29}$$

We also suppose that $A(t, x)$ and $\bar{\alpha}(t, x))$ satisfy the Itô condition

$$\|A(t, x)\| + \|\bar{\alpha}(t, x)\| < K(1 + \|x\|) \tag{14.30}$$

for some $K > 0$.

Theorem 14.33 *Let $A(t, x)$ and $\bar{\alpha}(t, x)$ be jointly continuous in t, x and satisfy (14.30). Then for every pair ξ_0, $v_0 \in \mathbb{R}^n$ there exists a weak solution of (14.29) with initial conditions $\xi(0) = \xi_0$ and $v(0) = v_0$.*

Proof. In $C^0([0, l], \mathbb{R}^n)$ introduce the σ-algebra $\tilde{\mathcal{F}}$ generated by cylindrical sets. By $\tilde{\mathcal{P}}_t$ denote the σ-algebra generated by cylindrical sets over $[0, t] \subset [0, l]$.

Consider the map $\bar{v} : [0, l] \times C^0([0, l], \mathbb{R}^n) \to \mathbb{R}^n$ defined by the formula

$$\bar{v}(t, x(\cdot)) = v_0 + \int_0^t \bar{\alpha}(\tau, x(\cdot)))\mathrm{d}\tau. \tag{14.31}$$

By construction this map is jointly continuous in $t \in [0, l]$ and $x(\cdot) \in C^0([0, l], \mathbb{R}^n)$. In addition it is obvious that if $x_1(\cdot)$ and $x_2(\cdot)$ coincide on $[0, t]$ then $\bar{v}(t, x_1(\cdot)) = \bar{v}(t, x_2(\cdot))$. This means that $\bar{v}(t, x(\cdot))$ is measurable with respect to $\tilde{\mathcal{P}}_t$ (see, e.g., [83]).

Taking into account (14.30) one can easily derive the inequality

$$\begin{aligned} \|\bar{v}(t, x(\cdot))\| &= \left\| \int_0^t \bar{\alpha}(\tau, x(\cdot)))\mathrm{d}\tau \right\| \\ &\le \int_0^t \|\bar{\alpha}(\tau, x(\cdot)))\|\mathrm{d}\tau \\ &\le K \int_0^t (1 + \|x(\tau)\|)\mathrm{d}\tau \end{aligned}$$

$$\leq K \int_0^t (1 + \|x(\cdot)\|_{C^0}) \mathrm{d}s$$
$$\leq lK(1 + \|x(\cdot)\|_{C^0}),$$

where $\|\cdot\|_{C^0}$ is the norm in $C^0([0,l],\mathbb{R}^n)$.

Introduce $A(t,x(\cdot))$ as $A(t,x(\cdot)) = A(t,x(t))$. Notice that $A(t,x(\cdot))$ is measurable with respect to $\tilde{\mathcal{P}}_t$ and that from (14.30) it follows that $\|A(t,x(\cdot))\| \leq K(1+\|x(\cdot)\|_{C^0})$. So, both $\bar{v}(t,x(\cdot))$ and $A(t,x(\cdot))$ satisfy the Itô condition in the form

$$\|\bar{v}(t,x(\cdot))\| + \|A(t,x(\cdot)\| \leq \bar{K}(1 + \|x(\cdot)\|_{C^0})$$

with $\bar{K} = \max(K, lK)$.

Now the pair $\bar{v}(t,x(\cdot))$ and $A(t,x(\cdot))$ satisfies all the conditions of [83, Theorem III.2.4], hence, the stochastic differential equation

$$x(t) = x_0 + \int_0^t \bar{v}(s,x(\cdot))\mathrm{d}s + \int_0^t A(s,x(\cdot))\mathrm{d}w(s) \tag{14.32}$$

has a weak solution on $[0,l]$. This means that there exist a probabilistic measure μ on $(C^0([0,l],\mathbb{R}^n),\ \mathcal{F})$ and a Wiener process in $\mathbb{R}^n$, given on $(C^0([0,l],\mathbb{R}^n),\mathcal{F},\mu)$ and adapted to $\mathcal{P}_t$, such that the coordinate process $x(t)$ on $(C^0([0,l],\mathbb{R}^n),\mathcal{F},\mu)$ and $w(t)$ satisfy (14.32). Let $v(t,x)$ be the regression $v(t,x) = E(\bar{v}(t,x(\cdot)) \mid x(t) = x)$. This together with the construction of the process $\bar{v}(t,x(\cdot))$ completes the proof of the Theorem. $\square$

The simple construction used in the proof of Theorem 14.33 can be generalized so that it is applicable in more complicated settings. First we consider the case where the force field is set-valued, lower semi-continuous but not necessarily convex valued.

Let $F(t,x)$ be a lower semi-continuous set-valued map $F : \mathbb{R}\times\mathbb{R}^n \multimap \mathbb{R}^n$ with closed images and $A(t,x) : \mathbb{R}^n \to \mathbb{R}^n$ be a field of single-valued linear operators jointly continuous in parameters $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$. We suppose that $F(t,x)$ and $A(t,x)$ satisfy the Itô condition, i.e., that there exists a constant $\Theta > 0$ such that

$$\|F(t,x)\| + \|A(t,x)\| < \Theta(1 + \|x\|) \tag{14.33}$$

for all $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$ where $\|A(t,x)\|$ is the operator norm and $\|F(t,x)\| = \sup_{y\in F(t,x)} \|y\|$.

The system of equations (14.30) is now replaced by the following analogous inclusion

$$\begin{cases} D\xi(t) = v(t,\xi(t)) \\ D_2\xi(t) = A(t,\xi(t))A^*(t,\xi(t)) \\ Dv(t,\xi(t)) \in F(t,x(t)). \end{cases} \tag{14.34}$$

In what follows we consider $\mathbb{R}^n$ and $\mathbb{R}$ with their Borel σ-algebras $\mathcal{B}^n$ and $\mathcal{B}$, respectively. Let $x(\cdot)$ be a continuous curve. Consider the set-valued vector field $F(t, x(t))$ along $x(\cdot)$ and denote by $\mathcal{P}F(\cdot, x(\cdot))$ the set of all measurable selectors of $F(t, x(t))$, i.e., the set of measurable maps $\{f : \mathbb{R} \to \mathbb{R}^n : f(x(t)) \in F(t, x(t))\}$. It is clear that, since condition (14.33) is satisfied, all such selectors are integrable on any finite interval in $\mathbb{R}$ with respect to Lebesgue measure. Denote by $\int \mathcal{P}F(\cdot, x(\cdot))$ the set of integrals with varying upper limits of these selectors.

We recall some facts and notions which will be used shortly. Let $l > 0$. In what follows we denote by λ the normalized Lebesgue measure on $[0, l]$, i.e., such that $\lambda([0, l]) = 1$.

Lemma 14.34 *Let (Ξ, d) be a separable metric space, X be a Banach space. Consider the space $Y = L^1(([0, l], \mathcal{B}, \lambda), X))$ of integrable maps from $[0, l]$ into X. If a set-valued map $G : \Xi \to Y$ is lower semicontinuous and has closed decomposable images, then it has a continuous selector.*

This is a particular case of the Bressan-Colombo Theorem (Theorem 4.9).

Denote by $C^0([0, l], \mathbb{R}^n)$ the Banach space of continuous maps from $[0, l]$ to $\mathbb{R}^n$ (i.e., continuous curves in $\mathbb{R}^n$, given on $[0, l]$).

Theorem 14.35 *As above, let $F(t, x)$ be a lower semi-continuous set-valued map $F : \mathbb{R} \times \mathbb{R}^n \multimap \mathbb{R}^n$ with closed values and $A(t, x) : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a field of single-valued linear operators jointly continuous in the parameters $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$. Let also (14.33) be fulfilled. Then for any $l > 0$, $x_0, v_0 \in \mathbb{R}^n$ inclusion (14.34) has a solution on $[0, l]$ with initial position x_0 and initial velocity v_0.*

Proof. In $C^0([0, l], \mathbb{R}^n)$ introduce the σ-algebra $\tilde{\mathcal{F}}$ generated by cylindrical sets. By $\tilde{\mathcal{P}}_t$ denote the σ-algebra generated by cylindrical sets over $[0, t] \subset [0, l]$.

Consider the set-valued mapping B sending $x(\cdot) \in C^0([0, l], \mathbb{R}^n)$ into $\mathcal{P}F(\cdot, x(\cdot))$. Since under condition (14.33) all selectors from $\mathcal{P}F(\cdot, x(\cdot))$ are integrable (see above), B takes values in the space $L^1(([0, l], \mathcal{B}, \lambda), \mathbb{R}^n)$. It is known (see, e.g., [155, Section 5.5]) that under the above-mentioned conditions $B : C^0([0, l], \mathbb{R}^n) \to L^1(([0, l], \mathcal{B}, \lambda), \mathbb{R}^n)$ is lower semicontinuous and for any $x(\cdot) \in C^0([0, l], \mathbb{R}^n)$ the set $\mathcal{P}F(\cdot, x(\cdot))$, i.e., the image $B(x(\cdot))$, is decomposable and closed. Thus, by Lemma 14.34, B has a continuous selector $b : C^0([0, l], \mathbb{R}^n) \to L^1(([0, l], \mathcal{B}, \lambda), \mathbb{R}^n)$.

For any $t \in [0, l]$ introduce the map $f_t : C^0([0, l], \mathbb{R}^n) \to C^0([0, l], \mathbb{R}^n)$ that sends a curve $x(\cdot) \in C^0([0, l], \mathbb{R}^n)$ into the curve

$$f_t(\tau, x(\cdot)) = \begin{cases} x(\tau) & \text{for } \tau \in [0, t] \\ x(t) & \text{for } \tau \in [t, l] \end{cases}.$$

Obviously the map f_t is continuous. Since $f_t(\tau, x(\cdot))$ belongs to $C^0([0, l], \mathbb{R}^n)$, the curve $b(f_t(\tau, x(\cdot))) \in L^1(([0, l], \mathcal{B}, \lambda), \mathbb{R}^n)$ is well-defined. By construction $b(f_t(\tau, x(\cdot))) \in F(\tau, x(\tau))$ for almost all $\tau \in [0, t]$ and this selector continuously depends on t in $L^1(([0, l], \mathcal{B}, \lambda), \mathbb{R}^n)$.

Consider the map $v : [0,l] \times C^0([0,l],\mathbb{R}^n) \to \mathbb{R}^n$ defined by the formula

$$v(t,x(\cdot)) = v_0 + \int_0^t b(f_t(\tau,x(\cdot)))\mathrm{d}\tau. \tag{14.35}$$

By construction this map is jointly continuous in $t \in [0,l]$ and $x(\cdot) \in C^0([0,l],\mathbb{R}^n)$. In addition it is clear that if $x_1(\cdot)$ and $x_2(\cdot)$ coincide on $[0,t]$ then $v(t,x_1(\cdot)) = v(t,x_2(\cdot))$. This means that $v(t,x(\cdot))$ is measurable with respect to $\tilde{\mathcal{P}}_t$ (see, e.g., [83]).

Taking into account (14.33) one can easily derive the inequality

$$\begin{aligned}\|v(t,x(\cdot))\| &= \left\|\int_0^t b(f_t(\tau,x(\cdot)))\mathrm{d}\tau\right\| \le \int_0^t \|b(f_t(\tau,x(\cdot)))\|\mathrm{d}\tau \\ &\le \int_0^t \|F(\tau,x(\tau))\|\mathrm{d}\tau \le \Theta\int_0^t (1+\|x(\tau)\|)\mathrm{d}\tau \\ &\le \Theta\int_0^t (1+\|x(\cdot)\|_{C^0})\mathrm{d}s \le l\Theta(1+\|x(\cdot)\|_{C^0}),\end{aligned}$$

where $\|\cdot\|_{C^0}$ is the norm in $C^0([0,l],\mathbb{R}^n)$.

Define $A(t,x(\cdot))$ by $A(t,x(\cdot)) = A(t,x(t))$. Notice that $A(t,x(\cdot))$ is measurable with respect to $\tilde{\mathcal{P}}_t$ and that from (14.33) it follows that $\|A(t,x(\cdot))\| \le \Theta(1+\|x(\cdot)\|_{C^0})$. So, both $v(t,x(\cdot))$ and $A(t,x(\cdot))$ satisfy the Itô condition in the form

$$\|v(t,x(\cdot)\| + \|A(t,x(\cdot)\| \le \bar{\Theta}(1+\|x(\cdot)\|_{C^0})$$

with $\bar{\Theta} = \max(\Theta, l\Theta)$.

Now the pair $v(t,x(\cdot))$ and $A(t,x(\cdot))$ satisfies all the conditions of [83, Theorem III.2.4], hence, the stochastic differential equation

$$x(t) = x_0 + \int_0^t v(s,x(\cdot))\mathrm{d}s + \int_0^t A(s,x(\cdot))\mathrm{d}w(s) \tag{14.36}$$

has a weak solution on $[0,l]$. This means that there exist a probabilistic measure μ on $(C^0([0,l],\mathbb{R}^n),\ \mathcal{F})$ and a Wiener process in $\mathbb{R}^n$, given on $(C^0([0,l],\mathbb{R}^n),\mathcal{F},\mu)$ and adapted to $\mathcal{P}_t$, such that the coordinate process $x(t)$ on $(C^0([0,l],\mathbb{R}^n),\mathcal{F},\mu)$ and $w(t)$ satisfy (14.36). This together with (14.35) completes the proof of the Theorem. □

To investigate this problem on a manifold we use the constructions of Section 7.7.3. The assumptions here are more restrictive than in the case of Euclidean space.

Let M be a stochastically complete Riemannian manifold on which a vector force field $\bar{\alpha}(t,m)$ independent of velocities is given. Thus the Newton law of the mechanical system takes the form

$$\frac{\mathrm{D}}{\mathrm{d}t}\dot{m}(t) = \bar{\alpha}(t,m(t)).$$

We suppose that the random perturbation of velocity takes the form $A(m)\dot{w}(t)$ where $A(m) : \mathbb{R}^k \to T_m M$ is a smooth field of linear operators. We suppose in addition that $A(m)A^*(m) = I$ where I is the unit operator in $T_m M$. This assumption can be interpreted as the fact that the Riemannian metric on M is determined by the diffusion coefficient generated by $A(m)$. In particular it means that we can apply the machinery of the equations with unit diffusion coefficient from Section 7.7.3.

The equation of motion for the system with random perturbation of velocity is given here in terms of mean derivatives on manifolds as introduced by Definition 9.1 and formula (9.18), and by the covariant mean derivative introduced by (9.15) in terms of stochastic parallel translation (see Sections 9.3 and 9.4). This equation takes the form

$$\begin{cases} \mathrm{D}\xi(t) = v(t, \xi(t)) \\ \mathrm{D}_2\xi(t) = I \\ \mathbf{D}v(t, \xi(t)) = \bar{\alpha}(t, \xi(t)). \end{cases} \tag{14.37}$$

Theorem 14.36 *Let the force field $\bar{\alpha}(t, m)$ be jointly continuous in t, m and be uniformly bounded, i.e., $\|\bar{\alpha}(t, m)\| < K$ for all $m \in M$ and $t \in [0, l] \subset \mathbb{R}$ for some $K > 0$. Then for every pair $m_0 \in M$, $v_0 \in T_{m_0}M$ there exists a solution of (14.37) with initial conditions $\xi(0) = m_0$, $v(0) = v_0$ that is well-defined on the entire interval $t \in [0, l]$.*

Proof. The idea of the proof is analogous to that for Langevin equations. We reduce (14.37) to the equation of velocity hodograph type in a single linear space. Then we show that the latter has a weak solution and that its Itô development satisfies (14.37). The difference is that here we use the velocity hodograph equation in terms of stochastic parallel translation (unlike the case of the Langevin equation where the ordinary parallel translation was applied) and so we have to appeal to the constructions of Section 7.7.3. In fact the proof uses the same argument as that of Theorem 7.95 and we refer the reader to the latter for a detailed explanation.

Consider the space $\tilde{\Omega} = C^0([0, l], T_{m_0}M)$ with the σ-algebra $\tilde{\mathcal{F}}$ generated by cylinder sets and a Wiener measure ν on $\tilde{\mathcal{F}}$. On the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \nu)$ the coordinate process $\tilde{w}(t, x(\cdot)) = x(t)$ is a Wiener process adapted to the family of σ-subalgebras $\mathcal{P}_t$ that for each t is generated by cylinder sets with bases on $[0, t]$ and is completed by all sets with ν-measure zero.

Since M is stochastically complete, the Itô development $R_I\tilde{w}(t)$ is well-defined on $[0, l]$ for ν-a.s. all curves in $\tilde{\Omega}$ and the parallel translation along ν-almost all sample paths of $R_I\tilde{w}(t)$ is also well-defined (see Section 7.6). Thus we can apply the operator Γ of parallel translation along $R_I\tilde{w}(\cdot)$ from each $R_I\tilde{w}(t)$ to $R_I\tilde{w}(0) = m_0$.

Introduce the process $\beta(t, x(\cdot)) = \int_0^t \Gamma\bar{\alpha}(s, R_I x(s))ds$ in $T_{m_0}M$. From the properties of parallel translation and of R_I it follows that $\beta(t)$ is uniformly bounded by the constant lK and that it is non-anticipative with respect to $\mathcal{P}_t$. In addition the density

$$\rho\left(x(\cdot)\right)=\exp\left(\int_0^l \langle \beta\left(t,x(\cdot)\right),\mathrm{d}\tilde{w}(t)\rangle-\frac{1}{2}\int_0^l \beta\left(t,x(\cdot)\right)^2\,\mathrm{d}t\right) \tag{14.38}$$

satisfies the equality

$$\int_{\tilde{\Omega}} \rho\,\mathrm{d}\nu = 1, \tag{14.39}$$

and so the measure μ on $(\tilde{\Omega},\tilde{\mathcal{F}})$ defined by the relation $\mathrm{d}\mu=\rho\mathrm{d}\nu$ is a probability measure.

Then the coordinate process $\bar{v}(t)$ on the probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\mu)$ satisfies the equation

$$\bar{v}(t)=v_0+\int_0^t \beta(s,\bar{v}(s))\mathrm{d}s+w(s) \tag{14.40}$$

where $w(t)$ is a Wiener process in $T_{m_0}M$, given on $(\tilde{\Omega},\tilde{\mathcal{F}},\mu)$, that is non-anticipative with respect to $\mathcal{P}_t$. The Itô development $\xi(t)=R_I\bar{v}(t)$ has the same sample paths as $R_I\tilde{w}(t)$ and so it is well-defined on the entire interval $[0,l]$.

Introduce the vector field $v(t,m)$ as the regression

$$v(t,x)=E(\Gamma_{0,t}\beta(t)\mid R_I\bar{v}(t)=x)$$

where $\Gamma_{t,0}$ is the operator of parallel translation along $R_I\bar{v}(\cdot)$ from $R_I\bar{v}(0)=m_0$ to $R_I\bar{v}(t)$. Taking into account the properties of the Itô development and of parallel translation as well as the construction of the covariant mean derivatives (9.15) and (9.18), one can easily show that $\xi(t)$ and $v(t,m)$ satisfy (14.37). □

Chapter 15
The Newton-Nelson Equation

The Newton-Nelson equation is a version of Newton's law formulated in terms of mixed symmetric second order mean derivatives. It describes the motion of a quantum particle in the framework of stochastic mechanics.

Nelson's stochastic mechanics is a subject based on the ideas of classical physics but giving the same predictions as quantum mechanics. Stochastic mechanics can be considered as a third method of quantization differing from the well-known Hamiltonian and Lagrangian (path integrals) methods. Each method has its own domain of applicability, and these domains have a large intersection (where the results are equivalent), but none of them includes any other completely.

The history of stochastic mechanics is presented, for example, in [45, 136, 187, 188, 190]. Apparently, the idea was first suggested by I. Fenyes [73], but it only became widely known, and obtained a natural form, following the appearance of Nelson's independently developed work [187, 188]. At present, among other things, a description of spinning particles, relativistic particles, the uncertainty principle and some parts of quantum field theory have been given in the language of stochastic mechanics (see, e.g., [45, 55, 136, 141, 142, 237, 238]). Nelson's book [190] surveys the main developments of the theory up to 1985.

In the version of stochastic mechanics pioneered by Nelson, the trajectory of a particle was assumed to be a Markov diffusion process. It is important to point out that in the works of V.P. Dmitriev and of H. Grabert, P. Hanggi and P. Talkner (see, e.g., [52, 137] and the references therein) physical reasoning is used to conclude that such a trajectory must be a non-Markov process. In addition, in his later work, E. Nelson discovered that in the framework of his "Markov" approach some higher momenta of two independent particles could be dependent.

We develop another approach to stochastic mechanics where the equation of motion is that of Nelson (the Newton-Nelson equation) but the trajectory is allowed to be an Itô diffusion type process that may not be a Markov process. Nevertheless we show that this version is related to quantum mechanics in an

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_15,

analogous way to Nelson's original version. Moreover, following the ideology of equations with mean derivatives of Section 8.1 (see Definition 8.21) we suggest an equation of Itô type for finding solutions of the Newton-Nelson equation such that for two independent particles it splits into two independent equations. Another (very simple) modification is that we find the diffusion term of the process from a separate equation with quadratic mean derivative and do not postulate its value *a priori*.

The intersection of the domains of applicability of stochastic mechanics and ordinary (Hamiltonian) quantum mechanics is as follows: for the case of forces for which the Schrödinger equation is well-defined (i.e., for potential forces, certain forces with friction and the magnetic field), there exists a natural interrelation between the processes of stochastic mechanics and wave functions, the solutions of the corresponding Schrödinger equations (see, e.g., [187, 188, 233]). Below in Section 15.1.1 we describe this interrelation for the case of potential forces. Note that in the paper [38] (see also [27, 192, 240]) this interrelation and the methods of partial differential equations were applied to obtain the existence of trajectories for non-relativistic stochastic mechanical systems with potential forces in $\mathbb{R}^n$ where the potential belongs to the so-called Rellich class. It should be pointed out that the trajectories obtained in [27, 38, 73, 240] are Markov diffusion processes, as was postulated in Nelson's pioneering works [187, 188].

15.1 Stochastic Mechanics in $\mathbb{R}^n$

This section consists of two subsections. In Subsection 15.1.1 we present the basic ideas of stochastic mechanics and, in particular, we describe the Newton-Nelson equation and its relation to the Schrödinger equation.

In Subsection 15.1.2, following [102, 105, 128, 129], we prove the existence of trajectories for non-relativistic stochastic mechanical systems under various initial conditions and with forces in $\mathbb{R}^n$ which are a sum of a force independent of velocities and a force linearly dependent on velocities. We do not assume the forces to be potential or gyroscopic; i.e., we include cases where other methods of quantization are not applicable (e.g., where the Schrödinger equation does not exist).

We apply the methods developed in Chapter 8.

15.1.1 Principal ideas of Nelson's stochastic mechanics

Let $\xi(t)$ be a stochastic process and assume that the mean derivatives (see Section 8.1) exist for $\xi(t)$.

Definition 15.1. The vector $\frac{1}{2}(DD_* + D_*D)\xi(t)$ is called the *acceleration* of the stochastic process $\xi(t)$.

Of course there exists a Borel vector field on $\mathbb{R}^n$ (the regression) such that the above acceleration is the composition of that vector field and $\xi(t)$.

Direct calculations show that

$$\frac{1}{2}(DD_* + D_*D)\xi(t) = (D_SD_S - D_AD_A)\xi(t) = D_Sv^\xi(t) - D_Au^\xi(t) \quad (15.1)$$

(see Definitions 8.15 and 8.16).

In this Section we shall mainly deal with a process of type (8.15), i.e., of the form $\xi(t) = \xi_0 + \int_0^t a(s)\mathrm{d}s + \sigma w(t)$. For such a process, from formulae (8.24) and (8.25) it follows that

$$D_Sv^\xi(t) = \frac{\partial}{\partial t}v^\xi(t) + (v^\xi(t)\cdot\nabla)v^\xi(t) \quad (15.2)$$

and

$$D_Au^\xi(t) = (u^\xi(t)\cdot\nabla)u^\xi(t) + \frac{\sigma^2}{2}\nabla^2u^\xi(t). \quad (15.3)$$

(Note that the right hand side of (15.2) has the form of an ordinary derivative of a non-autonomous vector field in the direction of itself in $\mathbb{R}^n$.) Thus, from (15.1)–(15.3) we obtain for a $\xi(t)$ of type (8.15) that

$$\begin{aligned}&\frac{1}{2}(DD_* + D_*D)\xi(t) \quad (15.4)\\ &= \left(\frac{\partial v^\xi(t)}{\partial t} + (v^\xi(t)\cdot\nabla)v^\xi(t)\right) - \left(\frac{\sigma^2}{2}\nabla^2u^\xi(t) + (u^\xi(t)\cdot\nabla)u^\xi(t)\right).\end{aligned}$$

Consider a Newtonian mechanical system in $\mathbb{R}^n$ with the vector force field $\bar{\alpha}(t,x,X)$. Here Newton's law (11.2) takes the following trivial form:

$$\ddot{x}(t) = \frac{1}{\mathrm{m}}\bar{\alpha}(t,x(t),\dot{x}(t)), \quad (15.5)$$

where m is the mass of a particle, $\ddot{x}(t)$ is the acceleration vector and $\dot{x}(t)$ is the velocity vector of the curve $x(t)$.

In stochastic mechanics the trajectory of a particle is assumed to be a stochastic process, not a deterministic curve. More precisely, we have the following:

Definition 15.2. A diffusion type process $\xi(t)$ in $\mathbb{R}^n$ is called a *stochastic-mechanical trajectory* of a particle with mass $\mathrm{m} > 0$ under the action of the force field $\bar{\alpha}(t,x,X)$ if it satisfies the system

$$\begin{cases}\frac{1}{2}(DD_* + D_*D)\xi(t) = \frac{1}{\mathrm{m}}\bar{\alpha}(t,\xi(t),v^\xi(t))\\ D_2\xi(t) = \frac{\hbar}{\mathrm{m}}I\end{cases} \quad (15.6)$$

where $\hbar = \frac{h}{2\pi}$, h is Plank's constant and I is the unit $n \times n$ matrix. In this case we also say that the stochastic-mechanical system with force $\bar{\alpha}(t, x, X)$ is given. The equality (15.6) is called the *Newton-Nelson equation*.

From the second equation of (15.6) it follows that $\xi(t)$ is a process of type (8.15) with σ such that $\frac{\sigma^2}{2} = \frac{\hbar}{2\mathrm{m}}$.

If $\sigma = 0$, the process $\xi(t)$ from Definition 15.2 turns into a deterministic curve and (15.6) becomes the ordinary Newton law (15.5). Without loss of generality we assume $\mathrm{m} = 1$.

Remark 15.3. Equality (15.6) for the Euclidean space $\mathbb{R}^n$ was first obtained by Nelson in [187]. It was also shown there that among all possible definitions of the acceleration of a stochastic process which are symmetric in time (i.e., well-defined physically) and coincide with the ordinary definition for smooth trajectories, only Definition 15.1 gives the correct result for some particular examples in quantum mechanics. Later equation (15.6) (for potential forces and in the form where the right-hand side is transformed according to (15.4)) was derived from some variational principles (see [190]).

Note that in stochastic mechanics one deals with the "quantization" of Newton's second law, while in the ordinary quantization procedures some other equations of motion (Euler-Lagrange or Hamilton) are involved.

The correspondence between stochastic mechanics and ordinary (Hamiltonian) quantum mechanics was established for potential forces (see, e.g., [187, 188, 190]) and for certain forces with friction [233] where both the Schrödinger equation and Newton-Nelson equation (15.6) are well-defined. We illustrate this correspondence with an example of potential forces. The arguments here are close to those in [187], but since we apply Lemmas 8.17 and 8.18 instead of classical results for Markov diffusions, we show that the correspondence is also valid under the assumption that the trajectories are Itô processes of diffusion type.

Let the force field $\bar{\alpha}$ of the mechanical system be a potential, i.e., it does not depend on velocity and $\bar{\alpha} = -\mathrm{grad}\, V$, where V is the potential energy. Let $\xi(t)$ be a trajectory of the stochastic-mechanical system as in Definition 15.2 with this force. Recall that for the osmotic velocity $u^\xi(t) = u^\xi(t, \xi(t))$ the vector field $u^\xi(t, x)$ is always described in the form $u^\xi = \sigma^2 \mathrm{grad}\, R$, where $R = \frac{1}{2} \log \rho^\xi(t, x)$ (see (8.18)). Let us suppose that for the current velocity $v^\xi(t) = v^\xi(t, \xi(t))$ the vector field $v^\xi(t, x)$ is also a gradient $v^\xi = \sigma^2 \mathrm{grad}\, S$ for some real function $S(t, x)$. Note that $S(t, x)$ is defined to within the functions depending only on t, i.e., whose gradient is zero. Consider the complex-valued function Ψ on M of the form $\Psi(t, x) = \exp(R + \mathrm{i}S)$.

Theorem 15.4 *Ψ satisfies the Schrödinger equation*

$$\frac{\partial \Psi}{\partial t} = \mathrm{i}\frac{\sigma^2}{2}\nabla^2 \Psi - \mathrm{i}\frac{1}{\hbar} V \Psi. \tag{15.7}$$

Proof. From Lemmas 8.17 and 8.18 it follows that $\frac{\partial u^\xi}{\partial t} = -\frac{\sigma^2}{2}\mathrm{grad}\,\mathrm{div}\, v^\xi - \mathrm{grad}\,(v^\xi \dot{u}^\xi)$. From (15.4) and (15.6) for $\bar{\alpha} = -\mathrm{grad}\, V$ it follows that $\frac{\partial v^\xi}{\partial t} = -\mathrm{grad}\, V - (v^\xi \cdot \nabla)v^\xi + (u^\xi \cdot \nabla)u^\xi + \frac{\sigma^2}{2}\nabla^2 u^\xi$. Then straightforward calculations show that

$$\frac{\partial \Psi}{\partial t} = \left(\frac{\partial R}{\partial t} + \mathrm{i}\frac{\partial S}{\partial t}\right)\Psi. \tag{15.8}$$

In order to find $(\frac{\partial R}{\partial t} + \mathrm{i}\frac{\partial S}{\partial t})$ we note that $\mathrm{grad}\,(\frac{\partial R}{\partial t} + \mathrm{i}\frac{\partial S}{\partial t}) = \frac{1}{\sigma^2}(\frac{\partial u^\xi}{\partial t} + \mathrm{i}\frac{\partial v^\xi}{\partial t})$ and so, by the above expressions for $\frac{\partial u^\xi}{\partial t}$ and $\frac{\partial v^\xi}{\partial t}$, since u^ξ and v^ξ are gradients, we obtain the formula

$$\begin{aligned}
&\mathrm{grad}\left(\frac{\partial R}{\partial t} + \mathrm{i}\frac{\partial S}{\partial t}\right)\\
&= -\mathrm{grad}\,\mathrm{div}\, v^\xi - \frac{1}{\sigma^2}\mathrm{grad}\,(v^\xi \cdot u^\xi) - \frac{\mathrm{i}}{\sigma^2}\mathrm{grad}\, V\\
&\quad -\frac{\mathrm{i}}{\sigma^2}(v^\xi \cdot \nabla)v^\xi + \frac{\mathrm{i}}{\sigma^2}(u^\xi \cdot \nabla)u^\xi + \frac{\mathrm{i}}{2}\nabla^2 u^\xi\\
&= \mathrm{grad}\left(-\frac{1}{2}\mathrm{div}\, v^\xi - \frac{1}{\sigma^2}(v^\xi \cdot u^\xi) - \frac{\mathrm{i}}{\sigma^2}V - \frac{\mathrm{i}}{2\sigma^2}(v^\xi)^2 + \frac{\mathrm{i}}{2\sigma^2}(u^\xi)^2 + \frac{\mathrm{i}}{2}\mathrm{div}\ u^\xi\right)\\
&= \mathrm{grad}\left(\frac{\mathrm{i}}{2}(\mathrm{div}\, u^\xi + \mathrm{i}\,\mathrm{div}\, v^\xi) + \frac{\mathrm{i}}{2\sigma^2}(u^\xi + \mathrm{i}v^\xi)^2 - \frac{\mathrm{i}}{\sigma^2}V\right).
\end{aligned}$$

Hence,

$$\left(\frac{\partial R}{\partial t} + \mathrm{i}\frac{\partial S}{\partial t}\right) = -\frac{\mathrm{i}}{2}(\mathrm{div}\ u^\xi + \mathrm{i}\,\mathrm{div}\, v^\xi) + \frac{1}{2\sigma^2}(u^\xi + \mathrm{i}\, v^\xi)^2 - \frac{\mathrm{i}}{\sigma^2}V + \mathrm{i}\, f(t) \tag{15.9}$$

where $f(t)$ is a real function, depending only on t, i.e., $\mathrm{grad}\, f(t) = 0$. On the other hand,

$$\nabla^2\Psi = \Psi\left(\frac{1}{\sigma^2}(u^2 + \mathrm{i}v^2)^2 + \frac{1}{\sigma^4}(\mathrm{div}\, u^\xi + \mathrm{i}\,\mathrm{div}\, v^\xi)\right). \tag{15.10}$$

Comparing (15.8), (15.9) and (15.10), we see that the following equality holds:

$$\frac{\partial \Psi}{\partial t} = \mathrm{i}\,\frac{\sigma^2}{2}\nabla^2\Psi - \mathrm{i}\frac{1}{\hbar}V\Psi + \mathrm{i}\, f(t)\Psi. \tag{15.11}$$

Adding to S an appropriate constant depending on t, one can obtain the equality $f = 0$. So (15.11) turns into the Schrödinger equation (15.7) with potential V. □

So, $\Psi(t, x)$ is a Schrödinger wave function corresponding to the above mechanical system. Conversely, let Ψ satisfy the Schrödinger equation (15.7) with potential V. Consider the functions $R = \mathrm{Re}\log\Psi$ and $S = \mathrm{Im}\log\Psi$ and the vector fields $u(t, x) = \sigma^2\mathrm{grad}\, R$, $v(t, x) = \sigma^2\mathrm{grad}\, S$ and $a(t, x) = v(t, x) + u(t, x)$.

Theorem 15.5 *Solutions $\xi(t)$ of the Itô equation*

$$d\xi(t) = a(t,\xi(t))dt + \sigma dw(t), \tag{15.12}$$

where $a(t,x)$ and σ are as introduced above, satisfy equation (15.6) with $\bar{\alpha} = -\operatorname{grad} V$.

We leave the proof to the reader as a simple exercise. Note that for a solution $\xi(t)$ of (15.12) we have $v^{\xi}(t,x) = v(t,x)$ and $u^{\xi}(t,x) = u(t,x)$, and it should be emphasized that in this case the assumption that $v^{\xi}(t,x)$ is a gradient is fulfilled automatically. Note also that strong solutions of (15.12) are Markov diffusion processes (see Definition 6.17). Thus we can formulate:

Proposition 15.6 *If a stochastic-mechanical trajectory $\zeta(t)$ with the force $\bar{\alpha} = -\operatorname{grad} V$ is an unique strong solution of (15.12) corresponding to a Schrödinger wave-function Ψ as described above, then $\zeta(t)$ is a Markov diffusion process.*

Indeed, $\rho^{\zeta}(t,x) = |\Psi(t,x)|^2 = \rho^{\xi}(t,x)$, where $\rho^{\xi}(t,x)$ is the probability density of the diffusion process $\xi(t)$, the solution of (15.12) constructed from $\Psi(t,x)$, which corresponds to $\zeta(t)$. The above diffusion process (trajectory) can be shown to exist when V belongs to the very broad so-called Rellich class (see [38]).

15.1.2 Existence theorems

In this Section we prove the existence of the trajectory assuming neither the Schrödinger equation to be well-defined nor its solution to exist. In this case the trajectory is not a diffusion process but an Itô process of type (8.15). First we consider deterministic initial data for the solution which leads to a singularity at $t = 0$ resembling the Big Bang (see Remark 15.10 below). Then we obtain another version of the construction that yields the existence of a solution, for random initial data, such that its distribution is nowhere zero. In the latter case no singularity at time zero arises.

Everywhere below in this Section we consider the vector force field $\bar{\alpha}$ of the form $\bar{\alpha}(t,x,X) = \bar{\alpha}_0(t,x) + \bar{\alpha}_1(t,x)X$ where $\bar{\alpha}_0(t,x)$ is a vector field on $\mathbb{R}^n$ depending on $t \in [0,l]$ and $\bar{\alpha}_1(t,x)$ is a linear operator in $\mathbb{R}^n$ depending on the parameters $t \in [0,l]$ and $x \in \mathbb{R}^n$, i.e., $\bar{\alpha}_1(t,x)$ is a (1.1)-tensor field on $\mathbb{R}^n$. Consider the derivative of $\bar{\alpha}_1(t,x)$, i.e., the field of bilinear operators $\bar{\alpha}_1'(t,x))(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, and the vector field $\mathrm{tr}\bar{\alpha}_1'(t,x)(\bar{\alpha}_1)ds = \sum_{i=1}^n \bar{\alpha}_1'(t,x)(\bar{\alpha}_1(t,x)e_i, e_i)$, where $e_1,\ldots,e_n$ is an arbitrary orthonormal frame in $\mathbb{R}^n$. We assume the following condition to be fulfilled:

Condition 15.7 *The vector field $\bar{\alpha}_0(t,x)$ and the tensor field $\bar{\alpha}_1(t,x)$ are Borel measurable jointly in t and x, $\mathrm{tr}\bar{\alpha}_1'(s,W(s))(\bar{\alpha}_1)$ exists and is also Borel measurable jointly in t, x and there exists a constant $C>0$ such that*

$$\int_0^l (\|\bar{\alpha}_0(t,x(t))\|^2 + \|\bar{\alpha}_1(t,x(t))\|^2 + \|\mathrm{tr}\bar{\alpha}_1'(t,x(t))(\bar{\alpha}_1)\|^2)\mathrm{d}t < C$$

for any continuous curve $x(t)$ in $\mathbb{R}^n$, $t\in[0,l]$, where $\|\bar{\alpha}_1(t,x)\|$ is the operator norm.

Condition 15.7 is fulfilled, for example, if $\bar{\alpha}_0(t,x)$ is continuous and uniformly bounded, $\bar{\alpha}_1(t,x)$ belongs to the functional space $C^1(\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n,\mathbb{R}^n)$ and has a bounded norm in that space.

For the sake of simplicity, without loss of generality we may assume $\sigma=1$. We shall deal with stochastic processes starting at a non-random point in $\mathbb{R}^n$, and for the sake of simplicity we assume this point to be equal to the origin.

Consider the probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\nu)$ where $\tilde{\Omega}=C^0([0,l],\mathbb{R}^n)$, $\tilde{\mathcal{F}}$ is the σ-algebra generated by the cylindrical sets, and ν is a Wiener measure (see Section 6.2.1). Denote by $\mathcal{B}_t$ the σ-algebra generated by the cylindrical sets with the bases over $[0,t]$; all the $\mathcal{B}_t$ are completed by the sets of ν-measure zero.

Recall that the coordinate process $W(t,x(\cdot))=x(t)$, $x(\cdot)\in\tilde{\Omega}$, on $(\tilde{\Omega},\tilde{\mathcal{F}},\nu)$ is a Wiener process in $\mathbb{R}^n$.

Let $t_0\in(0,l)$ and for $t\in[0,l]$ denote by $t_0(t)$ the function

$$t_0(t)=\begin{cases}\frac{1}{t_0}\ \text{if}\ t<t_0,\\ \\ \frac{1}{t}\ \text{if}\ t\geq t_0.\end{cases}\tag{15.13}$$

Recall that we are looking for a solution as a process of type (8.15) that under the above assumptions takes the form $\xi(t)=\int_0^t a(s)\mathrm{d}s+w(t)$ where $a(t)$ is to be found. Pick a deterministic initial condition $a_0\in\mathbb{R}^n$ for $a(t)$ and consider in $\mathbb{R}^n$ the equation

$$\begin{aligned}a(t) = a_0 &+ \int_0^t \bar{\alpha}_0(s,W(\cdot))\mathrm{d}s + \int_0^t \bar{\alpha}_1(s,W(s))\mathrm{d}W(s) + \frac{1}{2}\int_0^t \mathrm{tr}\bar{\alpha}_1'(\bar{\alpha}_1)\mathrm{d}s\\ &-\frac{1}{2}\int_0^t t_0(s)a(s)\mathrm{d}s + \frac{1}{2}\int_0^t t_0(s)\mathrm{d}W(s) - \frac{1}{2}t_0(t)W(t)\end{aligned}\tag{15.14}$$

Equation (15.14) has a unique strong solution for $t\in[0,l]$. Indeed, the coefficients of (15.14) are either Lipschitz continuous and have a linear growth with respect to a or do not depend on a. Since the solution is strong, it exists for any Wiener process and is non-anticipative with respect to $\mathcal{P}_t^W=\mathcal{B}_t$. In what follows we will consider $a(t)$ for the realization of $W(t)$ as the coordinate process on $(\tilde{\Omega},\tilde{\mathcal{F}},\nu)$.

From Condition 15.7 it follows that $a(t)$ satisfies (8.32).

Below in this Section we will use the notation introduced in Section 8.3. Consider the function $\theta(l)$ on $\tilde{\Omega}$ determined for the process $a(t)$ by formula (8.33). From the above condition it easily follows that $\theta(l)$ is a probability density on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \nu)$. Denote by μ^ξ the corresponding measure on $(\tilde{\Omega}, \tilde{\mathcal{F}})$, i.e., $\mathrm{d}\mu^\xi = \theta\, \mathrm{d}\nu$ (note that μ^ξ and ν are equivalent), and by $\xi(t)$ the coordinate process on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu^\xi)$. Then (see Section 8.3) the process $\xi(t)$ is expressed in the required form

$$\xi(t) = \int_0^t a(s)\mathrm{d}s + w(t) \tag{15.15}$$

where $w(t)$ is a Wiener process on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu^\xi)$ adapted to $\mathcal{P}_t^\xi = \mathcal{P}_t^W = \mathcal{B}_t$.

Since $\xi(t)$ and $W(t)$ coincide as coordinate processes on $(\tilde{\Omega}, \tilde{\mathcal{F}})$, (15.14) clearly turns into

$$\begin{aligned} a(t) = a_0 &+ \int_0^t \bar{\alpha}_0(s, \xi(s))\mathrm{d}s + \int_0^t \bar{\alpha}_1(s, \xi(s))a(s)\mathrm{d}s + \int_0^t \bar{\alpha}_1(s, \xi(s))\mathrm{d}w(s) \\ &+ \frac{1}{2}\int_0^t \mathrm{tr}\bar{\alpha}_1'(s, \xi(s))(\bar{\alpha}_1)\mathrm{d}s + \frac{1}{2}\int_0^t t_0(s))\mathrm{d}w(s) - \frac{1}{2}t_0(t)\xi(t). \end{aligned} \tag{15.16}$$

Lemma 15.8

(i) $D_*^\xi \bar{\alpha}_1(s, \xi(s))\mathrm{d}w(s) = -\mathrm{tr}\bar{\alpha}_1'(t, \xi(t))(\bar{\alpha}_1) + \bar{\alpha}_1(t, \xi(t))\frac{\xi(t)}{t}$

$$-\bar{\alpha}_1(t, \xi(t))E_t^\xi(\kappa(t)).$$

(ii) $D_*^\xi \int_0^t t_0(s)\mathrm{d}w(s) = \frac{\xi(t)}{t^2} - E_t^\xi\left(\frac{\kappa(t)}{t}\right)$ *for* $t > t_0$,

where $\kappa(t)$ *is as in Lemma* 8.35.

In order to prove (i) one should replace the Itô integral by a backward Itô integral according to formula (6.7) and then apply formula (8.37) with the same arguments as in the proof of Lemma 8.26(ii). Assertion (ii) follows from formula (8.37), the definition of the Itô integral, and the definition of $t_0(t)$ given by formula (15.13).

Theorem 15.9 *For* $t \in (t_0, l)$ *the process* $\xi(t)$ *satisfies* (15.6), *i.e., it is a trajectory of the stochastic mechanical system with force* $\bar{\alpha}(t, m, X)$.

Proof. Recall that for $t \in (t_0, l)$ we have $t_0(t) = \frac{1}{t}$. From (15.15) and Theorem 8.7 it follows that $D\xi(t) = E_t^\xi(a(t))$. By (8.36) we obtain $D_*\xi(t) = E_t^\xi(a(t)) + \frac{\xi(t)}{t} - E_t^\xi(\kappa(t))$. In particular

$$v^\xi(t) = D_S\xi(t) = E_t^\xi(a(t)) + \frac{1}{2}\frac{\xi(t)}{t} - \frac{1}{2}E_t^\xi(\kappa(t)) \tag{15.17}$$

and

$$u^\xi(t) = D_A\xi(t) = -\frac{1}{2}\frac{\xi(t)}{t} + \frac{1}{2}E_t^\xi(\kappa(t)). \tag{15.18}$$

Then for $t > t_0$, using formula (8.38) and equation (15.16), as well as Lemmas 8.36, 8.37 and 8.38, one can obtain

$$DD_*\xi(t) = \bar{\alpha}_0(t,\xi(t)) + \bar{\alpha}_1(t,\xi(t))E_t^\xi(a(t)) + \frac{1}{2}\mathrm{tr}\bar{\alpha}_1'(s,\xi(s))(\bar{\alpha}_1)$$
$$+E_t^\xi\left(\frac{a(t)}{t}\right) - \frac{\xi(t)}{t^2} - \frac{1}{2}E_t^\xi\left(\frac{a(t)}{t}\right) + \frac{1}{2}\frac{a(t)}{t^2}.$$

Analogously, using in addition formula (8.39) and Lemma 15.8, one can calculate

$$D_*D\xi(t) = \bar{\alpha}_0(t,\xi(t)) + \bar{\alpha}_1(t,\xi(t))E_t^\xi(a(t)) + \bar{\alpha}_1(t,\xi(t))\left(\frac{x(t)}{t}\right)$$
$$-\bar{\alpha}_1(t,\xi(t)E_t^\xi(\kappa(t)) - \frac{1}{2}\mathrm{tr}\bar{\alpha}_1'(t,\xi(t))(\bar{\alpha}_1) + \frac{1}{2}\frac{\xi(t)}{t^2} - \frac{1}{2}E_t^\xi\left(\frac{\kappa(t)}{t}\right)$$
$$-\frac{1}{2}E_t^\xi\left(\frac{a(t)}{t}\right) + \frac{1}{2}E_t^\xi\left(\frac{\kappa(t)}{t}\right).$$

So

$$\frac{1}{2}(DD_* + D_*D)\xi(t) = \bar{\alpha}_0(t,\xi(t)) \tag{15.19}$$
$$+\bar{\alpha}_1(t,\xi(t))\left[E_t^\xi(a(t) + \frac{1}{2}\frac{\xi(t)}{t} - \frac{1}{2}\kappa(t))\right].$$

From formula (15.17) it follows that (15.19) coincides with the first equation of (15.6).

Since $\xi(t)$ is a process of the form (8.15) (see above), by Theorem 8.12 $D_2\xi = I$ (recall that we set $\sigma = 1$) and so system (15.6) is fulfilled. □

Remark 15.10. Generally speaking, $\xi(t)$ for $t \in (0,t_0)$ does not satisfy (15.6). Thus $\xi(t)$ can be interpreted only as a trajectory of the stochastic mechanical system beginning at the instant t_0 in the random configuration $\xi(t_0)$ with the initial forward derivative $E_t^\xi(a(t_0))$. It is clear that t_0 may be chosen arbitrarily close to zero, and so we can bring the initial values of the trajectory to the origin for the configuration and, again as close as we want, to a_0 for the forward derivative. But we cannot put $t_0 = 0$, since the integral $\int_0^t \frac{1}{s}\mathrm{d}w(s)$ does not exist ($\int_0^t \frac{1}{s^2}\mathrm{d}s$ diverges, see, e.g., [162] and Section 6.2.2), i.e., when $t_0 = 0$ equations (15.14) and (15.16) are ill-posed. This behavior of the solution is interpreted in Section 15.3 as a possible description of the Big Bang (see Remark 15.51).

Remark 15.11. If $\bar{\alpha} = -\mathrm{grad}\,V$ then there may exist a trajectory of the stochastic mechanical system obtained by Carlen [38] (see above). Suppose that both our and Carlen's trajectories are connected with the same solution Ψ of the corresponding Schrödinger equation. Then those trajectories determine different probability measures on the space of sample paths but the densities ρ on $\mathbb{R}\times\mathbb{R}^n$ coincide and are equal to $|\Psi|^2$.

Since the function $t_0(t)$ is piecewise smooth (see its definition by formula (15.13)) we may consider its derivative $t_0'(t)$ defined by the formula

$$t_0'(t)=\begin{cases}0 & \text{if } t<t_0,\\ -\frac{1}{t^2} & \text{if } t\geq t_0.\end{cases} \tag{15.20}$$

Theorem 15.12 *Equations* (15.14) *and* (15.16) *are equivalent to the following equations of Stratonovich type*

$$\begin{aligned}a(t)=a_0&-\frac{1}{2}\int_0^t t_0(s)a(s)\mathrm{d}s-\frac{1}{2}\int_0^t t_0'(s)W(s)\mathrm{d}s\\&+\int_0^t\bar{\alpha}_0(s,W(\cdot))\mathrm{d}s+\int_0^t\bar{\alpha}_1(s,W(s))\circ\mathrm{d}W(s)\end{aligned} \tag{15.21}$$

and

$$\begin{aligned}a(t)=a_0-\frac{1}{2}\int_0^t t_0(s)a(s)\mathrm{d}s-\frac{1}{2}\int_0^t t_0'(s)\xi(s)\mathrm{d}s+\int_0^t\bar{\alpha}_0(s,\xi(\cdot))\mathrm{d}s\\+\int_0^t\bar{\alpha}_1(s,\xi(s))a(s)\mathrm{d}s+\int_0^t\bar{\alpha}_1(s,\xi(s))\circ\mathrm{d}w(s),\end{aligned} \tag{15.22}$$

respectively.

Proof. Indeed, by the Itô formula (6.10) for $f(t,x)=t_0(t)x$ we have

$$\mathrm{d}(t_0(t)W(t))=t_0'(t)W(t)\mathrm{d}t+t_0(t)\mathrm{d}W(t)$$

and

$$\mathrm{d}(t_0(t)\xi(t))=(t_0(t)a(t)+t_0'(t)\xi(t))\mathrm{d}t+t_0(t)\mathrm{d}w(t).$$

An application of formula (6.23) (which follows from (6.6)) to $A=\bar{\alpha}_1$ completes the proof. □

Remark 15.13. Equations (15.21) and (15.22) show that $a(t)$ is a vector (belongs to the tangent bundle), a fact which plays a significant role in the transition to manifolds. Note also that if $\bar{\alpha}_1=0$, $a(t)$ becomes a process with a.s. smooth sample paths.

Now we are in position to prove the existence of a solution of the Newton-Nelson equation (15.6) with random initial data. Here we shall obtain a process that is a solution for all positive times, i.e., without singularity at $t=0$. However, the initial density must be nowhere equal to zero.

The main technical difference from the above case is that here we can calculate some derivatives “beforehand” and then use them to construct a solution of (15.6).

As above, consider equation (15.6) with $\bar{\alpha}(t,x,X) = \bar{\alpha}_0(t,x) + \bar{\alpha}_1(t,x)X$ and suppose that $\bar{\alpha}_0$ and $\bar{\alpha}_1$ satisfy Condition 15.7. Consider the initial data for (15.6) of the form

$$\xi(0) = x_0, \qquad D\xi(0) = a_0, \tag{15.23}$$

where x_0 and a_0 are random elements with values in $\mathbb{R}^n$, x_0 has a smooth distribution ρ_0 with respect to Lebesgue measure on $\mathbb{R}^n$ with $\rho(x) \neq 0$ for all x and a_0 is bounded.

All processes below in this section are considered on some finite time interval $[0,l]$.

Let $W_0(t)$ be the standard Wiener process on $\mathbb{R}^n$, i.e., the coordinate process on the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \nu)$ introduced at the beginning of this section. The filtration $\mathcal{P}_t^{W_0}$ is generated by cylinder sets with bases on $[0,t]$.

Consider the process $W(t) = x_0 + W_0(t)$. Its density, denoted by $\rho^W(t,x)$, satisfies the diffusion equation

$$\frac{\partial}{\partial t}\rho^W(t,x) = \frac{1}{2}\Delta\rho^W(t,x)$$

with initial condition $\rho^W(0,x) = \rho_0$. Thus the density $\rho^W(t,x)$ can be considered as given *a priori* and we use it in the further construction. By Lemma 8.17 the osmotic velocity $u^W(t,W(t))$ of $W(t)$ can be found by the formula

$$u^W(t,x) = \frac{1}{2}\operatorname{grad}\log\rho^W(t,x).$$

Hence $u^W(t,x)$ is uniquely constructed from $\rho^W(t,x)$ (and, like $\rho^W(t,x)$, will also be used later in the construction).

Consider the Itô equation

$$\begin{aligned} a(t) = a_0 &+ \int_0^t \bar{\alpha}_0(s,W(s))\,\mathrm{d}s \\ &+ \int_0^t \bar{\alpha}_1(s,W(s))\,\mathrm{d}W(s) + \frac{1}{2}\int_0^t \operatorname{tr}\nabla\bar{\alpha}_1(\bar{\alpha}_1(s,W(s)))\,\mathrm{d}s \\ &+ \int_0^t D^W u^W(s,W(s))\,\mathrm{d}s + \int_0^t (a(s)\cdot\nabla)u^W(s,W(s))\,\mathrm{d}s, \end{aligned} \tag{15.24}$$

where D^W is defined in (8.7) and may be represented as $D^W = \frac{\partial}{\partial t} + \frac{1}{2}\Delta$ (see (8.22) and (8.24)). Thus $D^W u^W(s,W(s))$ is known and we can prove the existence of a solution $a(t)$ of equation (15.24) by imitating the method used for equation (15.14). Since (15.24) is linear in a, under the above conditions this equation has a unique strong solution that we shall again denote by $a(t)$. Let $\theta(l)$ be defined by formula (8.33) with $a(t)$ and $W(t)$ as above. Clearly $\theta(t)$ is a martingale with expectation 1. Introduce a new probability measure

μ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ by the relation $\mathrm{d}\mu = \theta(l)\,\mathrm{d}\nu$. On the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$ the coordinate process $W(t)$ turns into $\xi(t) = x_0 + \int_0^t a(s)\mathrm{d}s + w(t)$ where $w(t)$ is a "new" Wiener process on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$, adapted to $\mathcal{P}_t^W = \mathcal{P}_t^\xi$ (Girsanov's theorem). Note that by construction $a(t)$ is not anticipative with respect to $\mathcal{P}_t^\xi$, i.e., $\xi(t)$ is a diffusion type process.

Theorem 15.14 *The process $\xi(t)$ satisfies equation* (15.6) *and the initial conditions* (15.23).

Lemma 15.15 *The random element x_0 has the same probability distribution on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \nu)$ and on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$.*

Proof. Let f be an arbitrary bounded Borel measurable function. As in Section 8.3 we denote by $E^0(\psi \mid \mathcal{B})$ the conditional expectation for ψ with respect to $\mathcal{B}$ on the probability space $(\Omega, \mathcal{F}, \nu)$ and by $E'(\psi \mid \mathcal{B})$ the same conditional expectation on $(\Omega, \mathcal{F}, \mu)$. Then

$$E'[f(x_0)] = E^0[f(x_0)\theta(l)] = E^0[f(x_0)E_0^0\theta(l)] = E^0[f(\xi_0)],$$

because $E_0^0\theta(l) = \theta(0) = 1$ (since $\theta(t)$ is a martingale). The assertion of the Lemma follows from the fact that f is arbitrary. □

Proof. [of Theorem 15.14] On the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu)$ equation (15.24) turns into

$$\begin{aligned} a(t) = a_0 &+ \int_0^t \bar{\alpha}_0(s, \xi(s))\,\mathrm{d}s + \int_0^t \bar{\alpha}_1(s, \xi(s))a(s)\,\mathrm{d}s \\ &+ \int_0^t \bar{\alpha}_1(s, \xi(s))\mathrm{d}w(s) + \frac{1}{2}\int_0^t \operatorname{tr} \nabla\bar{\alpha}_1(s, \xi(s))\,\mathrm{d}s \\ &+ \int_0^t D^W u^W(s, \xi(s))\,\mathrm{d}s + \int_0^t (a(s) \cdot \nabla)u^W(s, \xi(s))\,\mathrm{d}s. \end{aligned}$$

Now the proof is reduced to the verification that the required formulae are satisfied. This is done by application of formulae from Section 8.3 (for details, see [128] and [129]). □

Let a_1 be an arbitrary bounded random element with values in $\mathbb{R}^n$.

Corollary 15.16 *Let there exist a $\delta > 0$ such that for all $x \in \mathbb{R}^n$ the inequality $\rho_0(x) > \delta$ holds and let in addition* $\operatorname{grad}\log\rho_0$ *be uniformly bounded. Then for $a_0 = a_1 + u^W(0, x_0)$ the process $\xi(t)$ from Theorem* 15.14 *satisfies equation* (15.6) *with initial conditions $\xi(0) = x_0$ and $D_S\xi(0) = a_1$.*

15.2 The Geometric Form of Stochastic Mechanics

15.2.1 Some comments on stochastic mechanics on Riemannian manifolds

Stochastic-mechanical systems are influenced by the geometry of configuration space even more so than the systems of classical mechanics. As in the classical case the Riemannian metric on the configuration space determines the kinetic energy of the system, but it also defines the field of Wiener processes in terms of which the motion is described, so that the quadratic derivative of a trajectory yields the $(2,0)$ metric tensor. In addition, the curvature of the configuration space is involved in the stochastic version of Newton's law.

Let M be a n-dimensional Riemannian manifold. We study the exponential map, parallel translation and other geometric objects on M generated by the Levi-Civitá connection.

Let $\xi(t)$ be a stochastic process in M and assume that the mean derivatives in the sense of Definition 9.1 and formulae (9.14) and (9.15) exist for $\xi(t)$.

Definition 15.17. The vector $\frac{1}{2}(\mathbf{D}D_* + \mathbf{D}_*D)\xi(t)$ is called the *acceleration* of $\xi(t)$ (cf. Definition 15.1).

As in Section 15.1.1 there exists a Borel vector field on M (the regression) such that the acceleration is the composition of that field and $\xi(t)$.

On determining the covariant mean derivatives $\mathbf{D}_S = \frac{1}{2}(\mathbf{D} + \mathbf{D}_*)$ and $\mathbf{D}_A = \frac{1}{2}(\mathbf{D} - \mathbf{D}_*)$, we then obtain the following analog of formula (15.1)

$$\frac{1}{2}(\mathbf{D}D_* + \mathbf{D}_*D)\xi(t) = (\mathbf{D}_S D_S - \mathbf{D}_A D_A)\xi(t) = \mathbf{D}_S v^\xi(t) - \mathbf{D}_A u^\xi(t). \quad (15.25)$$

Now suppose that $D_2\xi(t) = \sigma^2\bar{g}$ where D_2 is the quadratic mean derivative, $\sigma > 0$ is a constant and $\bar{g}$ is the metric $(2,0)$-tensor. According to the material of Section 9.4 this means that we determine the forward and backward mean derivatives with respect to the Levi-Civitá connection of Riemannian metric given by $D_2\xi(t)$. In this case we derive from formulas (9.17) and (9.18) the following analogs of formulas (15.2) and (15.3):

$$\mathbf{D}_S v^\xi(t) = \frac{\partial}{\partial t}v^\xi(t) + \nabla_{v^\xi(t)}v^\xi(t), \quad (15.26)$$

$$\mathbf{D}_A u^\xi(t) = \nabla_{u^\xi(t)}u^\xi(t) + \frac{1}{2}\sigma^2\nabla^2 u^\xi(t), \quad (15.27)$$

where ∇ is the covariant derivative of the Levi-Civitá connection and ∇^2 is the Laplace-Beltrami operator (see Definition 2.58). Thus from formulae (15.25)–(15.27) it follows that for such $\xi(t)$ the following formula

$$\frac{1}{2}(\mathbf{D}D_* + \mathbf{D}_*D)\xi(t) = \left(\frac{\partial}{\partial t}v^{\xi}(t) + \nabla_{v^{\xi}(t)}v^{\xi}(t)\right) - \left(\frac{\sigma^2}{2}\nabla^2 u^{\xi}(t) + \nabla_{u^{\xi}(t)}u^{\xi}(t)\right) \tag{15.28}$$

holds (this is the analog of formula (15.4)).

Let M be the configuration space of a mechanical system as in Section 11.1 with a force field $\bar{\alpha}(t, m, X)$, i.e., the trajectory of the system is governed by equation (11.2). It seems to be natural to determine Newton's law for stochastic mechanics in this case by complete analogy with equation (15.6), i.e., by setting the acceleration equal to the force. But in doing this we will not obtain the correspondence with the solutions of the Schrödinger equations analogous to those described in Section 15.1.1. More precisely, the situation is as follows. One can easily make a minimal modification of the construction and prove an analog of Theorem 15.4 (we leave this to the reader as a simple exercise) but one obtains the Laplace-Beltrami operator ∇^2 in the analog of equation (15.7), while physicists would normally use the Laplace-de Rham operator $\mathit{\Delta} = \mathrm{d}\delta + \delta\mathrm{d}$ from Definition 1.73 (this difference seems to have been first highlighted in [45]).

So, in order to obtain the correspondence mentioned above we might replace the Laplace-Beltrami operator by the Laplace-de Rham operator in the right hand side of (15.28), i.e., change the above definition of acceleration (a variant of such a change is discussed below in Remark 15.19). Instead we take into account Weitzenbock's formula (2.37) and define a stochastic mechanical system on M as follows (cf. Definition 15.2):

Definition 15.18. A process $\xi(t)$ in M is called a *stochastic-mechanical trajectory* in M of a particle with mass m, under the action of the force field $\bar{\alpha}(t, m, X)$, if it satisfies the system

$$\begin{cases} \frac{1}{2}(\mathbf{D}D_* + \mathbf{D}_*D)\xi(t) = \frac{1}{\mathrm{m}}\bar{\alpha}(t, \xi(t), v^{\xi}(t)) + \frac{\hbar}{2\mathrm{m}}\widehat{\mathrm{Ric}}(\xi(t)) \circ u^{\xi}(t), \\ \qquad\qquad D_2\xi(t) = \frac{\hbar}{\mathrm{m}}\bar{g}, \end{cases} \tag{15.29}$$

where $\bar{g}$ is an autonomous positive definite symmetric $(2, 0)$-tensor and all mean derivatives and the Ricci tensor in the first equality of (15.29) are determined with respect to the Levi-Civitá connection of the Riemannian $(0, 2)$-metric tensor inverse to $\bar{g}$. In this case we say that a stochastic-mechanical system with force $\bar{\alpha}(t, m, X)$ is given on M. Relation (15.29) is called the *Newton-Nelson equation* on M.

Note that (15.29) determines the Riemannian metric which defines the kinetic energy for a classical mechanical system whose quantization is described by (15.29). In particular, the Newton law (11.2) for the latter system is given in terms of the covariant derivative of the Levi-Civitá connection of that metric.

In what follows we shall look for solutions of (15.29) in the class of Itô processes that are Itô developments of processes in tangent spaces of the form

$\zeta(t) = \int_0^t a(s)\mathrm{d}s + \sigma w(t)$ where $\sigma = \sqrt{\frac{\hbar}{m}}$ (see Section 9.3). In this case it is evident that if $\frac{\hbar}{\mathrm{m}} \approx 0$ (i.e., if the mass is very big in comparison with Plank's constant), similarly to the trajectories in $\mathbb{R}^n$, $\xi(t)$ turns into a deterministic curve and the first equality of (15.29) becomes the classical Newton law (11.2) if the metric determined from the second equality of (15.29) remains unchanged (recall that for a deterministic curve its quadratic derivative is zero).

We shall not deal with limits as $\frac{\hbar}{\mathrm{m}} \to 0$ and so without loss of generality in what follows we shall suppose the mass $\frac{\hbar}{m} = 1$.

Now the correspondence between stochastic mechanics and ordinary quantum mechanics on manifolds is analogous to that of Section 15.1.1 and is proved via the same argument (cf. [190]).

Remark 15.19. In order to make natural the definition of acceleration in the form $(\frac{\partial}{\partial t}v^\xi(t) + \nabla_{v^\xi(t)}v^\xi(t)) - (-\frac{1}{2}\Delta u + \nabla_{u^\xi(t)}u^\xi(t))$ with the Laplace-de Rham operator Δ (see above), in [53, 54] the construction of the parallel translation along stochastic processes was modified so that the parallel translation obtained takes into account the deviation of geodesics. Having defined $\bar{\mathbf{D}}$ and $\bar{\mathbf{D}}_*$ by formula (9.15), where the new parallel translation is involved, one obtains the Newton-Nelson equation in the form

$$\frac{1}{2}(\bar{\mathbf{D}}D_* + \bar{\mathbf{D}}_*D)\xi(t) = \bar{\alpha}(t, \xi(t), v^\xi(t)), \tag{15.30}$$

completely analogous to the usual form. Of course, (15.30) is equivalent to (15.29). Note that the Newton-Nelson equation in the form (15.30) is given in [53, 54, 190]. We do not use the form (15.30) since our constructions below are based on the usual parallel translation.

15.2.2 Existence theorems

The main purpose of this section is to generalize the existence theorem of Section 15.1.2 to a rather broad class of Riemannian manifolds, not necessarily Euclidean spaces (see [106, 107, 115]).

Our generalization follows the same scheme as the basic existence theorems of Section 15.1.2, the necessary modification is based on the methods developed in Chapters 8 and 9. Using parallel translation along stochastic processes we construct a special stochastic equation in the tangent space at the initial configuration of the motion (a certain stochastic version of the velocity hodograph equation (11.20), different from (14.10)) and prove the existence of its solutions. Then we show that the developments of the solutions satisfy the Newton-Nelson equation (at least after a certain non-zero instant, fixed in advance, if the initial configuration is deterministic). Note that the trajectory is an Itô process of diffusion type on a manifold, not necessarily a diffusion process.

We consider a vector force field $\bar{\alpha}$ of the same kind as in Section 15.1.2, namely of the form $\bar{\alpha}(t,m,X) = \bar{\alpha}_0(t,m) + \bar{\alpha}_1(t,m)X$ where $\bar{\alpha}_0(t,m)$ is a vector field on M depending on $t \in [0,l]$ and $\bar{\alpha}_1(t,m)$ is a linear operator in T_mM depending on the parameter $t \in [0,l]$, i.e., $\bar{\alpha}_1(t,m)$ is a $(1,1)$-tensor field on M. Examples of such $\bar{\alpha}$ include potential and gyroscopic force fields (e.g., a magnetic field or an electromagnetic field on a general relativistic space-time).

As in Section 15.1.2, for the sake of simplicity (and without loss of generality) we set $\frac{\hbar}{\mathrm{m}} = 1$ (i.e., we work in the system of units with this property).

Let us introduce some notation. For a $(1,1)$-tensor field $v(t,m)$ we will deal with its covariant derivative, the $(1,2)$-tensor field $\nabla v(t,m)(\cdot,\cdot) : T_mM \times T_mM \to T_mM$, as well as with the tensor $\nabla v(t,m)(v(t,m)\cdot,\cdot) : T_mM \times T_mM \to T_mM$ and its trace, the vector field

$$\mathrm{tr}\nabla v(t,m)(v) = \sum_{i=1}^{n} \nabla v(t,m)(v(t,m)e_i, e_i),$$

where $e_1,\dots,e_n$ is an orthonormal frame in T_mM (see formula 6.11).

Examples of the above $(1,1)$-tensor fields are $\widehat{\mathrm{Ric}}(m)$ and $\bar{\alpha}_1(t,m)$, for which we consider the vector fields (traces) $\mathrm{tr}\nabla\widehat{\mathrm{Ric}}(m)(\widehat{\mathrm{Ric}})$ and $\mathrm{tr}\nabla\bar{\alpha}_1(t,m)(\bar{\alpha}_1)$. Note that the field $\mathrm{tr}\nabla\widehat{\mathrm{Ric}}(m)(\widehat{\mathrm{Ric}})$ is C^∞-smooth since the tensor field $\widehat{\mathrm{Ric}}$ is also C^∞-smooth.

In what follows in this Section we assume the next two conditions to be fulfilled.

Condition 15.20 *The Riemannian manifold M is complete in the usual sense (see Definition 1.49 and the Hopf-Rinow Theorem (Theorem 3.68)). The Ricci tensor, the linear operator $\widehat{\mathrm{Ric}}(m) : T_mM \to T_mM$, is bounded uniformly in m with respect to the operator norm generated in the tangent spaces by the Riemannian metric $\langle\cdot,\cdot\rangle$. The vector field $\mathrm{tr}\nabla\widehat{\mathrm{Ric}}(\widehat{\mathrm{Ric}})$ on M is also uniformly bounded with respect to the norm generated by the Riemannian metric.*

Condition 15.21 *The vector field $\bar{\alpha}_0(t,m)$ and the tensor field $\bar{\alpha}_1(t,m)$ are Borel measurable jointly in t and m, the vector field $\mathrm{tr}\nabla\bar{\alpha}_1(t,x)(\bar{\alpha}_1)$ exists and is also Borel measurable jointly in t and m, and there exists a constant $C > 0$ such that*

$$\int_0^l \left(\|\bar{\alpha}_0(t,m(t))\|^2 + \|\bar{\alpha}_1(t,m(t))\|^2 + \|\mathrm{tr}\nabla\bar{\alpha}_1(t,x)(\bar{\alpha}_1)\|^2\right) \mathrm{d}t < C$$

for any continuous curve $m(t)$ on M, $t \in [0,l]$, and where $\|\bar{\alpha}_1(t,m)\|$ is the operator norm.

Remark 15.22. From Theorem 7.80 it follows that under Condition 15.20 the Riemannian manifold M is stochastically complete.

Let $m_0 \in M$ and introduce the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \nu)$ where $\tilde{\Omega} = C_0([0,l], T_{m_0}M)$, $\tilde{\mathcal{F}}$ is the σ-algebra generated by the cylinder sets and ν is the Wiener measure (recall that $T_{m_0}M$ is an n-dimensional Euclidean space with respect to the metric tensor $\langle\cdot,\cdot\rangle$ at m_0). Denote by $\tilde{\mathcal{B}}_t$ the σ-algebra generated by the cylinder sets with bases over $[0,t]$, all the $\tilde{\mathcal{B}}_t$ being completed by the sets of ν-measure zero.

Recall that the coordinate process $W(t,\omega(\cdot)) = \omega(t)$, $\omega(\cdot) \in \tilde{\Omega}$, is a Wiener process in $T_{m_0}M$. Since by Remark 15.22 M is stochastically complete, the Itô development $R_I W(t)$ is well-defined for $t \in [0,l]$ and Riemannian parallel translation is also well-defined along $R_I W(t)$ (see Section 7.6.1). Further on we will use the operator $\Gamma_{t,s}$ of parallel translation introduced in Section 7.7.1.

Thus for any vector or tensor field $v(t,m)$ we can consider the vector (tensor) field $\Gamma_{0,t} v(t, R_I W(t))$ at m_0 to be the result of parallel translation of $v(t, R_I W(t))$ along $R_I W(\cdot)$ from the (random) point $R_I W(t)$ to the point $R_I W(0) = m_0$. Since $R_I W(t)$ is an extension of Cartan's development R_c from the class of piecewise smooth curves onto ν-almost all continuous curves in $T_{m_0}M$ and the analogous fact is valid for parallel translation along $R_I W(\cdot)$ (see Remarks 7.67 and 7.84), $\Gamma_{0,t} v(t, R_I W(t)$ is determined along ν-almost all continuous curves in $T_{m_0}M$. This field along $\omega(\cdot) \in \tilde{\Omega}$ will be denoted by $(\Gamma_{0,t} v)(t, \omega(\cdot))$. Note that it may depend on t (which is compatible with the notation) even if $v(m)$ is autonomous.

Pick $t_0 \in (0,l)$ and for $t \in [0,l]$ consider the function $t_0(t)$ defined by formula (15.13).

Let $a_0 \in T_{m_0}M$ and consider in $T_{m_0}M$ the equation

$$\begin{aligned} a(t) = a_0 &+ \int_0^t (\Gamma_{0,s}\bar{\alpha}_0)(s, W(\cdot))\mathrm{d}s - \frac{1}{4}\int_0^t \left(\Gamma_{0,s}\mathrm{tr}\nabla\widehat{\mathrm{Ric}}\left(\widehat{\mathrm{Ric}}\right)\right)(s, W(\cdot))\mathrm{d}s \\ &+\frac{1}{2}\int_0^t (\Gamma_{0,s}\mathrm{tr}\nabla\bar{\alpha}_1(\bar{\alpha}_1))(s, W(s))\mathrm{d}s \\ &+\frac{1}{2}\int_0^t \left(\left(\Gamma_{0,s}\widehat{\mathrm{Ric}}\right)(s, W(\cdot)) - t_0(s)\right) a(s)\mathrm{d}s \\ &+\int_0^t \left((\Gamma_{0,s}\bar{\alpha}_1)(s, W(\cdot)) - \frac{1}{2}\left(\Gamma_{0,s}\widehat{\mathrm{Ric}}\right)(s, W(\cdot)) + \frac{1}{2}t_0(s)\right)\mathrm{d}W(s) \\ &-\frac{1}{2}t_0(t)W(t) \end{aligned} \tag{15.31}$$

(an analog of equation (15.14)).

Equation (15.31) has a unique strong solution for $t \in [0,l]$. Indeed, the coefficients of (15.31) are either Lipschitz continuous and have linear growth with respect to a, or do not depend on a. Since the solution is strong, it exists for any Wiener process and is non-anticipative with respect to $\mathcal{P}_t^W$. In what follows we will consider $a(t)$ for the realization of $W(t)$ as the coordinate process on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \nu)$. From Condition 15.21 it follows that $a(t)$ satisfies (7.54).

Consider the probability density $\theta(l)$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ defined for the above a by formula (8.33) (cf. Section 15.1.2). Introduce the measure $\mu^{\zeta} = \theta \mathrm{d}\nu$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ and denote by $\zeta(t)$ the coordinate process on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu^{\zeta})$. Then (see Sections 8.3 and 15.1.2) $\zeta(t)$ is expressed in the form

$$\zeta(t) = \int_0^t a(s)\mathrm{d}s + w(t) \tag{15.32}$$

where $w(t)$ is a Wiener process on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mu^{\zeta})$ adapted to $\mathcal{P}_t^W = \mathcal{P}_t^{\zeta} = \mathcal{B}_t$. Since $\zeta(t)$ and $W(t)$ coincide as coordinate processes on $(\tilde{\Omega}, \tilde{\mathcal{F}})$, (15.31) turns into

$$\begin{aligned} a(t) = a_0 &+ \int_0^t (\Gamma_{0,s}\bar{\alpha}_0)(s,\zeta(\cdot))\mathrm{d}s - \frac{1}{4}\int_0^t \left(\Gamma_{0,s}\mathrm{tr}\nabla\widehat{\mathrm{Ric}}\left(\widehat{\mathrm{Ric}}\right)\right)(s,\zeta(\cdot))\mathrm{d}s \\ &+\frac{1}{2}\int_0^t (\Gamma_{0,s}\mathrm{tr}\nabla\bar{\alpha}_1(\bar{\alpha}_1))(s,\zeta(\cdot))\mathrm{d}s + \int_0^t (\Gamma_{0,s}\bar{\alpha}_1)(s,\zeta(\cdot))\circ a(s)\mathrm{d}s \\ &+ \int_0^t ((\Gamma_{0,s}\bar{\alpha}_1)(s,\zeta(\cdot)) - \frac{1}{2}\left(\Gamma_{0,s}\widehat{\mathrm{Ric}}\right)(s,\zeta(\cdot)) + \frac{1}{2}t_0(s))\mathrm{d}w(s) \\ &-\frac{1}{2}t_0(t)\zeta(t) \end{aligned} \tag{15.33}$$

(an analog of equation (15.16)).

Since by Remark 15.22 M is stochastically complete and since $a(t)$ satisfies (7.54), from Theorem 7.98 it follows that for $\zeta(t)$ the Itô development $\xi(t) = R_I\zeta(t)$ is well-defined on the entire interval $t \in [0, l]$.

Theorem 15.23 *For $t \in (t_0, l)$ the above-mentioned process $\xi(t)$ satisfies (15.29) with the force $\bar{\alpha}(t,m,X) = \bar{\alpha}_0(t,m) + \bar{\alpha}_1(t,m)X$, i.e., it is a trajectory of the stochastic mechanical system with that force.*

Proof. By construction $\xi(t)$ satisfies equation (9.22). This allows us to derive some technical statements which follow from the results of Section 7.7 and from those on the calculation of mean derivatives in Chapters 8 and 9.

Lemma 15.24

(i) $D\xi(t) = E_t^{\xi}(\Gamma_{t,0}a(t))$.

(ii) $D_*\xi(t) = E_t^{\xi}(\Gamma_{t,0}a(t)) + E_t^{\xi}\left[\Gamma_{0,t}\left(\frac{\zeta(t)}{t} - \kappa(t)\right)\right]$, *where $\kappa(t)$ is as defined in Lemma* 8.35.

Proof. Assertion (i) is in fact formula (9.25), and (ii) is (9.29). □

Corollary 15.25

(i) $D_S\xi(t) = E_t^{\xi}(\Gamma_{t,0}a(t)) + \frac{1}{2}E_t^{\xi}\left[\Gamma_{t,0}\left(\frac{\zeta(t)}{t} - \kappa(t)\right)\right]$.

(ii) $\quad D_A\xi(t) = -\frac{1}{2}E_t^\xi\left[\Gamma_{t,0}\left(\frac{\zeta(t)}{t} - \kappa(t)\right)\right].$

Corollary 15.25 can be obtained from formulae (9.30) and (9.31).

Lemma 15.26

(i) $\quad \mathbf{D}D_*\xi(t) = D^\xi\left(\Gamma_{t,0}a(t)\right) + E_t^\xi\left(\Gamma_{t,0}\frac{a(t)}{t}\right) - E_t^\xi\left(\Gamma_{t,0}\frac{\zeta(t)}{t^2}\right).$

(ii) $\quad \mathbf{D}_*D\xi(t) = D_*^\xi(\Gamma_{t,0}a(t)).$

The assertion of Lemma 15.26 follows from Lemma 15.24, Lemma 8.36 (which can evidently be generalized to the processes in M) and Lemma 8.38.

So the proof of Theorem 15.23 is reduced to the calculation of mean derivatives for $\Gamma_{t,0}a(t)$ which can be done by calculating the derivatives for the summands in (15.33).

Lemma 15.27 *For $t \geq t_0$:*

(i) $\quad \mathbf{D}_*^\xi\Gamma_{t,0}\left(\int_0^t t_0(s)\mathrm{d}w(s)\right) = E_t^\xi\left[\Gamma_{t,0}\left(\frac{\zeta(t)}{t^2} - \frac{\kappa(t)}{t}\right)\right];$

(ii) $\quad \mathbf{D}^\xi\Gamma_{t,0}\left(\frac{\zeta(t)}{t}\right) = E_t^\xi\left[\Gamma_{t,0}\left(\frac{a(t)}{t} - \frac{\zeta(t)}{t^2}\right)\right];$

(iii) $\quad \mathbf{D}_*\xi\Gamma_{t,0}\left(\frac{\zeta(t)}{t}\right) = E_t^\xi\left[\Gamma_{t,0}\left(\frac{a(t)}{t} - \frac{\kappa(t)}{t}\right)\right].$

The proof of Lemma 15.27(i) is analogous to that of Lemma 15.8(ii) with a modification based on formulae (9.15). The proofs of (ii) and (iii) are analogous to that of Lemma 8.37.

Lemma 15.28

(i) $$\mathbf{D}_*^\xi\Gamma_{t,0}\left(\int_0^t\left(\Gamma_{0,t}\bar{\alpha}_1\right)(s,\zeta(\cdot))\,\mathrm{d}w(s)\right) = \\ -\Gamma_{t,0}\left(\Gamma_{0,t}\mathrm{tr}\nabla\bar{\alpha}_1\left(\bar{\alpha}_1\right)\right)(t,\zeta(t)) \\ +\Gamma_{t,0}\left(\Gamma_{0,t}\bar{\alpha}_1\right)(t,\zeta(\cdot))\circ E_t^\xi\left[\Gamma_{t,0}\left(\frac{\zeta(t)}{t} - \kappa(t)\right)\right].$$

(ii) $$\mathbf{D}_*^\xi\Gamma_{t,0}\left(\int_0^t\left(\Gamma_{0,t}\widehat{\mathrm{Ric}}\right)(s,\zeta(\cdot))\,\mathrm{d}w(s)\right) = \\ -\Gamma_{t,0}\left(\Gamma_{0,t}\mathrm{tr}\nabla\widehat{\mathrm{Ric}}\left(\widehat{\mathrm{Ric}}\right)\right)(s,\zeta(s)) \\ +\Gamma_{t,0}\left(\Gamma_{0,t}\widehat{\mathrm{Ric}}\right)(t,\zeta(\cdot))\circ E_t^\xi\left[\Gamma_{t,0}\left(\frac{\zeta(t)}{t} - \kappa(t)\right)\right].$$

To prove Lemma 15.28 one should apply formulae (9.15), the definitions of the integrands and the expression for Itô integrals via the backward integrals (6.7) and (6.26) (cf. Lemma 15.8(i)).

All other summands in (15.32) are differentiated directly according to (9.15).

By the definition of the processes $\zeta(t)$ and $\xi(t)$ and of the operator of parallel translation $\Gamma_{t,s}$ the following relations hold:

$$\Gamma_{t,0}(\Gamma_{0,t}\bar{\alpha})(t,\zeta(\cdot)) = \bar{\alpha}(t,\xi(t));$$
$$\Gamma_{t,0}(\Gamma_{0,t}\widehat{\mathrm{Ric}})(t,\zeta(\cdot)) = \widehat{\mathrm{Ric}}(\xi(t)).$$

Thus by Lemmas 15.26–15.28 and formula (15.33)

$$
\begin{aligned}
&\frac{1}{2}(\mathbf{D}D_* + \mathbf{D}_* D)\xi(t)\\
&= \bar{\alpha}_0(t,\xi(t)) + \bar{\alpha}_1(t,\xi(t)) \circ a(t) + \frac{1}{2}E_t^{\xi}\left[\Gamma_{t,0}\left(\frac{\zeta(t)}{t} - \kappa(t)\right)\right]\\
&\quad + \frac{1}{2}\widehat{\mathrm{Ric}}(\xi(t)) \circ \left\{-E_t^{\xi}\left[\Gamma_{t,0}\left(\frac{\zeta(t)}{t} - \kappa(t)\right)\right]\right\}.
\end{aligned}
\tag{15.34}
$$

Taking into account Corollary 15.25 one sees that (15.34) coincides with (15.29). Theorem 15.23 follows. □

Remark 15.29. As in Remark 15.11 we should point out that $\xi(t)$ does not satisfy (15.29) for $t \in (0, t_0)$ because here $t_0(t) = \frac{1}{t_0}$ and $E_t^{\xi}(\Gamma_{t,0}\frac{\zeta(t)}{t\cdot t_0}) \neq E_t^{\xi}(\Gamma_{t,0}\frac{\zeta(t)}{t^2})$. Thus $\xi(t)$ can be interpreted only as a trajectory of the stochastic mechanical system beginning at the instant t_0 from the random configuration $\xi(t_0)$ with the initial mean forward derivative $E_t^{\xi}(\Gamma_{t,0}a(t_0))$ (see the details in Remark 15.10). A certain analog of this situation with 'big bang' will be described below in Section 15.3.2, which is devoted to general relativity.

Remark 15.30. One can easily see that $E_t^{\xi}(a(t))$ is the hodograph of the forward mean derivative for the process $\xi(t)$ and that equation (15.33) is a direct analog of the velocity hodograph equation (11.20).

We now turn to the proof of an existence theorem analogous to Theorem 15.14.

Theorem 15.31 *Let Conditions* 15.20 *and* 15.21 *be fulfilled. Let η be a random element taking values in M with density $\rho(m) \neq 0$ for all $m \in M$. Let a be a bounded Borel measurable vector field on M. Then there exists a solution $\xi(t)$ of* (15.29) *with initial conditions $\xi(0) = \eta$ and $D\xi(0) = a(\eta)$. If there exists a $\delta > 0$ such that $\rho_0(m) > \delta$ for all $m \in M$ and* $\mathrm{Grad}\log\rho_0$ *(see* (8.20)*) is bounded, then there exists a solution $\xi(t)$ of* (15.29) *with initial conditions $\xi(0) = \eta$ and $D_S\xi(0) = a(\eta)$.*

Proof. Denote by $W(t)$ the coordinate process on the probability space $(C^0([0,l],\mathbb{R}^n), \mathcal{F}, \nu)$ taking values in $\mathbb{R}^n$ where $\mathcal{F}$ is the σ-algebra generated by cylinder sets and ν is a Wiener measure. As above, it is clear that $W(t)$ is a standard Wiener process in $\mathbb{R}^n$. Take some Borel measurable cross-section b_0 of the bundle OM and consider the Itô development $R_I W(t)$ with initial data $b_0(\eta)$ (see Definition 7.68). $R_I W(t)$ exists for $t \in [0,l]$ by Theorem 7.99 since, by Remark 15.22, from Condition 15.20 it follows that M is stochastically complete. Consider the following objects along $R_I W(t)$: the vector field $u^W(t,m)$ constructed for $R_I W(t)$ by formula (8.20), the vector field $D^W u^W(t,m)$ and the tensor field $\nabla u^W(t,m) = K \circ T(u^W(t,m))$ (the operator of covariant derivative applied to $u^W(t,m)$). Using the construction of parallel translation along $R_I W(t)$ (see Section 7.6), in this section the operator $\Gamma_{t,0}$ is applied to a tensor $v(t, R_I W(t))$ yielding the tensor field $\Gamma_{t,0} v(t, W(t))$ along $W(t)$ in $\mathbb{R}^n$, which is well-defined along a.s. all continuous curves in $\mathbb{R}^n$ with respect to Wiener measure (see Section 7.6).

Denote by a_0 the vector $b(0)^{-1}[a(\eta) + u^W(0, W(0))]$ and consider the following stochastic differential equation in $\mathbb{R}^n$:

$$\begin{aligned} a(t) = a_0 &+ \int_0^t \Gamma_{0,s}\alpha_0(s, W(s))\mathrm{d}s + \int_0^t \Gamma_{0,s}\alpha_1(s, W(s)) \circ \mathrm{d}W(s) \\ &+ \int \Gamma_{0,s} D^W u^W(s, W(s))\mathrm{d}s + \int_0^t \big(\Gamma_{0,s}\nabla u^W(s, W(s)) \circ a(s)\big)\, \mathrm{d}s \\ &- \frac{1}{2}\int_0^t \Gamma_{0,s}\widehat{\mathrm{Ric}}(s, W(s)) \circ \mathrm{d}W(s), \end{aligned} \tag{15.35}$$

where $\circ\, \mathrm{d}W(s)$ denotes Stratonovich integration in composition with parallel translation, i.e.,

$$\begin{aligned} &\int_0^t \Gamma_{0,s} A(s, W(s)) \circ \mathrm{d}W \\ &= \int_0^t \Gamma_{0,s} A(s, W(s))\mathrm{d}W(s) + \frac{1}{2}\int_0^t \Gamma_{0,s}\mathrm{tr}\nabla A(t,m)(A)\mathrm{d}s \end{aligned}$$

for a $(1,1)$-tensor field A on M.

Since (15.35) is linear in a, it has a unique solution $a(t)$. From Condition 15.21 it follows that $a(t)$ satisfies (7.54). Set

$$\theta(t) = \exp\left\{-\frac{1}{2}\int_0^t a(s)^2\mathrm{d}s + \int_0^t a\cdot \mathrm{d}W.\right\}$$

Under these assumptions $\theta(t)$ is a martingale with $E\theta(t) = 1$. Thus we can introduce a new probability measure μ on $(C^0([0,l], \mathbb{R}^n), \mathcal{F})$ by the relation $\mathrm{d}\mu = \theta(l)\mathrm{d}\nu$. Denote by $\zeta(t)$ the coordinate process on $(C^0([0,l], \mathbb{R}^n), \mathcal{F}, \mu)$. By Girsanov's theorem $\zeta(t) = \int_0^t a(s)\mathrm{d}s + w(t)$ where $w(t)$ is a Wiener process on $C^0([0,l], \mathbb{R}^n), \mathcal{F}, \mu)$ adapted to $\zeta(t)$ and $a(t)$ is not anticipative with respect to $\mathcal{P}_t^\zeta$, i.e., ζ is a process satisfying the hypothesis of Theorem 7.99.

Since M is stochastically complete and $a(t)$ satisfies (7.54), there exists an Itô development $\xi(t) = R_I\zeta(t)$ with initial condition η.

Lemma 15.32 *The process $\xi(t)$ is the solution of equation (15.29) with initial conditions $\xi(0) = \eta$ and $D_S\xi(0) = a(\eta)$.*

Proof. From the definition of Itô development it evidently follows that

$$\Gamma_{t,0}\Gamma_{0,t}u^W(t, \xi(t)) = u^W(t, \xi(t)); \qquad \Gamma_{t,0}\Gamma_{0,t}\nabla u^W \circ a = \nabla_{\Gamma_{t,0}a} u^W$$

$$\Gamma_{t,0}\Gamma_{0,t}\alpha_0 = \alpha_0; \quad \Gamma_{t,0}\Gamma_{0,t}\alpha_1 = \alpha_1; \quad \Gamma_{t,0}\Gamma_{0,t}\mathrm{Ric} = \mathrm{Ric};$$
$$\Gamma_{t,0}\Gamma_{0,t}\mathrm{tr}\nabla\alpha_1(t,\xi(t))(\alpha_1) = \mathrm{tr}\nabla\alpha_1(t,\xi(t))(\alpha_1);$$
$$\Gamma_{t,0}\Gamma_{0,t}\mathrm{tr}\nabla\mathrm{Ric}(t,\xi(t))(\mathrm{Ric}) = \mathrm{tr}\nabla\mathrm{Ric}(t,\xi(t))(\mathrm{Ric})$$

and so on. Thus the fact that $\xi(t)$ satisfies (15.29) follows by direct calculation. □

To obtain a solution with initial conditions $\xi(0) = \eta$ and $D\xi(0) = a(\eta)$, in the above construction one should take a_0 in the form $b(0)^{-1}(a(\eta))$. □

15.3 Relativistic Stochastic Mechanics

15.3.1 Stochastic mechanics in Minkowski space

In this section we present a modification of stochastic mechanics which is well-posed for describing relativistic particles in Minkowski space (see Example 13.4). In particular we obtain generalizations of the existence theorems of Section 15.1 to the relativistic case. In our constructions, following [103], we apply the relativistic definition of mean derivatives suggested by Guerra and Ruggiero [55, 142] and Zastawniak's idea of transition from stochastic processes in Minkowski space to those in the underlying Euclidean space (see, e.g., [237, 238]). The material of this section forms the basis for the construction of stochastic mechanics in the general relativistic case in Section 15.3.2 below (see [130]).

For the sake of simplicity, in this section we work in a system of units in which the speed of light $c = 1$. We also consider particles with rest mass 1.

In this subsection we denote by M^4 the Minkowski space with inner product $(\cdot,\cdot)$ (see Example 13.4). Choose a certain orthonormal frame (with respect to $(\cdot,\cdot)$) in M^4. Let $\mathbb{R}^4$ be the (underlying) Euclidean space in which the above frame is orthonormal in the Euclidean sense. The (positive-definite) inner product in $\mathbb{R}^4$ will be denoted by a dot $\cdot$. The main idea here is to consider Itô processes $\xi(\tau)$ of the form (8.15) as processes in M^4 while Wiener processes are defined with respect to $\mathbb{R}^4$. This idea originated in the work of Zastawniak (see [237]). Here τ is an invariant parameter which plays the role of proper time. For such $\xi(\tau)$, according to an idea of Guerra and Ruggiero [55, 142], we define the *relativistic forward mean derivative* $D_+\xi(\tau)$ and *relativistic backward mean derivative* $D_-\xi(\tau)$ as follows:

$$\begin{aligned} &D_+\xi(\tau) \\ &= \lim_{\Delta\tau\downarrow 0} E\left(\frac{\xi(\tau+\Delta\tau)-\xi(\tau)}{\Delta\tau}\,\middle|\,\mathcal{N}_t^{\xi}, (\xi(\tau+\Delta\tau)-\xi(\tau))^2 \le 0\right) \\ &\quad + \lim_{\Delta\tau\downarrow 0} E\left(\frac{\xi(\tau)-\xi(\tau-\Delta\tau)}{\Delta\tau}\,\middle|\,\mathcal{N}_t^{\xi}, (\xi(\tau)-\xi(\tau-\Delta\tau))^2 \ge 0\right); \end{aligned} \tag{15.36}$$

$$
\begin{aligned}
&D_-\xi(\tau)\\
&= \lim_{\Delta\tau\downarrow 0} E\left(\frac{\xi(\tau)-\xi(\tau-\Delta\tau)}{\Delta\tau} \middle| \mathcal{N}_t^\xi, (\xi(\tau)-\xi(\tau-\Delta\tau))^2 \le 0 \right) \\
&\quad + \lim_{\Delta\tau\downarrow 0} E\left(\frac{\xi(\tau+\Delta\tau)-\xi(\tau)}{\Delta\tau} \middle| \mathcal{N}_t^\xi, (\xi(\tau+\Delta\tau)-\xi(\tau))^2 \ge 0 \right).
\end{aligned} \tag{15.37}
$$

It should be noted that (15.36) and (15.37) are covariant under the Lorentz transformations. Below we will deal with both D_+ and D_- on the one hand and D, D_* and D_2 on the other hand, applied to a process $\xi(\tau)$ where D, D_* and D_2 will be calculated in $\mathbb{R}^4$ by formulae (8.1), (8.2) and (8.13), respectively.

Of course, $D_+\xi(\tau) = Y_+(\tau,\xi(\tau))$ for the vector field

$$
\begin{aligned}
&Y_+(\tau,x)\\
&= \lim_{\Delta\tau\downarrow 0} E\left(\frac{\xi(\tau+\Delta\tau)-\xi(\tau)}{\Delta\tau} \middle| \xi(\tau)=x, (\xi(\tau+\Delta\tau)-\xi(\tau))^2 \le 0 \right) \\
&\quad + \lim_{\Delta\tau\downarrow 0} E\left(\frac{\xi(\tau)-\xi(\tau-\Delta\tau)}{\Delta\tau} \middle| \xi(\tau)=x, (\xi(\tau)-\xi(\tau-\Delta\tau))^2 \ge 0 \right)
\end{aligned}
$$

and $D_-\xi(\tau) = Y_-(\tau,\xi(\tau))$ for the vector field

$$
\begin{aligned}
&Y_-(\tau,x)\\
&= \lim_{\Delta\tau\downarrow 0} E\left(\frac{\xi(\tau)-\xi(\tau-\Delta\tau)}{\Delta\tau} \middle| \xi(\tau)=x, (\xi(\tau)-\xi(\tau-\Delta\tau))^2 \le 0 \right) \\
&\quad + \lim_{\Delta\tau\downarrow 0} E\left(\frac{\xi(\tau+\Delta\tau)-\xi(\tau)}{\Delta\tau} \middle| \xi(\tau)=x, (\xi(\tau+\Delta\tau)-\xi(\tau))^2 \ge 0 \right).
\end{aligned}
$$

Consider the coordinate decomposition of some process $\zeta(\tau)$ with respect to the frame mentioned above: $\zeta(\tau) = (\zeta^0(\tau), \bar{\zeta}(\tau))$ where $\bar{\zeta}(\tau) = (\zeta^1(\tau), \zeta^2(\tau), \zeta^3(\tau))$. Then clearly

$$
D_+\zeta(\tau) = (D^\zeta \zeta^0(\tau), D_*^\zeta \bar{\zeta}(\tau)) \tag{15.38}
$$

and

$$
D_-\zeta(\tau) = (D_*^\zeta \zeta^0(\tau), D^\zeta \bar{\zeta}(\tau)), \tag{15.39}
$$

where $D\zeta$ and $D_*\zeta$ are defined in $\mathbb{R}^4$ by (8.7) and (8.8), respectively.

Let us introduce the relativistic current velocity $\bar{v}^\zeta(\tau) = \frac{1}{2}(D_+\zeta(\tau) + D_-\zeta(\tau))$ and the relativistic osmotic velocity $\bar{u}^\zeta(\tau) = \frac{1}{2}(D_+\zeta(\tau) - D_-\zeta(\tau))$. We have $\bar{v}^\zeta(\tau) = \bar{v}^\zeta(\tau,\zeta(\tau))$ where $\bar{v}^\zeta(\tau,x) = \frac{1}{2}(Y_+(\tau,x) + Y_-(\tau,x))$ and $\bar{u}^\zeta(\tau) = \bar{u}^\zeta(\tau,\zeta(\tau))$ where $\bar{u}^\zeta(\tau,x) = \frac{1}{2}(Y_+(\tau,x) - Y_-(\tau,x))$.

Lemma 15.33 *The relativistic current velocity $\bar{v}(\tau)$ and the usual current velocity $v(\tau)$, calculated in $\mathbb{R}^4$ by Definition* 8.16*, coincide.*

To prove that $\bar{v}^\zeta(\tau) = v^\zeta(\tau)$, compare formulae (15.38) and (15.39).

Lemma 15.34 *The relativistic osmotic velocity takes the form*

$$\bar{u}^\zeta(\tau) = \left(D_A^\zeta \zeta^0(\tau), -D_A^\zeta \bar{\zeta}(\tau)\right)$$

while the usual osmotic velocity, calculated in $\mathbb{R}^4$ *by Definition* 8.16, *takes the form*

$$u^\zeta(\tau) = \left(D_A^\zeta \zeta^0(\tau), D_A^\zeta \bar{\zeta}(\tau)\right),$$

and so they do not coincide.

Lemma 15.34 follows immediately from (15.38) and (15.39).

Lemma 15.35

$$\bar{u}^\zeta(\tau, x) = -\frac{1}{2}\sigma^2 \text{Grad } \log\, \rho^\zeta(\tau, x) \tag{15.40}$$

where $\rho^\zeta(\tau, x)$ *is the density of* $\zeta(\tau)$ *with respect to Lebesgue measure, defined in Section* 8.1, *and* Grad *is the gradient calculated with respect to the inner product* $(\cdot,\cdot)$ *in* M^4.

Lemma 15.35 follows from the above coordinate decomposition for $\bar{u}^\zeta(\tau)$, formula (8.18) and the definition of the inner product $(\cdot,\cdot)$ in M^4 (note in particular, that $\text{Grad}\, f = (-\frac{\partial f}{\partial x^0}, \frac{\partial f}{\partial x^1}, \frac{\partial f}{\partial x^2}, \frac{\partial f}{\partial x^3})$ while $\text{grad}\, f = (\frac{\partial f}{\partial x^0}, \frac{\partial f}{\partial x^1}, \frac{\partial f}{\partial x^2}, \frac{\partial f}{\partial x^3})$ for any function f on M^4).

Lemma 15.36 *Formula* (8.19) *remains true for* $\bar{v}^\zeta$ *and* ρ^ζ

Indeed, this follows from the fact that $\bar{v}^\zeta(\tau) = v^\zeta(\tau)$ (see above).

For a vector field $Z(\tau, x)$ on M^4 define the relativistic forward $D_+ Z(\tau, \zeta(\tau))$ and backward $D_- Z(\tau, \zeta(\tau))$ mean derivatives along $\zeta(\tau)$ by the formulae

$$\begin{aligned} &D_+ Z(\tau, \zeta(\tau)) = \\ &\lim_{\Delta\tau \downarrow 0} E\left(\frac{Z(\tau+\Delta\tau, \zeta(\tau+\Delta\tau)) - Z(\tau, \zeta(\tau))}{\Delta\tau} \middle| \mathcal{N}_t^\zeta, (\zeta(\tau+\Delta\tau) - \zeta(\tau))^2 \le 0 \right) \\ &+ \\ &\lim_{\Delta\tau \downarrow 0} E\left(\frac{Z(\tau, \zeta(\tau)) - Z(\tau-\Delta\tau, \zeta(\tau-\Delta\tau))}{\Delta\tau} \middle| \mathcal{N}_t^\zeta, (\zeta(\tau) - \zeta(\tau-\Delta\tau))^2 \ge 0 \right) \end{aligned} \tag{15.41}$$

and

$$\begin{aligned} &D_- Z(\tau, \zeta(\tau)) = \\ &\lim_{\Delta\tau \downarrow 0} E\left(\frac{Z(\tau, \zeta(\tau)) - Z(\tau-\Delta\tau, \zeta(\tau-\Delta\tau))}{\Delta\tau} \middle| \mathcal{N}_t^\xi, (\zeta(\tau) - \zeta(\tau-\Delta\tau))^2 \le 0 \right) \\ &+ \\ &\lim_{\Delta\tau \downarrow 0} E\left(\frac{Z(\tau+\Delta\tau, \zeta(\tau+\Delta\tau)) - Z(\tau, \zeta(\tau))}{\Delta\tau} \middle| \mathcal{N}_t^\zeta, (\zeta(\tau+\Delta\tau) - \zeta(\tau))^2 \ge 0 \right) \end{aligned} \tag{15.42}$$

(a modification of (8.22) and (8.23), respectively, by analogy with (7.46) and (7.47)). Taking into account (8.24) and (8.25), one can easily derive

$$D_+Z(\tau,\zeta(\tau)) = \frac{\partial}{\partial\tau}Z + (Y_+,\nabla)Z + \frac{\sigma^2}{2}\Box Z,$$
$$D_Z(\tau,\xi(\tau)) = \frac{\partial}{\partial\tau}Z + (Y_-,\nabla)Z - \frac{\sigma^2}{2}\Box Z, \qquad (15.43)$$

where $\Box = -\frac{\partial^2}{\partial x^{0\,2}} + \frac{\partial^2}{\partial x^{1\,2}} + \frac{\partial^2}{\partial x^{2\,2}} + \frac{\partial^2}{\partial x^{3\,2}}$ is the wave operator (the d'Alembertian, cf. (8.24) and (8.25)).

Definition 15.37. The vector $\frac{1}{2}(D_+D_- + D_-D_+)\xi(\tau)$ is called the 4-*acceleration* of the process ξ at τ.

Below we use the ordinary quadratic mean derivative D_2 introduced in Definition 8.10. The introduction of a special relativistic quadratic mean derivative is not necessary in the problems under consideration for the following reasons.

Introduce "relativistic" increments of the process $\zeta(\tau)$ as follows:

$$\varDelta_+\zeta(\tau) = (\zeta^0(\tau+\varDelta\tau) - \zeta^0(\tau), \bar{\zeta}(\tau) - \bar{\zeta}(\tau-\varDelta\tau))$$

and

$$\varDelta_-\xi(\tau) = (\zeta^0(\tau) - \zeta^0(\tau-\varDelta\tau), \bar{\zeta}(\tau+\varDelta\tau) - \bar{\zeta}(\tau))$$

which are related to ordinary increments in the same way as relativistic forward and backward mean derivatives are related to ordinary forward and backward mean derivatives.

Lemma 15.38 *Let $\zeta(\tau)$ be an Itô diffusion type process on M^4 whose diffusion term is given by a Borel measurable field of linear operators $A(x)$: $\mathbb{R}^k \to M^4$ that is independent of the past and so depends only on $x \in M^4$. Denote by D_2 the ordinary quadratic mean derivative as in Definition* 8.10. *Then*

$$D_2\zeta(\tau) = \lim_{\varDelta\tau\to+0} \frac{\varDelta_+\zeta(\tau)\otimes\varDelta_+\zeta(\tau)}{\varDelta\tau}$$
$$= \lim_{\varDelta\tau\to+0} \frac{\varDelta_-\zeta(\tau)\otimes\varDelta_-\zeta(\tau)}{\varDelta\tau} = A(\zeta(\tau))A^*(\zeta(\tau)). \qquad (15.44)$$

The proof of Lemma 15.38 is based on the fact that in the case under consideration the diffusion term of $\zeta(\tau)$ and of the process with inverse time direction coincide. Note that in (15.44) we do not use conditional expectation since $A(x)$ is Borel measurable and does not depend on the past.

We should recall (see Remark 13.20) that in relativistic dynamics the 4-force necessarily depends on the 4-velocity. In addition, by definition it is independent of the proper time τ since it is an absolute object (independent of a reference frame: recall that proper time is well-defined only in the reference frame of an observer). This is why, below, we consider a force field on M^4 only in the form $\bar{\alpha} = \bar{\alpha}(x)X$ where $\bar{\alpha}(x)$ is a $(1,1)$-tensor, i.e., a linear operator in M^4, depending on $x \in M^4$ (an analog of $\bar{\alpha}_1$ from Section 15.1.2).

Definition 15.39. (cf. Definition 15.2) The *relativistic Newton-Nelson equation* with 4-force $\bar{\alpha}(x)X$ is the system in the Minkowski space M^4 of the form

$$\begin{cases} \frac{1}{2}(D_+D_- + D_-D_+)\xi(t) = \bar{\alpha}(\xi(\tau))v^{\xi}(\tau) \\ D_2\xi(t) = \frac{\hbar}{\mathrm{m}}I \end{cases} \tag{15.45}$$

where m is the mass (rest mass) of a particle.

We shall look for solutions of (15.45) among all Itô processes. From (15.45) one can easily see that the stochastic-mechanical world line of a relativistic particle with rest mass m is an Itô process $\xi(\tau)$ in $\mathbb{R}^4$ of the form (8.15) with $\frac{\sigma^2}{2} = \frac{\hbar}{2\mathrm{m}}$ such that it satisfies the first equation of (15.45) as a process in M^4.

Recall that everywhere in this Subsection we assume $\mathrm{m} = 1$ without loss of generality.

Theorem 15.40 *For $\xi(\tau)$ as above the following equality holds*

$$\frac{1}{2}(DD_* + D_*D)\xi(\tau) = \frac{1}{2}(D_+D_- + D_-D_+)\xi(\tau). \tag{15.46}$$

Proof. This statement is proved by straightforward calculation. Indeed, by formulae (15.38), (15.39), (15.41) and (15.42) we have

$$D_+D_-\xi(\tau) = \big(D^{\xi}_*D^{\xi}\xi^0(\tau), D^{\xi}D^{\xi}_*\bar{\xi}(\tau)\big)$$

and

$$D_-D_+\xi(\tau) = \big(D^{\xi}D^{\xi}_*\xi^0(\tau), D^{\xi}_*D^{\xi}\bar{\xi}(\tau)\big)$$

which leads to (15.46). □

Let the force field $\bar{\alpha}(x)X$ on M^4 satisfy Condition 15.7 (with $\bar{\alpha}_0(\tau, x) = 0$ and $\bar{\alpha}_1 = \bar{\alpha}$) relative to the Euclidean metric of $\mathbb{R}^4$. Then for any initial forward derivative $a_0 \in M^4$ and for the function $t_0(\tau)$ defined by formula (13.13), we can apply the construction of $a(\tau)$ and θ of Section 15.1.2 and so obtain the corresponding Itô process $\xi(\tau)$ in $\mathbb{R}^4$.

The following statement is a simple corollary of Theorems 15.9 and 15.40.

Theorem 15.41 *For $\tau \in (t_0, l)$ the process $\xi(\tau)$ in M^4, mentioned above, satisfies* (15.45) *and so it is a stochastic-mechanical world line.*

Let the tensor field $\alpha(x)$ be a closed 2-form, i.e., $\alpha = \mathrm{d}\omega$ where ω is a 1-form on M^4 (a particular case is an electromagnetic field $F = \mathrm{d}A$, see Section 13.1.5). The trajectory of the stochastic-mechanical system with 4-force $\bar{\alpha}X$ on M^4, where as usual $\bar{\alpha}$ is physically equivalent to α with respect to $(\cdot,\cdot)$, is connected with the solution of the corresponding Klein-Gordon equation, so there is a link between relativistic stochastic mechanics and relativistic quantum mechanics. This is a generalization of the construction of Section 15.1.1 for the non-relativistic case. We describe it briefly, following the scheme of Guerra and Ruggiero [55, 142] developed for the electromagnetic field. In order to avoid confusion, in the case of an electromagnetic field we assume (without loss of generality) that the electric charge e of the particle is equal to 1.

Let $\xi(\tau)$ be a stochastic-mechanical world line. Let us make the additional assumption that there exists a function S such that $\sigma^2 \mathrm{Grad}\, S = \bar{v}^{\xi} + \bar{\omega}$ where $\bar{\omega}$ is the vector physically equivalent to ω with respect to $(\cdot,\cdot)$ and Grad is defined in Lemma 15.35. Introduce the complex-valued function $\Psi = \exp(R+\mathrm{i}S)$ where $R = \frac{1}{2}\log\rho^{\xi}$. Then, taking into account Lemmas 15.35 and 15.36, imitating the proof of Theorem 15.4 one can deduce that Ψ satisfies the equality

$$\mathrm{i}\,h\frac{\partial\Psi}{\partial\tau} = \frac{\sigma^2}{2}\left(\nabla - \frac{\mathrm{i}}{\hbar}\bar{\omega}, \nabla - \frac{\mathrm{i}}{\hbar}\bar{\omega}\right)\Psi \tag{15.47}$$

where in the right hand side we have the formal inner product (with respect to $(\cdot,\cdot)$) of differential operators, considered as formal vectors (cf. Lemma 8.20 where the definition of ∇ is also given). Note that $(\nabla,\nabla) = \Box$.

Suppose that $\Psi(\tau,x) = \exp(-\mathrm{i}\frac{\mathrm{m}}{2\hbar}\tau)\varphi(x)$. Then from (15.47) it follows that $\varphi(x)$ satisfies the equality

$$\left(\left(\nabla - \frac{\mathrm{i}}{\hbar}\bar{\omega}, \nabla - \frac{\mathrm{i}}{\hbar}\bar{\omega}\right) - \frac{1}{h^2}\right)\varphi = 0. \tag{15.48}$$

Formula (15.48) is the Klein-Gordon equation. In particular, if $\omega = 0$, (15.48) takes the form

$$\left(\Box - \frac{1}{h^2}\right)^2\varphi = 0. \tag{15.49}$$

The inverse construction is more straightforward. We describe it briefly according to [237, 238] where it was derived for the electromagnetic field.

Let a complex-valued function φ on M^4 satisfy (15.48). Represent it in the form $\varphi(x) = \exp(R+\mathrm{i}\,S)$ and consider the vector fields $\bar{v}(x) = \sigma^2\mathrm{Grad}S - \bar{\omega}$ and $\bar{u}(x) = \sigma^2\mathrm{Grad}\,R = \frac{1}{2}\sigma^2(\mathrm{Grad}\ \log\ \rho) = \frac{1}{2}\sigma^2\frac{\mathrm{Grad}\,\rho}{\rho}$ on M^4 for $\rho = \varphi\bar{\varphi}$.

Theorem 15.42 *Equation* (15.48) *is equivalent to the following system:*

$$\hbar\,\mathrm{div}\,\bar{u} + (\bar{u},\bar{u}) - (\bar{v},\bar{v}) = 1, \tag{15.50}$$

$$\hbar\,\mathrm{div}\,\bar{v} + 2(\bar{u},\bar{v}) = 0. \tag{15.51}$$

To prove Theorem 15.42 one should substitute $\varphi = \exp(R+\mathrm{i}\,S)$ into (15.48) and then, after natural transformations similar to those in the proof of Theorem 15.4, separate real and imaginary parts.

The vector fields $\bar{v}$ and $\bar{u}$ generate the vector fields $Y_+(x) = \bar{v}(x) + \bar{u}(x) = \operatorname{Grad} S + \operatorname{Grad} R + \bar{\omega}$ and $Y_-(x) = \bar{v}(x) - \bar{u}(x) = \operatorname{Grad} S - \operatorname{Grad} R + \bar{\omega}$. Having made the coordinate decompositions $Y+(x) = (Y_+^0, \bar{Y}_+)$ and $Y_-(x) = (Y_-^0, \bar{Y}_-)$, introduce the vector fields $Y(x) = (Y_+^0, \bar{Y}_-)$ and $Y_*(x) = (Y_-^0, \bar{Y}_+)$. Consider the diffusion process $\xi(\tau)$ on $\mathbb{R}^4$ satisfying the Itô equation

$$\mathrm{d}\xi(\tau) = Y(\xi(\tau))\mathrm{d}\tau + \sigma\,\mathrm{d}w(\tau) \tag{15.52}$$

where $\sigma > 0$ is as above.

Lemma 15.43 *For $\xi(\tau)$ from (15.52) $D_+\xi(\tau) = Y_+(\xi(\tau))$, $D_-\xi(\tau) = Y_-(\xi(\tau))$ and so $\bar{v}^\xi = \bar{v}$ and $\bar{u}^\xi = \bar{u}$.*

The assertion of Lemma 15.43 is a trivial consequence of the construction.

Theorem 15.44 *$\xi(\tau)$ from (15.52) satisfies the Newton-Nelson equation (15.45) with $\bar{\alpha}X$ physically equivalent to $\mathrm{d}\omega(\cdot, X)$ with respect to $(\cdot,\cdot)$.*

Proof. By straightforward calculations of $D_+D_-\xi(\tau)$ and $D_-D_+\xi(\tau)$ one obtains the formula

$$\frac{1}{2}(D_+D_- + D_-D_+)\xi(\tau) = -\frac{1}{2}\nabla(\hbar \operatorname{div} \bar{u} + (\bar{u},\bar{u}) - (\bar{v},\bar{v})) + \bar{\alpha}v \tag{15.53}$$

where $\bar{\alpha}$ is physically equivalent to $\mathrm{d}\omega(\cdot, X)$ with respect to $(\cdot,\cdot)$. From (15.50) it follows that the first summand on the right-hand side of (15.53) vanishes so that (15.53) is equal to (15.45) with the above $\bar{\alpha}$. □

Theorem 15.45 *Let the Klein-Gordon equation (15.48) be well-defined for a given 4-force $\bar{\alpha}$ and let $\xi(\tau)$ be a solution of the corresponding Newton-Nelson equation (15.45). Then*

$$E((v^\xi)^2 + (u^\xi)^2) = -1. \tag{15.54}$$

Proof. Clearly $(v^\xi)^2 + (u^\xi)^2 = (-\hbar \operatorname{div} u^\xi + (v^\xi)^2 - (u^\xi)^2) + (\hbar \operatorname{div} u^\xi + 2(u^\xi)^2)$. Direct calculations show that $(\hbar \operatorname{div} u^\xi + 2(u^\xi)^2) = \frac{\sigma^4}{2}\frac{\nabla^2\rho^\xi}{\rho^\xi}$ where $\nabla^2 = \nabla\cdot\nabla$ and so

$$E\left(\hbar \operatorname{div} u^\xi + 2(u^\xi)^2\right) = \frac{\sigma^4}{2}\int_{\mathbb{R}^4}\frac{\nabla^2\rho^\xi}{\rho^\xi}\rho^\xi\mathrm{d}\lambda = \frac{\sigma^4}{2}\int_{\mathbb{R}^4}\nabla^2\rho^\xi\mathrm{d}\lambda.$$

Note that $\xi(\tau)$, as a solution of (15.52), is a diffusion process in $\mathbb{R}^4$. Thus the *Kolmogorov-Fokker-Planck equation*

$$\frac{\partial}{\partial\tau}\rho^\xi = \frac{\sigma^2}{2}\nabla^2\rho^\xi - \operatorname{div}(\rho^\xi Y)$$

is valid for ρ^ξ where Y is as in (15.52). By construction ρ^ξ is independent of τ. Hence $E(\hbar \operatorname{div} u^\xi + 2(u^\xi)^2) = \sigma^2\int_{M^4}\operatorname{div}(\rho^\xi Y)\mathrm{d}\lambda$ where the latter integral is equal to zero by a standard application of the divergence formula. So (15.54) follows from (15.50). ☐

Remark 15.46. Equality (15.54) may be considered as the characteristic feature for τ to be a proper time along a stochastic-mechanical world-line $\xi(\tau)$ (cf. Definition 13.12). This idea, as well as the proof of (15.54), was apparently first suggested by Zastawniak [237].

15.3.2 Stochastic mechanics in the space-times of general relativity

In this subsection M^4 is a 4-dimensional Lorentz manifold with metric $(\cdot,\cdot)$ whose signature is $(-,+,+,+)$ (see Section 13.1.1). For the sake of simplicity we assume from now on that M^4 is orientable and oriented in time. In other words, a well-defined 'future' time direction is specified in every tangent space T_mM^4, $m \in M^4$.

Consider the principal bundle $L(M^4)$ with structural group L^+_-, the proper orthochronous Lorentz group (see [72]). The action of L^+_- on Minkowski space preserves the standard and time orientation. The bundle $L(M^4)$ is a subbundle of the principal bundle of Lorentz-orthonormal frames. Denote by H the restriction of the Levi-Civitá connection to $L(M^4)$ and by V the vertical distribution on $L(M^4)$. As on OM in Section 2.7, the bundles H and V over $L(M^4)$ are trivial. In particular H is trivialized by the basic vector fields on $L(M^4)$.

The generalization of Itô processes to the Lorentz manifold M^4 requires some modification with respect to the case of Minkowski space. Choose a point $m_0 \in M^4$ and a Lorentz orthonormal frame b in $T_{m_0}M^4$. Introduce a Euclidean structure in $T_{m_0}M^4$ by setting b to be orthonormal in the Euclidean sense. We may now consider a Wiener process $w(\tau)$ in $T_{m_0}M^4$ as well as Itô processes with this $w(\tau)$. One can easily show that the entire construction of Itô developments on manifolds can be clearly generalized to the above case of processes on the Lorentz manifold M^4 by using connections on $L(M^4)$ instead of on OM. Those developments will be called *Itô processes on the Lorentz manifold* M^4.

The above parameter τ will play the role of proper time. So for a process $\xi(\tau)$ we may expect that (15.54) is fulfilled.

In order to avoid any possible confusion we assume in this section that M^4 is stochastically complete. This means that the development of a Wiener process in the above-mentioned sense exists for $\tau \in [0,\infty)$ (see Section 7.6.2). Unfortunately nothing like the criterion of stochastic completeness of Theorem 7.80 is known for Lorentz manifolds.

Itô processes on M^4, which are the developments of Itô processes of diffusion type of the form $z = \int_0^\tau a(s)\mathrm{d}s + \sigma w(\tau)$ with $\sigma = \sqrt{\frac{\hbar}{\mathrm{m}}}$, are of particular interest to us. Let $\xi(\tau)$ be such a process in M^4. In a chart (U,φ) we can apply formulae (15.36) and (15.37) to define relativistic forward and backward, respectively, mean derivatives $D_+\xi(\tau)$ and $D_-\xi(\tau)$ and represent them as compositions of ${}^{\varphi}Y^0_+(\tau,m)$ and ${}^{\varphi}Y^0_-(\tau,m)$ (determined in the chart (U,φ)) with $\xi(\tau)$ (cf. Section 9.1). As in Section 9.1 we determine $X^0_+(\tau,m)$ and $X^0_-(\tau,m)$ by formulae $X^0_+(\tau,m) = Y^0(\tau,m)_n$ and $X^0_-(\tau,m) = Y^0_*(\tau,m)_n$, respectively, for any $m \in M^4$ where n denotes the calculations in the normal chart at m. Then we define the relativistic forward and backward mean derivatives of ξ on M^4 with respect to H by the formulae

$$D_+\xi(\tau) = X^0_+(\tau,\xi(\tau)), \qquad D_-\xi(\tau) = X^0_-(\tau,\xi(\tau)) \tag{15.55}$$

(cf. Definition 9.1).

The derivative $\bar{D}_S = \frac{1}{2}(D_+ + D_-)$ is called the *relativistic symmetric mean derivative*. Consider the vector $\bar{v}^{\xi}(\tau,m) = \frac{1}{2}(X^0_+(\tau,m) + X^0_-(\tau,m))$; $\bar{v}^{\xi}(\tau,\xi(\tau)) = \bar{D}_S\xi(\tau)$ is called the *relativistic current* 4*-velocity* of the process $\xi(\tau)$; $\bar{D}_A = \frac{1}{2}(D_+ - D_-)$ is called the *relativistic antisymmetric mean derivative*. Consider for $\xi(\tau)$ as above the vector $\bar{u}^{\xi}(\tau,m) = \frac{1}{2}(X^0_+(\tau,m) - X^0_-(\tau,m))$; $\bar{u}^{\xi}(\tau,\xi(\tau)) = \bar{D}_A\xi(\tau)$ is called the *relativistic osmotic velocity* of $\xi(\tau)$.

Let $Y(\tau,m)$ be a C^2-smooth vector field on M^4. Making the same modification for formulae (9.15) as above in this section, we can define the covariant relativistic mean derivatives $\mathbf{D}_+Y(\tau,\xi(\tau))$ and $\mathbf{D}_-Y(\tau,\xi(\tau))$. Here we use parallel translation with respect to the Levi-Civitá connection on $L(M^4)$.

We can also apply the Levi-Civitá connection on $L(M^4)$ (i.e., its normal charts on M^4, parallel translation of vectors in TM^4, etc.) to define $D\xi(\tau)$, $D_*\xi(\tau)$, $D^{\xi}\eta(\tau)$, $D^{\xi}_*\eta(\tau)$, $\mathbf{D}Y(\tau,\xi(\tau))$ and $\mathbf{D}_*Y(\tau,\xi(\tau))$ as in Chapters 8 and 9. In spite of the fact that these objects differ from those in the above sections, since the connections are different, we use the old notation without any ambiguity.

Let us pick a Lorentz orthonormal frame in T_mM^4 at some m and represent a vector in terms of its coordinates with respect to the frame $X = (X^0, \bar{X})$ where X^0 denotes the 'time-like' component and $\bar{X}$ the 3-dimensional 'space-like' component. Note that this coordinate decomposition is covariant with respect to Lorentz transformations in T_mM^4 for $D_+\xi(\tau)$ and $D_-\xi(\tau)$ and not covariant for $D\xi(\tau)$ and $D_*\xi(\tau)$. Nevertheless $v^{\xi}(\tau)$ is covariant, since one can can easily see that $v^{\xi}(\tau) = \bar{v}^{\xi}(\tau)$. Indeed, direct calculations show that both $v^{\xi}(\tau)$ and $\bar{v}^{\xi}(\tau)$ have the same coordinate decomposition: $(D^{\xi}_S\xi^0(\tau), D^{\xi}_S\bar{\xi}(\tau))$ (cf. Section 15.3.1).

As in Section 15.3.1, for the osmotic velocity $\bar{u}^{\xi}(\tau)$ we have

$$\bar{u}^{\xi}(\tau) = \left(D^{\xi}_A\xi^0(\tau), -D^{\xi}_A\bar{\xi}(\tau)\right)$$

while $u^{\xi}(\tau) = (D^{\xi}_A\xi^0(\tau), D^{\xi}_A\bar{\xi}(\tau))$. The following formula holds:

$$\bar{u}^{\xi}(\tau,x) = \frac{1}{2}\sigma^2 \,\mathrm{Grad}\log\rho^{\xi}(\tau,x), \tag{15.56}$$

where Grad is the gradient calculated with respect to the Lorentz metric.

Definition 15.47. The vector $\frac{1}{2}(\mathbf{D}_+D_- + \mathbf{D}_-D_+)\xi(\tau)$ is called the 4-*acceleration* of the process ξ at τ.

Theorem 15.48 *For an Itô process $\xi(\tau)$ in M^4 as above the following equality holds*

$$\frac{1}{2}(\mathbf{D}D_* + \mathbf{D}_*D)\xi(\tau) = \frac{1}{2}(\mathbf{D}_+D_- + \mathbf{D}_-D_+)\xi(\tau). \tag{15.57}$$

As for formula (15.46), this statement is proved by direct calculation. Indeed,

$$\mathbf{D}_+D_-\xi(\tau) = (\mathbf{D}^{\xi}_*D^{\xi}\xi^0(\tau), \mathbf{D}^{\xi}D^{\xi}_*\bar{\xi}(\tau))$$

and also

$$\mathbf{D}_-D_+\xi(\tau) = (\mathbf{D}^{\xi}D^{\xi}_*\xi^0(\tau), \mathbf{D}^{\xi}_*D^{\xi}\bar{\xi}(\tau)),$$

which yields (15.57).

Here we consider the relativistic Newton-Nelson equation of the form

$$\begin{cases} \frac{1}{2}(\mathbf{D}_+D_- + \mathbf{D}_-D_+)\xi(\tau) = \bar{\alpha}(\xi(\tau), \bar{v}^{\xi}(\tau,\xi(\tau))), \\ D_2\xi(\tau) = \frac{\hbar}{\mathrm{m}}I, \end{cases} \tag{15.58}$$

where $\bar{\alpha}$ is a 4-force, i.e., it does not depend on τ and necessarily depends on 4-velocity. This is why we consider below the force field on M^4 in the form $\bar{\alpha}(m,X) = \bar{\alpha}_1(m)X$, where $\bar{\alpha}_1(x)$ is a linear operator in T_mM^4, as above.

There is no $\widehat{\mathrm{Ric}}$ term in (15.58) (for the same reasons that such a term appeared in (15.29); see Section 15.2.1). In the relativistic case a natural relation between the Newton-Nelson equation and the Klein-Gordon equation is established (for Minkowski space this was illustrated in Section 15.3.1, see also, e.g., [55, 142, 189, 190, 228, 237, 238]). But in the Klein-Gordon equation (unlike the Schrödinger equation on a Riemannian manifold) the wave operator $\Box = \nabla^* \cdot \nabla$ is involved, where ∇ is the covariant derivative of the Levi-Civitá connection on M^4 (see, e.g., [49]). Since we need not obtain $\Delta = \mathrm{d}\delta + \delta\mathrm{d}$ in the corresponding equation of Schrödinger type, here we do not include the $\widehat{\mathrm{Ric}}$ term in (15.58) (cf. Section 15.2.1).

Definition 15.49. An Itô process $\xi(\tau)$ on M^4 of the type mentioned above is called a *trajectory* of a relativistic stochastic-mechanical system with force field $\bar{\alpha}(m,X)$ if it satisfies (15.58).

Let M^4 be stochastically complete. Specify a point $m_0 \in M^4$ and consider the Itô development $R_IW(\tau)$ of a Wiener process $W(\tau)$ in $T_{m_0}M^4$ as described above in this section. Choose a Lorentz-orthonormal frame b in $T_{m_0}M^4$.

Let the force field $\bar{\alpha}(m,X) = \bar{\alpha}_1(m)X$ on M^4 be such that the tensor $\bar{\alpha}_1(m)$ and its covariant derivative have bounded absolute values of components with respect to the frames, parallel to b along $R_I W(\tau)$. The simplest example of such an $\bar{\alpha}$ is $\bar{\alpha} = 0$, it holds when no external forces but gravitation alone are under consideration.

Make the parallel translation of $\bar{\alpha}_1(R_I W(\tau))$ and $\mathrm{tr}\nabla\bar{\alpha}_1(\bar{\alpha}_1)$ along $R_I(W(\cdot))$ to m_0. Introduce the Euclidean inner product in $T_{m_0}M^4$ by considering b as an orthonormal frame in the Euclidean sense. Fix some $t_0 \in (0,l)$. Now, having omitted all summands with a $\widehat{\mathrm{Ric}}$ term in equations (15.31), (15.33), etc., one can apply the construction of Section 15.2.2 and easily obtain the existence of the corresponding $\zeta(\tau)$ and its development $\xi(\tau)$.

Theorem 15.50 *For $\tau \in (t_0,l)$ the process $\xi(\tau)$ satisfies* (15.58).

This is a simple corollary of (15.57) and the equality $v^\xi(\tau) = \bar{v}^\xi(\tau)$.

Remark 15.51. It should be pointed out that $\xi(\tau)$ does not satisfy (15.58) for $\tau \in (0,t_0)$ since here $t_0(\tau) = \frac{1}{t_0}$ and $E^\xi_\tau(\Gamma_{\tau,0}\theta^{-1}\frac{\zeta(\tau)}{\tau\cdot t_0}) \neq E^\xi_\tau(\Gamma_{\tau,0}\theta^{-1}\frac{\zeta(\tau)}{\tau^2})$. Thus $\xi(\tau)$ can be interpreted only as a trajectory of a stochastic mechanical system beginning at the instant t_0 of proper time from a random configuration $\xi(t_0)$ with initial mean forward derivative $E^\xi_\tau(\Gamma_{\tau,0}a(t_0))$. It is clear that t_0 may be chosen arbitrarily close to zero, and so we can bring the initial values to m_0 and a_0 as closely as desired. However, we cannot set $t_0 = 0$, since the integral $\int_0^\tau \frac{1}{s}\mathrm{d}W(s)$ does not exist (indeed, $\int_0^\tau \frac{1}{s^2}\mathrm{d}s$ diverges, see, e.g., [162]), i.e., when $t_0 = 0$ the analogs of equations (15.31) and (15.33) are ill-posed. We suggest the hypothesis that this situation may be thought of as a description of the 'big bang', the initial point for all the trajectories. Indeed, all the equations of physical laws are ill-posed at the 'big bang'; we can set initial conditions for any time greater than the instant of the 'big bang', but cannot do this for the 'big bang' itself.

The following analog of Theorem 15.31 holds.

Theorem 15.52 *Let the Levi-Civitá connection on $L(M^4)$ be stochastically complete and the $(1,1)$-tensor field $\alpha(m)$ be continuous and have compact support. Then for η as in Theorem 15.31 and for a continuous vector field $a(m)$ with compact support there exists a solution $\xi(\tau)$ of equation (15.58) with initial conditions $\xi(0) = \eta$ and $\bar{D}_S\xi(0) = a(\eta)$.*

The analogous statement with initial condition for the forward mean derivative $D\xi(0) = a(\eta)$ is also true.

Taking into account formula (15.57) and the fact that $\bar{v}^\xi(\tau) = v^\xi(\tau)$ (see above) it is easy to see that Theorem 15.52 is a simple generalization of Theorem 15.31 (here we use another bundle and another connection). Note that the hypothesis of Theorem 15.52 has been expressed in a form that is easy to verify. It can be replaced by the corresponding analogs of Conditions 15.20 and 15.21.

Chapter 16
Hydrodynamics

16.1 The Lagrangian Formalism of the Hydrodynamics of an Ideal Barotropic Fluid

Following the lines of Chapter 11, we can apply the results of Section 5.1 to study mechanical systems on the configuration spaces $\mathcal{D}^s(M)$, $H^s(M,M)$, or $H^s(M,N)$ with kinetic energy given by the (weak) Riemannian metric. Here we analyze those systems which are naturally related to certain problems of hydrodynamics. Note that according to the Lagrangian formalism, a trajectory of such a system gives the flow of a fluid.

16.1.1 Diffuse matter

In what follows, we use the notation and the hypotheses of Section 5.1. In particular, M denotes a compact Riemannian manifold without boundary and $\langle\cdot,\cdot\rangle$ is its Riemannian metric.

Consider a mechanical system on $H^s(M,M)$ with zero potential energy and kinetic energy $\mathcal{K}(X) = \frac{(X,X)}{2}$, where $(\cdot,\cdot)$ is given by (5.1). Then Newton's equation for the system is

$$\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\,\dot{g}(t) = 0, \tag{16.1}$$

where $\frac{\bar{\mathrm{D}}}{\mathrm{d}t}$ is defined as in Section 5.1.

Definition 16.1. The mechanical system defined above by (16.1) is called a *Lagrangian hydrodynamical system* (*LHS*) of diffuse matter (with zero external force).

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9_16,

It is clear by (16.1) that the trajectory of every particle of diffuse matter is a geodesic of $\langle\cdot,\cdot\rangle$ on M. In other words, the trajectory of the LHS with initial condition $e = \mathrm{id}$ is given by the vector field $X \in T_e\mathcal{D}^s(M)$ of the initial velocities and the metric $\langle\cdot,\cdot\rangle$.

The kinetic energy is constant along a trajectory of the LHS on $H^s(M, M)$ and the trajectory is an extremal of the action functional with the Lagrangian $L = \mathcal{K}$.

Similarly, one may define an LHS of diffuse matter on a manifold M with boundary. This time, however, the motion takes place on a larger manifold N without boundary ($\dim N = \dim M$). The construction still holds if we take $H^s(M, N)$ as the configuration space. One may also consider an LHS of diffuse matter with an external force.

Note that by Proposition 3.67 the curve $(g(t), \dot{g}(t))$ in $T\mathcal{D}^s(M)$ is an integral curve of the geodesic spray $\bar{\mathcal{Z}}$ (see (5.4)), i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}t}(g(t), \dot{g}(t)) = \mathcal{Z}(g(t), \dot{g}(t)). \tag{16.2}$$

Proposition 16.2 *For every $X \in T_e\mathcal{D}^s(M)$ there exists a unique solution $g(t)$ of (16.1) with initial conditions $g(0) = e$ and $\frac{\mathrm{d}}{\mathrm{d}t}g(0) = X$ that is well-defined for t in a sufficiently small interval $[0, \varepsilon)$.*

Indeed, (16.1) is equivalent to (16.2) (which has a smooth right-hand side).

Recall that the group $\mathcal{D}^s(M)$ is an open neighborhood of e in $H^s(M, M)$ and so a solution of (16.1) that starts at e, for t in a sufficiently small interval $[0, \varepsilon)$, belongs to $\mathcal{D}^s(M)$. A key role in the Euler description of diffuse matter is played by the vector $v(t) \in T_e\mathcal{D}^s(M)$ obtained by right translation of the velocity $\dot{g}(t) \in T_{g(t)}\mathcal{D}^s(M)$ to the tangent space at the unit, i.e., $v(t) = TR^{-1}_{g(t)}\dot{g}(t)$ (see Section 11.2). Specify some t and consider the right-invariant vector field $\bar{v}(t)$ on $\mathcal{D}^s(M)$ constructed by right translations of $v(t)$ to all points of $\mathcal{D}^s(M)$. Note that the derivative $\frac{\partial}{\partial t}v(t)$ is a vertical vector in $T_{(e,v(t))}T\mathcal{D}^s(M)$ and from (16.2) it follows that $\frac{\partial}{\partial t}v(t)$ equals the difference between $\bar{\mathcal{Z}}$ and its component "tangent" to the submanifold $\bar{v}(t)$ at the given t with respect to the decomposition $T_{(e,v(t))}T\mathcal{D}^s(M) = \mathsf{V}_{(e,v(t))} \oplus T_{(e,v(t))}\bar{v}(t)$. In fact, if $v(t)$ is an H^s-vector field on M, $\bar{v}(t)$ is only continuous (see Section 5.1), but if it is H^{s+k} for some integer $k > 0$, then $\bar{v}(t)$ is C^k-smooth and so it really has the tangent space at $(e, v(t))$.

We introduce the notation $\bar{g}(t + \Delta t) = R^{-1}_{g(t)}g(t + \Delta t)$. From the construction of the geodesic spray and of $\bar{v}(t)$ it follows that

$$\bar{\mathcal{Z}}(e, v(t)) = \lim_{\Delta t \to +0} \frac{\bar{v}(t, \bar{g}(t + \Delta t)) - v(t)}{\Delta t} - \bar{K} \lim_{\Delta t \to +0} \frac{\bar{v}(t, \bar{g}(t + \Delta t)) - v(t)}{\Delta t}$$

where the first summand on the right-hand side is tangent to $\bar{v}(t)$. Then $\frac{\partial}{\partial t}v(t)$, as the difference between $\bar{\mathcal{Z}}$ and the latter summand takes the form

$$\frac{\partial}{\partial t}v(t) = -\bar{K} \lim_{\Delta t \to +0} \frac{\bar{v}(t, \bar{g}(t+\Delta t)) - v(t)}{\Delta t}.$$

Since $\bar{K} \lim\limits_{\Delta t \to +0} \frac{\bar{v}(t,\bar{g}(t+\Delta t))-v(t)}{\Delta t} = \nabla_{v(t)}v(t)$, taking into account that the vector $v(t) \in T_e\mathcal{D}^s(M)$ is a vector field $v(t,m)$ on M and that the connector $\bar{K}$ is defined via the connector K of the Levi-Civitá connection on M by formula (5.2), we obtain that $v(t)$ satisfies the equation

$$\frac{\partial}{\partial t}v + \nabla_v v = 0 \tag{16.3}$$

on M. This is the so-called *Hopf equation* describing the motion of diffuse matter in the framework of Euler's approach. Sometimes it is also called the *Burgers equation without viscous term.* Note that on the flat torus (16.4) takes the form

$$\frac{\partial}{\partial t}v + (v \cdot \nabla)v = 0. \tag{16.4}$$

It should be pointed out that, unlike (16.1), equations (16.3) and (16.4) lose derivatives, as is typical in Euler's approach.

The motion of diffuse matter is not of much interest for hydrodynamics and we shall use it only as a starting point for our further analysis.

16.1.2 A barotropic fluid

Let us now turn to Lagrangian hydrodynamical systems of an ideal barotropic fluid. The major difference between such an LHS and that described in Section 16.1.1 is the presence of a force field, the potential of which is called the *internal energy* of the fluid. Strictly speaking, an LHS of an ideal barotropic fluid is not covered by the general construction of Chapter 11 because the force field loses smoothness, i.e., it is an H^{s-1}-smooth "vector field" on M. If we considered only C^∞-diffeomorphisms, then the force field would be C^∞-smooth. However, then the configuration space $\mathcal{D}^\infty(M)$ would be modeled on a locally convex space, rather than on a Banach space. As a consequence, the whole analysis would become much more complicated.

Here we briefly outline some definitions and results on LHSs of a barotropic fluid on a closed manifold. A more detailed account can be found in [59, 210, 211].

Let M be a compact Riemannian manifold without boundary and let $\mathcal{D}^s(M)$ be the group of H^s-diffeomorphisms of M with $s > n/2 + 2$. Denote by $\mathcal{V}^{s-1}$ the space of H^{s-1}-smooth volume forms ν on M such that

$$\int_M \nu = \int_M \mu,$$

where μ is the Riemannian volume form. Following Smolentsev [210], consider the map $\Psi_\mu\colon D^s(M) \to \mathcal{V}^{s-1}$ defined as $\Psi_\mu(g) = (g^{-1})^*\mu$, where $(g^{-1})^*\mu$ is the pull-back of μ under the map g^{-1}. Since any two volume forms differ by a multiplier (see Section 1.6) which is a scalar function on M, we have $\Psi_\mu(g) = \rho(g)\mu$, where $\rho(g)\colon M \to (0, +\infty)$ is an H^{s-1} function called the density of the fluid at $g \in \mathcal{D}^s(M)$.

Let $\mathcal{U}_1\colon (0,\infty) \to (0,\infty)$ be a smooth function. The composition $\mathcal{U}_1(\rho)\colon M \to (0,\infty)$ is called the *specific internal energy* of the system. Then the internal energy $\mathcal{U}\colon \mathcal{D}^s(M) \to (0,\infty)$ is defined as

$$\mathcal{U}(g) = \int_M \mathcal{U}_1(\rho)\,\rho\,\mu = \int_M \mathcal{U}_1(\rho)\,\nu,$$

where $\nu = \rho\mu = \Psi_\mu(g)$. In a true physical system, the function $\mathcal{U}_1$ depends on the properties of the fluid.

Consider also the function $p\colon M \to \mathbb{R}$ given by $p(\rho) = \rho^2\,\frac{d\mathcal{U}_1}{d\rho}$, which is known as the *state equation* in mechanics. The function p is called the *pressure* of the fluid at g, where $\Psi_\mu(g) = \rho\mu$.

Remark 16.3. To explain the terminology, we emphasize that the fluid under consideration is compressible, since we have taken the entire group $\mathcal{D}^s(M)$ as the configuration space. In mechanics a compressible fluid is said to be *barotropic* if the pressure depends only on the density.

The gradient of $\mathcal{U}$ with respect to $(\cdot,\cdot)$ on $\mathcal{D}^s(M)$ might not exist because $(\cdot,\cdot)$ is just a weak Riemannian metric. However, as the following theorem shows, the gradient exists in the class of H^{s-1}-vector fields on M.

Theorem 16.4 *Let F be the vector field on $\mathcal{D}^s(M)$ defined by the equation*

$$F_g = TR_g\left(\frac{1}{\rho}\operatorname{grad} p(\rho)\right),$$

where $\Psi_\mu(g) = \rho\mu$. Then for any $Y \in T_gD^s(M)$, we have $\mathrm{d}\mathcal{U}(Y) = (Y, F_g)$, i.e., $F = \operatorname{grad}\mathcal{U}$.

Proof. [219] Let $g(t)$ be the flow of Y on M. Differentiating the equation $\mu = g^*(\rho(t)\mu)$, we see that $\frac{\mathrm{d}}{\mathrm{d}t}\,g^*(\rho(t)\mu) = 0$ or, equivalently, $\frac{\partial\rho}{\partial t} + \operatorname{div}(\rho V) = 0$ where $V = TR^{-1}_{g(t)}\,Y_{g(t)}$. Therefore, we have

$$
\begin{aligned}
\mathrm{d}\mathcal{U}(Y) = {} & \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{U}(g(t))\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\int_M \mathcal{U}_1(\rho)\,\rho\mathrm{d}\mu\Big|_{t=0} \\
&= \int_M \frac{\partial \mathcal{U}_1}{\partial \rho}\,(-\mathrm{div}\,(\rho Y))\;\rho\mathrm{d}\mu + \int_M \mathcal{U}_1(\rho)\,(-\mathrm{div}\,(\rho Y))\;\rho\mathrm{d}\mu \\
&= \int_M \left\langle \mathrm{grad}\,\left(\rho\frac{\mathrm{d}\mathcal{U}_1}{\mathrm{d}\rho}\right), Y\right\rangle \rho\mathrm{d}\mu + \int_M \langle \mathrm{grad}\,\mathcal{U}_1, Y\rangle\,\rho\mathrm{d}\mu \\
&= \int_M \left\langle \mathrm{grad}\,\frac{p}{\rho}, Y\right\rangle \rho\mathrm{d}\mu + \int_M \frac{\mathrm{d}\mathcal{U}_1}{\mathrm{d}\rho}\langle \mathrm{grad}\,\rho, Y\rangle\rho\mathrm{d}\mu \\
&= \int_M \langle \mathrm{grad}\,p, Y\rangle\,\rho\mathrm{d}\mu - \int_M \frac{p}{\rho^2}\langle \mathrm{grad}\,p, Y\rangle\,\rho\mathrm{d}\mu \\
&\quad + \int_M \frac{p}{\rho^2}\langle \mathrm{grad}\,p, Y\rangle\,\rho\mathrm{d}\mu = \int_M \left\langle \frac{\mathrm{grad}\,p}{\rho}, V\right\rangle \rho\mathrm{d}\mu \\
&= \left(TR_g\left(\frac{\mathrm{grad}\,p(\rho)}{\rho}\right), TR_g V\right)_g = (F, Y)_g. \quad \Box
\end{aligned}
$$

Definition 16.5. An *LHS of an ideal barotropic fluid without external force* is the mechanical system on $\mathcal{D}^s(M)$ with kinetic energy $\mathcal{K}(X) = \frac{(X,X)}{2}$ given by (5.1) and potential energy $\mathcal{U}$.

The force field in such an LHS is $-\mathrm{grad}\,\mathcal{U}$, so that Newton's equation takes the form

$$\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\,\dot{g}(t) = -\mathrm{grad}\,\mathcal{U}. \tag{16.5}$$

Using standard properties of the Levi-Civitá connection, it is not hard to show that the total energy $E = \mathcal{K} + \mathcal{U}\circ\pi$ is constant along a trajectory of the LHS and that every trajectory is an extremal with fixed end-points of the action functional with the Lagrangian $L = \mathcal{K} - \mathcal{U}\circ\pi$.

Let $\varPhi$ be a vector field on M and $\bar{\varPhi}$ the induced right-invariant vector field on $\mathcal{D}^s(M)$.

Definition 16.6. An *LHS of an ideal barotropic fluid with the external force* $\varPhi$ is the mechanical system on $\mathcal{D}^s(M)$ with $\mathcal{K}$ as in Definition 16.5 and the total force field $-\mathrm{grad}\,\mathcal{U} + \bar{\varPhi}$. Newton's equation for this system is

$$\frac{\bar{\mathrm{D}}}{\mathrm{d}t}\,\dot{g}(t) = -\mathrm{grad}\,\mathcal{U} + \bar{\varPhi}. \tag{16.6}$$

Let us now show how to pass to Euler's equation for a barotropic ideal fluid. Let $g(t)$ be a trajectory of the LHS (16.6). Consider the curve $u(t) = TR^{-1}_{g(t)}\dot{g}(t)$ in $T_e\mathcal{D}^s(M)$. According to the definitions given in Section 11.2, the curve $u(t)$ satisfies Euler's equation, which, in this particular case, coincides with Euler's equation of an ideal barotropic fluid. To see this, we observe that by the definition of $\bar{\varPhi}$ and Theorem 16.4, the right-hand side of (16.6) turns into $-\frac{1}{\rho}\mathrm{grad}\,p + \varPhi$ after being right-translated to $T_e\mathcal{D}^s(M)$. Then, applying the same arguments as in Section 16.1.1, we obtain that

$\frac{\mathrm{d}}{\mathrm{d}t}u(t,m) = -\nabla_{u(t,m)}u(t,m) - \frac{1}{\rho}\mathrm{grad}\, p + \Phi$ where $u(t,m)$ is $u(t)$ considered as a vector field on M. On the other hand, similarly to the proof of Theorem 16.4, we can deduce that $\frac{\partial \rho}{\partial t} + \mathrm{div}(\rho u) = 0$. As a result, we obtain the system of *Euler equations*

$$\begin{cases} \frac{\partial u}{\partial t} + \nabla_u u + \frac{1}{\rho}\mathrm{grad}\, p = \Phi \\ \qquad\quad \frac{\partial \rho}{\partial t} + \mathrm{div}\,(\rho u) = 0. \end{cases} \tag{16.7}$$

Remark 16.7. We finish this section by giving some general references on the material presented here. The manifold of C^∞-diffeomorphisms was studied in [210, 211]. In this case the principle of least action in Maupertuis' form has been proved and the integrals of motion have been analyzed. The manifold $\mathcal{D}^s(M)$ was studied in [59]. Note that, in general, it is harder to work with $\mathcal{D}^s(M)$ than with $\mathcal{D}^\infty(M)$, yet the existence theorems have been proved only for $\mathcal{D}^s(M)$. In [59], an LHS of an ideal barotropic fluid was regarded as a system with a strong constraint force given by a potential having a minimum on the manifold $\mathcal{D}^s_\mu(M)$ of volume preserving diffeomorphisms. The latter group is the configuration space for an ideal incompressible fluid. It has been shown that a trajectory on $\mathcal{D}^s(M)$ approaches the submanifold $\mathcal{D}^s_\mu(M)$ as the parameters of the system go to infinity, so that the fluid becomes incompressible.

16.2 Lagrangian Hydrodynamical Systems of an Ideal Incompressible Fluid

The LHS of an ideal incompressible fluid is described as the LHS of diffuse matter subjected to a constraint in the sense of Section 11.6. This constraint is described as follows. Both in the case of a finite-dimensional manifold M without boundary and with boundary introduce the notation $\beta = T_e\mathcal{D}^s_\mu(M) \subset T_e\mathcal{D}^s(M)$ and translate this subspace to tangent spaces at all points of $\mathcal{D}^s(M)$. As a result we obtain the distribution $\bar{\beta}$, a sub-bundle of $T\mathcal{D}^s(M)$ that plays the role of a constraint. This constraint is holonomic, its integral manifold passing through e is $\mathcal{D}^s_\mu(M)$ and so we shall generally restrict the constraint equations to $T_e\mathcal{D}^s(M)$. However, in our considerations below it is sometimes more convenient to deal with $\bar{\beta}$ as a constraint rather than restricting ourselves to $\mathcal{D}^s_\mu(M)$.

Note that in this case the covariant derivatives $\widetilde{\nabla} = \bar{P} \circ \bar{\nabla}$ and $\frac{\widetilde{\mathrm{D}}}{\mathrm{d}t} = \bar{P} \circ \frac{\bar{\mathrm{D}}}{\mathrm{d}t}$ introduced by formulae (5.8), where the operator $\bar{P}$ is defined by formula (5.6), coincide with the reduced covariant derivatives by means of Section 11.1.

Let F be an H^s-smooth vector force field on M, i.e., a vector in $T_e\mathcal{D}^s(M)$. Construct the right-invariant field $\bar{F}$ on $\mathcal{D}^s(M)$ ($\bar{F}_e = F$) and consider the Newton law on $\mathcal{D}^s(M)$ of the obtained system with constraint in the form

$$\frac{\tilde{\mathrm{D}}}{\mathrm{d}t}\dot{g}(t) = \bar{P} \circ \bar{F}. \tag{16.8}$$

Remark 16.8. $\bar{P} \circ \bar{F}$ is a right-invariant vector field on $\mathcal{D}^s(M)$ such that $(\bar{P} \circ \bar{F})_e = P \circ F$ is the divergence-free (and tangent to the boundary if M is with boundary) component of F in the Hodge decomposition (5.7) ((5.10), respectively). For this reason, without loss of generality, we may consider F as divergence-free (and tangent to the boundary, respectively), i.e., as belonging to $T_e\mathcal{D}^s_\mu(M)$.

Since, as mentioned above, $\bar{\beta}$ is integrable, following the general theory we can restrict the system to the integral manifold $\mathcal{D}^s_\mu(M)$ that yields a mechanical system without constraint on that integral manifold.

Definition 16.9. A *Lagrangian hydrodynamical system of an ideal incompressible fluid without external forces* is a mechanical system with configuration space $\mathcal{D}^s_\mu(M)$ and kinetic energy $\mathcal{K}(X) = \frac{1}{2}(X, X)$, $X \in T\mathcal{D}^s_\mu(M)$ where $(\cdot,\cdot)$ is a weak Riemannian metric (5.1), and so Newton's law takes the form

$$\frac{\tilde{\mathrm{D}}}{\mathrm{d}t}\dot{g}(t) = 0, \tag{16.9}$$

where $\frac{\tilde{\mathrm{D}}}{\mathrm{d}t}$ is as defined in (5.8).

Note that $\frac{\tilde{\mathrm{D}}}{\mathrm{d}t}$ is the covariant derivative of the Levi-Civitá connection of the metric $(\cdot,\cdot)$ on $\mathcal{D}^s_\mu(M)$, as follows from general results of Riemannian geometry. Thus the trajectories of the LHS (16.9) are the geodesics of that metric (i.e., in particular, they are extremals with fixed ends for the action functional with Lagrangian $L = \mathcal{K}$).

Let $F \in T_e\mathcal{D}^s_\mu(M)$ and $\bar{F}$ be the corresponding right-invariant vector field on $\mathcal{D}^s_\mu(M)$.

Definition 16.10. A *Lagrangian hydrodynamical system of ideal incompressible fluid with external force* F is a mechanical system with configuration space and kinetic energy as in Definition 16.9, with external force $\bar{F}$ and hence with Newton's law in the form

$$\frac{\tilde{\mathrm{D}}}{\mathrm{d}t}\dot{g}(t) = \bar{F}. \tag{16.10}$$

Taking into account Remark 16.8, one can easily see that we do not lose generality by choosing the external forces from $T_e\mathcal{D}^s_\mu(M)$.

Recall that the geodesics of the Levi-Civitá connection of the metric $(\cdot,\cdot)$ on $\mathcal{D}^s_\mu(M)$ are described by the geodesic spray $\mathcal{S}$ (see Theorem 5.9).

Theorem 16.11 *For every $X \in T_e\mathcal{D}^s_\mu(M)$ there exists a unique solution $g(t)$ of equation (16.9) with $\dot{g}(0) = X$ that is well-defined on some sufficiently small interval $t \in [0, \varepsilon)$.*

Indeed, the solutions of (16.9), and only these solutions, are described as curves $\pi\gamma(t)$ where $\pi : T\mathcal{D}^s_\mu(M) \to \mathcal{D}^s_\mu(M)$ is the natural projection and $\gamma(t)$ is an integral curve of the geodesic spray $\mathcal{S}$ on $T\mathcal{D}^s_\mu(M)$. Such a curve $\gamma(t)$ locally exists and is unique for every initial condition $X \in T\mathcal{D}^s_\mu(M)$ (in particular, $X \in T_e\mathcal{D}^s_\mu(M)$) since $\mathcal{S}$ is smooth.

Remark 16.12. Since $\mathcal{S}$ is right-invariant, the choice of initial condition $X \in T_e\mathcal{D}^s_\mu(M)$) does not restrict generality. Indeed, by a right shift the initial point of a trajectory can be translated to e.

Theorem 16.13 *Let F be a divergence-free (and tangent to the boundary if M is with boundary) H^l-vector field on M where $l > \frac{n}{2}+2$. If $l \geq s > \frac{n}{2}+1$, for every $X \in T_e\mathcal{D}^s_\mu(M)$ the solution $g(t)$ of (16.10) with $g(0) = e$ and $\dot{g}(0) = X$, exists for $t > 0$ small enough, and is unique.*

Proof. F generates on $T\mathcal{D}^s_\mu(M)$ the second order differential equation (special vector field) $\mathcal{S} + (\bar{F})^l$ (where $(\bar{F})^l$ is the vertical lift of $\bar{F}$). Its integral curves are sent by the mapping π to the solutions of (16.10). The field $\mathcal{S} + \bar{F}^l$ on $T\mathcal{D}^s_\mu(M)$ is at least C^1-smooth (see Theorem 5.4) which yields the local existence and uniqueness of a solution. If $s = l > \frac{n}{2} + 2$, the existence and uniqueness of a solution on $T\mathcal{D}^s_\mu(M)$ follow from its existence and uniqueness on $T\mathcal{D}^{s-1}_\mu(M)$ and from the smooth dependence of the flows on TM on the initial data (cf. Theorem 5.5). □

Note the following hydrodynamical interpretation of Theorem 5.10: the flow of an ideal incompressible fluid without external forces can realize every configuration of the fluid that is sufficiently close to the initial configuration.

Now we can derive the Euler equation in “space coordinates” (see Section 11.2) corresponding to the LHS (16.9) or (16.10). Let $g(t)$ be a solution of (16.10) and consider $u(t) = \mathbb{R}_g^{-1} \circ \dot{g}(t) \in T_e\mathcal{D}^s_\mu(M)$. As in the derivation of the Euler equation (16.7), using the definition of $\frac{\tilde{\mathrm{D}}}{\mathrm{d}t} = \bar{P}\frac{\bar{\mathrm{D}}}{\mathrm{d}t}$, one can easily show that $u(t)$ in $T_e\mathcal{D}^s_\mu(M)$ satisfies the equation $P(\frac{\partial u}{\partial t} + \nabla_u u) = F$. Since $u \in T_e\mathcal{D}^s_\mu(M)$, we have $P\frac{\partial u}{\partial t} = \frac{\partial u}{\partial t}$. Hence, since $P(\nabla_u u) = \nabla_u u + \operatorname{grad} p$ (see (5.7) and (5.10)), we obtain the classical Euler equation of the motion of an ideal incompressible fluid:

$$\frac{\partial u}{\partial t} + \nabla_u u + \operatorname{grad} p = F. \tag{16.11}$$

Remark 16.14. Usually one finds the coefficient $\frac{1}{\rho}$ before $\operatorname{grad} p$ in the Euler equation, where ρ is the fluid density. Since the fluid is incompressible, its density is constant and we can conveniently adopt a system of units in which $\rho = 1$.

Remark 16.15. The condition $\operatorname{div} u = 0$ (and the condition that u is tangent to the boundary if M is with boundary) is also incorporated into the Euler equation. We emphasize that this condition is equivalent to $u \in T_e\mathcal{D}^s_\mu(M)$.

Remark 16.16. When we prove the existence and uniqueness of solutions in Theorems 16.11 and 16.13 in the framework of Lagrangian formalism, we use the fact that the field $\mathcal{S}$ on $T\mathcal{D}^s_\mu(M)$ is smooth and right-invariant. When we pass to the Euler description, i.e., to the Euler equation (16.11) in $T_e\mathcal{D}^s_\mu(M)$, we lose information about $\mathcal{S}$. We emphasize that the Euler equation is a partial differential equation that is well-defined only on the everywhere dense subset $T_e\mathcal{D}^{s+1}_\mu(M)$ of $T_e\mathcal{D}^s_\mu(M)$. The proof of the existence of solutions for the Euler equation (16.11), which is independent of Theorems 16.11 and 16.13, is rather complicated.

The Newton Law

$$\frac{\widetilde{\mathrm{D}}}{\mathrm{d}t}\dot{g}(t) = \bar{F}(t, g(t), \dot{g}(t)) \tag{16.12}$$

with non-right-invariant vector force field $\bar{F}(t, g, Y)$, $Y \in T_g\mathcal{D}^s_\mu(\mathcal{T}^n)$ describes fluid motion under a force that depends on the fluid configuration and on velocity. Equation (16.12) can obviously be reduced to the first order equation on $T\mathcal{D}^s_\mu(\mathcal{T}^n)$ with special vector field $\mathcal{S}+\bar{F}^l(t, g, Y)_{|(g,Y)}$. Recall that $\mathcal{S}$ is C^∞-smooth. Hence, if $\bar{F}(t, g, Y)$ is smooth enough, (16.12) has a solution $g(t)$ of the Cauchy problem $g(0) = e$, $\dot{g}(0) = X_0 \in T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$, well-defined on a sufficiently small time interval. Denote by $u(t)$ the curve in $T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$ (i.e., a divergence-free vector field on $\mathcal{T}^n$) obtained by right translations of vectors $\dot{g}(t)$, i.e., $u(t) = \dot{g}(t) \circ g^{-1}(t) = TR^{-1}_{g(t)}\dot{g}(t)$. Applying the above arguments one can easily see that $u(t)$ satisfies the Euler equation

$$\frac{\partial}{\partial t}u + (u \cdot \nabla)u - \mathrm{grad}p = TR^{-1}_{g(t)}\bar{F}(t, g(t), u(t, g(t))). \tag{16.13}$$

16.3 The Regularity Theorem and a Review of Results on the Existence of Solutions

As we have shown, a trajectory of an LHS of an ideal incompressible fluid exists locally on $\mathcal{D}^s_\mu(M)$ provided that $s > n/2+1$. (See Theorem 16.11 and, for a system with external force, Theorem 16.13.) Passing from the LHS to (16.11), we obtain the local existence of solutions of Euler's equation. The solutions belong to H^s with $s > n/2 + 1$, and so are smooth in the standard sense. The existence of solutions on the interval $(-\infty, \infty)$ has so far been proved only on two-dimensional manifolds. For various two-dimensional problems, results of this kind were obtained in [157, 230, 236]. In higher dimensions, proving global existence is an important and still unsolved problem. This difference between hydrodynamics on two- and three-dimensional manifolds has its roots in the fact that the geometric properties of the group of volume preserving diffeomorphisms $\mathcal{D}^s_\mu(M)$ change drastically as we pass from $\dim M = 2$ to higher dimensions. (See [3, Appendix 2] and [178] for a detailed discussion of this matter.)

Consider an LHS of an ideal incompressible fluid without external force on a manifold M. Geometrically, the existence of trajectories on $(-\infty, \infty)$ means that the weak Riemannian manifold $\mathcal{D}^s_\mu(M)$ with the metric given by (5.1) is geodesically complete. Note that here we have no analog of the Hopf-Rinow theorem, i.e., geodesic completeness does not mean that any two points of $\mathcal{D}^s_\mu(M)$ can be connected by a geodesic. Later in this section we discuss the problem of whether or not any element of $\mathcal{D}^s_\mu(M)$ can be obtained as a flow of an incompressible fluid with the initial condition $e \in \mathcal{D}^s_\mu(M)$.

Now let us turn to the regularity theorem, which is a very important result in higher-dimensional hydrodynamics. This theorem claims that on the interval where the flow of the fluid exists, the diffeomorphisms forming the flow are as smooth as the initial conditions. (In two-dimensional hydrodynamics the regularity theorem follows from the existence and uniqueness theorem.)

Let M be a compact orientable Riemannian manifold, possibly with boundary, and let $n = \dim M$. Assume also that $s > n/2 + 1$, $q > 0$, and the external force $f_0(t) \in T_e\mathcal{D}^{s+q}_\mu(M) \subset T_e\mathcal{D}^s_\mu(M)$ is continuous in t in the topology of $T_e\mathcal{D}^{s+q}_\mu(M)$. Select the initial condition $X_0 \in T_e\mathcal{D}^s_\mu(M)$ to be an H^{s+k}-smooth vector field on M with $0 \leq k \leq q$. Denote by $\eta(t) \in \mathcal{D}^s_\mu(M)$ the flow of an ideal fluid on M with external force f_0 such that $\eta(0) = e$ and $\dot\eta(0) = X_0$.

For the case of M with boundary we need the following technical statement. Let N be a compact orientable Riemannian manifold without boundary, $\dim N = \dim M$, such that M is a compact submanifold of N. For example, if M is a given Riemannian manifold, we may set $N = \operatorname{double} M$ (see Section 1.1) and equip N with a Riemannian metric which coincides with the original metric on one of the copies of $M \subset N$. Assume that $s > n/2 + 1$ and denote by $\mathcal{D}^s_\mu(M)$ and $\mathcal{D}^s_\mu(N)$ the groups of volume preserving diffeomorphisms of M and N, respectively. Also, let $j\colon \operatorname{Vect} N \to \operatorname{Vect} M$ be the restriction morphism of vector fields, $\mathcal{S}$ the spray on $T\mathcal{D}^s_\mu(N)$ (Theorem 5.9), and $\pi\colon T\mathcal{D}^s_\mu(N) \to \mathcal{D}^s_\mu(N)$ the natural projection.

Theorem 16.17 [18, 19] *There exists a C^∞-smooth right-invariant subbundle Ξ^s of $T\mathcal{D}^s_\mu(N)$ and a C^∞-smooth fiber-wise right-invariant projection $\bar R\colon T\mathcal{D}^s_\mu(N) \to \Xi^s$ with the following properties:*

(i) *The projection $j\colon \Xi^s_e \to T_e\mathcal{D}^s_\mu(M)$ is an isomorphism. (Here Ξ^s_e is the fiber of Ξ^s over e.)*

(ii) *The distribution Ξ^s is non-holonomic. The fibers of Ξ^s have infinite dimension and infinite codimension in the fibers of $TD^s_\mu(N)$.*

(iii) *Denote by $T\bar R\colon TT\mathcal{D}^s_\mu(N) \to T\Xi^s$ the tangent map of $\bar R$. Let $X(t)$ be the integral curve of $T\bar R \circ \mathcal{S}$ with the initial condition $X(0) = Y \in \Xi^s_e$. Then the curve $\eta(t) = \pi X(t)$ consists of diffeomorphisms which preserve M, and $\eta(t)\mid_M$ is a curve in $\mathcal{D}^s_\mu(M)$. This curve is a trajectory* *of the LHS of an ideal incompressible fluid (without external force) on M with the initial condition $Y_0 = j(Y)$.*

A detailed proof of Theorem 16.17 is given in [106, Section 26].

Note that the family of diffeomorphisms $\eta(t)$ does not correspond to the motion of the fluid on $N\backslash M$, even though $\eta(N \setminus M) = N \setminus M$. We should also point out that a "free" geodesic on $\mathcal{D}^s_\mu(N)$ with initial condition Y is a flow of the fluid on N that mixes M and $N \setminus M$.

Let $f(t) \in \Xi^s_e$ be a time-dependent external force and $\bar{f}(t)$ the right-invariant vector field on $\mathcal{D}^s_\mu(N)$ corresponding to $f(t)$. To obtain the results on the flow of a fluid with external force, we just need to slightly refine (ii). We replace the field $T\bar{R} \circ \mathcal{S}$ by the field $T\bar{R} \circ (\mathcal{S} + \bar{f}(t)^l)$, where $\bar{f}(t)^l$ is the vertical lift of $\bar{f}(t)$ to $T\mathcal{D}^s_\mu(N)$. With this modification in mind, we have:

Corollary 16.18 *The curve $\eta(t)\ |_M$ in $\mathcal{D}^s_\mu(M)$ is a trajectory of the LHS of an ideal incompressible fluid on M with the external force $f_0(t) = j\, f(t)$. Theorem 16.17 (suitably modified) still holds for $\eta(t)$ as above.*

It is clear from (i) that $f(t)$ and Y are entirely given by specifying $f_0(t)$ and Y_0, respectively.

Remark 16.19. It is clear by definition that the vector bundle Ξ^k defined for all $k \geq s$ is the intersection of Ξ^s with $T\mathcal{D}^k_\mu(N) \subset T\mathcal{D}^s_\mu(N)$ and $\bar{R}$ is a vector bundle morphism $T\mathcal{D}^k_\mu(N) \to \Xi^k$.

Note that we have two equivalent descriptions of the flow of an ideal incompressible fluid on a manifold M with boundary. The first one uses an LHS on $\mathcal{D}^s_\mu(M)$, while the second one is in terms of an LHS on $\mathcal{D}^s_\mu(N)$ with constraint Ξ. This means that the solutions of flow equations in both descriptions must exist for the same values of t simultaneously. It turns out that the use of the second description sometimes simplifies the argument. We apply this in the proof of the following.

Theorem 16.20 (Regularity theorem) *Let X_0 and f_0 belong to the space $T_e\mathcal{D}^{s+k}_\mu(M) \subset T_e\mathcal{D}^s_\mu(M)$ and $\eta(t)$ be the flow of an ideal incompressible fluid on M with external force f_0 such that $\eta(0) = e$ and $\dot{\eta}(0) = X_0$. For M both with and without boundary the diffeomorphism $\eta(t)$ belongs to $\mathcal{D}^{s+k}_\mu(M)$ for all t for which the flow exists in $\mathcal{D}^s_\mu(M)$. Equivalently, the solution $X(t)$, $X(0) = X_0$, of Euler's equation with external force f_0 is an H^{s+k}-smooth vector field on M (i.e., $X(t) \in T_e D^{s+k}_\mu(M)$) for all t such that $X(t)$ exists as an element of H^s.*

Proof. Following [15, 17, 18, 19], we first analyze the more complex case where $\partial M \neq \emptyset$, and then conclude the proof by indicating what modifications can be made when $\partial M = \emptyset$.

Let, as before, N be a closed Riemannian n-manifold and let M be isometrically embedded in N via an embedding i. Recall that one may, for example, take the double of M as N. Then the metric on N is chosen to coincide with the original metric on one of the copies of $M \subset N$. First, let us use the constraint Ξ^s introduced in Theorem 16.17. Let $Y_0 = \bar{R}_e \circ i(X_0)$ and $f = \bar{R}_e \circ i(f_0)$. By definition, we have (see Remark 16.19)

$$Y_0 \in \Xi_e^s \cap T_e D_\mu^{s+k}(N) = \Xi_e^{s+k} \quad \text{and} \quad f \in \Xi_e^s \cap T_e D_\mu^{s+q}(N) = \Xi_e^{s+q} \ .$$

Denote by $\bar{Y}_0$ the right-invariant vector field on $\mathcal{D}_\mu^s(N)$ such that $(\bar{Y}_0)_e = Y_0$. Clearly, $\bar{Y}_0$ is a C^k-section of Ξ^s which is C^k-smooth as a vector field on $\mathcal{D}_\mu^s(N)$ (see Theorem 5.4).

Consider the mechanical system with the constraint Ξ^s and the vector field $T\bar{R}(\mathcal{S} + \bar{f}^l)$ on Ξ^s introduced above. (Recall that $\bar{f}^l$ is the vertical lift of the right-invariant vector field $\bar{f}(t)$.) By definition, $T\bar{R}(\mathcal{S} + \bar{f}^l)$ is right-invariant on $\mathcal{D}_\mu^s(N)$. This field is also C^k-smooth, since $\bar{f}(t)$ is C^k-smooth on $\mathcal{D}_\mu^s(N)$ and $T\bar{R}$ and $\mathcal{S}$ are both C^∞-smooth. Denote by ϕ_t the flow of $T\bar{R}(\mathcal{S} + \bar{f}^l)$ on Ξ^s. The local existence of ϕ_t is guaranteed by the smoothness of the field. In other words, for any initial condition $V \in \Xi^s$, the map $t \mapsto \phi_t(V)$ is defined for a sufficiently small t. Since the field is $\mathcal{D}_\mu^s(N)$-right-invariant, so is the flow ϕ_t. For a fixed t, the diffeomorphism ϕ_t is a C^q-smooth map $\Xi^s \to \Xi^s$. Since $\bar{Y}_0$ and ϕ_t are both right-invariant, the domain of the function $t \mapsto \phi_t(V)$ is independent of $V \in \bar{Y}_0$. For the same reason, the field $\bar{Y}(t) = \phi_t(\bar{Y}_0)$ is right-invariant for every t. Note also that because $\bar{Y}_0$ is a C^k-submanifold of Ξ^s, the field $\bar{Y}(t)$ is C^k-smooth on $\mathcal{D}_\mu^s(N)$ for every t. Denote by $Y_e(t)$ the vector of $\bar{Y}(t)$ that belongs to $\Xi_e^s \subset T_e\mathcal{D}_\mu^s(N)$. The results of Chapter 5 yield that $Y_e(t)$ is an H^{s+k}-smooth vector field on N for all t such that $\phi_t(\bar{Y}_0)$ exists, i.e., $Y_e(t) \in T_e\mathcal{D}_\mu^s(N)$.

Let $\tilde{\eta}(t)$ be the flow of $Y_e(t)$ on N. It follows from the results of Chapter 5 that $\tilde{\eta}(t) \in \mathcal{D}^{s+k}(N)$ as long as $Y_e(t) \in T_e\mathcal{D}_\mu^s(N)$. Furthermore, $\tilde{\eta}(t)$ is a C^k-smooth curve on $\mathcal{D}_\mu^s(N)$ (Section 5.1). In particular,

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\tilde{\eta}(t) = Y_e(t) \circ \tilde{\eta}(t) = \bar{Y}(t) \mid_{\tilde{\eta}(t)} .$$

By Corollary 16.18, we see that $\eta(t) = \tilde{\eta}(t) \mid_M$ is a trajectory of the LHS of an ideal incompressible fluid on M and $X(t) = Y_e(t) \mid_M$ is a solution of Euler's equation. It is clear that $\eta(t) \in \mathcal{D}_\mu^{s+k}(M)$ and $X(t) \in T_e\mathcal{D}_\mu^{s+k}(M)$ for all t for which $\eta(t)$ and $X(t)$ exist and belong to H^s.

If $\partial M = \emptyset$, the proof is easier; there is no need to pass to Y_0, f, and the fields on Ξ^s. Instead, working with X_0, f_0, and $\mathcal{S} + \bar{f}^l$ on $T\mathcal{D}_\mu^s(M)$, one can apply the same argument. $\square$

Corollary 16.21 *Let the force $f_0(t)$ be a continuous curve in $T_e\mathcal{D}_\mu^\infty(M)$ and let $X_0 \in T_e\mathcal{D}_\mu^\infty(M)$. Then $\eta(t) \in \mathcal{D}_\mu^\infty(M)$ as long as $\eta(t) \in T_e\mathcal{D}_\mu^s(M)$. Equivalently, $X(t)$ is C^∞-smooth as long as it is H^s-smooth.*

Remark 16.22. The idea of our proof of Theorem 16.20 was originally used in [61] to prove the regularity theorem on a closed manifold or on $M \setminus \partial M$. It is essential that in our method (developed in [17, 18]) we deal with vector fields on Ξ^s right-invariant under $\mathcal{D}^s_\mu(N)$ and that the latter group can "move" the boundary ∂M. On the other hand, the group $\mathcal{D}^s_\mu(M)$ used in [61] preserves ∂M. As a consequence, one cannot obtain regularity in the normal directions to ∂M by working only with the fields on $T\mathcal{D}^s_\mu(M)$.

Remark 16.23. In [151], the regularity for a manifold with boundary was proved in the particular case of a potential external force. The proof of this result can be reduced to the study of the flow of a free fluid. In our proof the force is only assumed to be divergence-free. As pointed out in Remark 16.8, the general case can be formally derived from the case we have analyzed. Note also that the regularity theorem for a general external force on a bounded domain in $\mathbb{R}^n$ was announced in [219, 220]. However, as the author later pointed out (see [221]), the proof was incomplete.

Let us now turn to the problem of whether or not two given elements of $\mathcal{D}^s_\mu(M)$ can be connected by a flow of an ideal incompressible fluid without external force, i.e., by a geodesic of the weak Riemannian metric. For $\dim M = 2$ and $\dim M = 3$, this problem was studied by Shnirelman [209] in the following context. Let $\eta \in \mathcal{D}^s_\mu(M)$ and let there exist a piecewise smooth curve $\eta(t)$, $t \in [0,1]$, in $\mathcal{D}^s_\mu(M)$ which joins id with η (i.e., $\eta(0) = \mathrm{id}$ and $\eta(1) = \eta$). In other words, we assume that η belongs to the path-connected component of id. Denote by $l_0(\eta(\cdot))\mid_0^1$ the length of the curve $\eta(t)$ evaluated with respect to the H^0-metric defined by (5.1). Thus,

$$l_0\left(\eta(\cdot)\right)\Big|_0^1 = \int_0^1 \sqrt{(\dot\eta(t), \dot\eta(t))}\,\mathrm{d}t.$$

Taking into account that the flow of an ideal incompressible fluid is a geodesic of (5.1), we see that the question is whether or not there exists a smooth extremal of l_0 with fixed end-points id and η.

The main result of [209] (Theorem 1.1) is as follows. Let M be the three-dimensional cube. Then there is a diffeomorphism η in the path-connected component of id such that for any piecewise smooth path $\eta(\cdot)$ with $\eta(0) = \mathrm{id}$ and $\eta(1) = \eta$, there exists a path $\eta'(\cdot)$ with the same end-points and strictly smaller length: $l_0(\eta'(\cdot))\mid_0^1 < l_0(\eta(\cdot))\mid_0^1$. As a consequence, η cannot be joined with id by a flow of the fluid.

For a two-dimensional manifold M, it is still unknown whether or not a given diffeomorphism η from the connected component of id in $\mathcal{D}^s_\mu(M)$ can be connected with id by a flow of the ideal fluid. However, it was conjectured in [209] that such a flow always exists.

The proof of the main theorem of [209] is based on the following important, though technical, results. Let $\mathrm{dist}(\xi_1, \xi_2)$ be the infimum of the l_0-lengths over all curves in $\mathcal{D}^s_\mu(M)$ which connect ξ_1 and ξ_2. As in the finite-dimensional case, dist is a metric (i.e., a Riemannian distance) on $\mathcal{D}^s_\mu(M)$. This metric induces the weak (i.e., H^0-) topology on $\mathcal{D}^s_\mu(M)$. (According to a result of

[209], the closure of $\mathcal{D}^s_\mu(M)$ with respect to dist contains no interior points.) It is shown in [209] that the diameter of $\mathcal{D}^s_\mu(M)$ with respect to dist is finite if M is three-dimensional and contractible, and infinite if M is a two-dimensional domain. When M is three-dimensional, we have the following estimate for dist. Let $\Delta_\xi(x) = \rho(x, \xi(x))$ be the distance from x to $\xi(x)$ on M. It is not hard to see that $\Delta_\xi \in L^2(M)$. As shown in [209], there exist constants $\alpha > 0$ and $C > 0$, which depend on M only, such that for every $\xi \in \mathcal{D}^s_\mu(M)$

$$\operatorname{dist}(\mathrm{id}, \xi) \le C \left(\|\Delta_\xi\|_{L^2}\right)^\alpha .$$

Let $n = \dim M$ and $s > n/2 + 1$. Recall that, by Theorem 5.14 (proved originally in [61]), for any manifold M, there exists an H^s-neighborhood W of $\mathrm{id} \in \mathcal{D}^s_\mu(M)$ such that every element of W belongs to a flow of a free ideal incompressible fluid starting at id. It is shown in [13] that W is also filled out by flows of the fluid with an external force, provided that the following smoothness hypothesis holds. We assume first that $s > n/2+1$ and either the external force f is independent of time and H^{s+1}-smooth (i.e., $f \in T_e\mathcal{D}^{s+1}_\mu(M)$) or, if f is time-dependent, then $f(t)$ is H^{s+1}-smooth for every t, continuous in t in $T_e\mathcal{D}^{s+1}_\mu(M)$ and C^1-smooth in t in $T_e\mathcal{D}^s_\mu(M)$. Furthermore, if $s > n/2 + 2$, then these assumptions can be relaxed: $f \in T_e\mathcal{D}^s_\mu(M)$ for an autonomous f; or, otherwise, $f(t)$ is continuous in t in $T_e\mathcal{D}^s_\mu(M)$.

To prove the latter assertion one passes to the group $\mathcal{D}^{s-1}_\mu(M)$ and then applies the regularity theorem (Theorem 16.20) together with the relative version of the theory of topological degree [13].

Remark 16.24. As is shown in [14], the neighborhood W is also covered by flows of a viscous incompressible fluid. The problem of determining the size of W seems to be interesting and important. The aforementioned results of [209] show that in the three-dimensional case W is strictly smaller than the connected component of id in $\mathcal{D}^s_\mu(M)$.

In conclusion, let us prove that the flow of a free incompressible fluid has a first integral. This result is analogous to the angular momentum conservation law for the motion of a rigid body with a stationary point. (See Section 11.2 and, in particular, Remark 11.4.) The existence of such a first integral is known in hydrodynamics as the circulation conservation law [3]. Apparently, this integral was originally considered in [211] by means of the Lagrangian approach for the group of C^∞-diffeomorphisms. Similarly to the finite-dimensional case, the existence of a first integral follows from the fact that the metric (the Lagrangian) is invariant with respect to the group structure (Noether's theorem). The proof below is obtained by adapting the standard finite-dimensional argument to the infinite-dimensional setting on $\mathcal{D}^s_\mu(M)$ (cf., [3]). An essentially new point in the proof is that now we apply the regularity theorem (Theorem 16.20).

Let $s > n/2 + 1$. Consider the set β^{s-1} of H^{s-1}-smooth divergence-free vector fields on M which are tangent to the boundary. The fields from β^{s-1}

are continuous on M, but may not be C^1-smooth. Thus, the vanishing of the divergence means only that β^{s-1} is orthogonal to the space of exact forms in the Hodge decomposition. (See (5.5) and (5.9).)

Denote by $\bar{\beta}^{s-1}$ the bundle over $\mathcal{D}^s_\mu(M)$ obtained by right-translations of $\beta^{s-1} = \bar{\beta}^{s-1}_e$, i.e., with fibers $\bar{\beta}^{s-1}_g = \{X \circ g \mid X \in \beta^{s-1},\ g \in \mathcal{D}^s_\mu(M)\}$. For $V \in \beta^{s-1}$, let $\bar{V}$ be the left-invariant cross-section of $\bar{\beta}^{s-1}$, i.e., $\bar{V}_g = Tg \circ V$. Observe that the fibers of $\bar{\beta}^{s-1}$ inherit the right-invariant inner product given by (5.1). Since $T_e\mathcal{D}^s_\mu(M) \subset \bar{\beta}^{s-1}$, the inner product $(\bar{V}_g, X_g)_g$, where $X_g \in T_g\mathcal{D}^s_\mu(M)$, is well-defined.

We emphasize that the metric given by (5.1) is simply the restriction of the metric on the fibers of $\bar{\beta}^{s-1}$ to the fibers of $T\mathcal{D}^s_\mu(M)$.

Let $X_0 \in T_e\mathcal{D}^s_\mu(M)$ and let $\eta(t)$ be the geodesic on $\mathcal{D}^s_\mu(M)$ (the flow of fluid) with initial conditions $\eta(0) = \mathrm{id}$ and $\dot{\eta}(0) = X_0$.

Theorem 16.25 *For any $V \in \beta^{s-1}$, the inner product $(\bar{V}, \dot{\eta})$ is constant along η.*

Proof. First, let us assume that

$$V \in \beta = T_e\mathcal{D}^s_\mu(M) \quad \text{and} \quad X_0 \in T_e\mathcal{D}^{s+1}_\mu(M) \subset T_e\mathcal{D}^s_\mu(M).$$

Then by Theorem 16.20, we have $\eta(t) \in \mathcal{D}^{s+1}_\mu(M)$ as long as $\eta(t)$ exists as an element of $\mathcal{D}^s_\mu(M)$. Hence, $\bar{V}_g = Tg \circ V \in T_g\mathcal{D}^s_\mu(M)$. Denote the flow of V by $g(\tau)$. In what follows we regard $g(\tau)$ as a one-parameter subgroup of $\mathcal{D}^s_\mu(M)$ and consider the right action $R_{g(\tau)}$ on $\mathcal{D}^s_\mu(M)$. Since $\eta \in \mathcal{D}^s_\mu(M)$, we have

$$\frac{\mathrm{d}}{\mathrm{d}\tau} R_{g(\tau)} \circ \eta = \frac{\mathrm{d}}{\mathrm{d}\tau} \big(\eta \circ g(\tau)\big) = T\eta \circ \frac{\mathrm{d}}{\mathrm{d}\tau} g(\tau) = T\eta \circ V = \bar{V}_\eta.$$

Thus, $\bar{V}$ is the generator of $R_{g(\tau)}$. Since the spray $\mathcal{S}$ on $T\mathcal{D}^s_\mu(M)$ is right-invariant under the action of $\mathcal{D}^s_\mu(M)$ (see Section 5.1), the curve $t \mapsto R_{g(\tau)}\eta(t)$ is a geodesic for every fixed τ.

Let

$$\eta(t,\tau) = R_{g(\tau)}\eta(t) \quad \text{and} \quad \dot{\eta}(t,\tau) = \frac{\mathrm{d}}{\mathrm{d}t}\, \eta(t,\tau).$$

By definition, we have $\frac{\mathrm{d}}{\mathrm{d}\tau}\eta(t,\tau) = \bar{V}$. It is easy to see that the fields $\dot{\eta}(t,\tau)$ and $\bar{V}$ commute on the intersection of their domains, i.e., $\big[\dot{\eta}(t,\tau), \bar{V}\big] = 0$. Clearly,

$$\bar{V}\,(\dot{\eta}(t,\tau), \dot{\eta}(t,\tau)) = \frac{\mathrm{d}}{\mathrm{d}\tau}\, (\dot{\eta}(t,\tau), \dot{\eta}(t,\tau)) = 0,$$

where $\bar{V}(\dot{\eta}(t,\tau), \dot{\eta}(t,\tau))$ is the derivative of the scalar function $(\dot{\eta}(t,\tau), \dot{\eta}(t,\tau))$ in the direction of $\bar{V}$. Because $\dot{\eta}(t,\tau)$ and $\bar{V}$ commute and the Levi-Civitá connection is torsion-free, we have

$$\tilde{\nabla}_{\dot{\eta}(t,\tau)}\bar{V} = \tilde{\nabla}_{\bar{V}}\dot{\eta}(t,\tau).$$

Observe that

$$\nabla_{\dot\eta(t,\tau)}\dot\eta(t,\tau) = \frac{\tilde{\mathrm{D}}}{\mathrm{d}t}\dot\eta(t,\tau) = 0.$$

Then, by the definition of the Riemannian connection (see formula (2.29)),

$$\begin{aligned} 0 = \bar V(\dot\eta,\dot\eta) &= 2\big(\tilde\nabla_V\dot\eta,\dot\eta\big) = 2\left(\tilde\nabla_{\dot\eta}\bar V,\dot\eta\right) \\ &= 2\left(\tilde\nabla_{\dot\eta}\bar V,\dot\eta\right) + 2\left(\tilde\nabla_{\dot\eta}\dot\eta,\bar V\right) = 2\frac{\mathrm{d}}{\mathrm{d}t}\left(\bar V,\dot\eta(t,\tau)\right) \end{aligned}$$

along $\eta(t,\tau)$. Therefore, $\frac{\mathrm{d}}{\mathrm{d}t}\big(\bar V,\dot\eta(t)\big) = 0$ along the geodesic $\eta(t) = \eta(t,0)$. This completes the proof of the particular case.

Now let us turn to the general case where $V \in \beta^{s-1}$ and $X_0 \in T_e\mathcal{D}^s_\mu(M)$. Recall that $\beta = T_e\mathcal{D}^s_\mu(M)$ is dense in β^{s-1} and $T_e\mathcal{D}^{s+1}_\mu(M)$ is dense in $T_e\mathcal{D}^s_\mu(M)$. Thus, there exists a sequence $V_j \in \beta^{s-1}$ converging to V in the H^{s-1}-topology and a sequence $X_i \in T_e\mathcal{D}^{s+1}_\mu(M)$ converging to X_0 in the H^s-topology. We pass to the limit as follows.

Let $\eta_i(t)$ be the geodesic in $\mathcal{D}^s_\mu(M)$ with $\dot\eta_i(0) = X_i$. Since the solution of a differential equation depends continuously on its initial condition, $\eta_i(t)$ converges to $\eta(t)$ uniformly on every finite interval. Note that for $\eta_i(t)$, X_i, and $\bar V_j$, the theorem has already been proved.

It is not hard to see that the linear map $V \mapsto (V,\cdot)$, where $(\cdot,\cdot)$ is the weak inner product (5.1), is a continuous embedding of β into $T^*_e\mathcal{D}^s_\mu(M)$. The convergence of the vector fields in the H^{s-1}-topology implies the convergence in the space H^{-s}. Recall that in the theory of Sobolev spaces the latter space is identified with the dual to H^s by means of the H^0-inner product. It is clear that the sequence $(V_j,\dot\eta_i)$ converges to $(V,\dot\eta)$ as $i,j \to \infty$. Since $(V_j,\dot\eta_i)$ is constant along η_i, the function $(V,\dot\eta)$ must also be constant along η. $\square$

16.4 Description of Deterministic Viscous Hydrodynamics Via a Stochastic Version of Newton's Law on Groups of Diffeomorphisms

Everywhere in this section we deal with fluids moving in a flat n-dimensional torus $\mathcal{T}^n$. Recall that $\mathcal{T}^n$ is the quotient space of $\mathbb{R}^n$ with respect the integral lattice where the Riemannian metric is inherited from $\mathbb{R}^n$ (see Sections 5.2 and 10.2).

16.4.1 General construction

This section is devoted to the approach to hydrodynamics in terms of the geometry of groups of diffeomorphisms, suggested for perfect fluids by Arnold

[2] and Ebin and Marsden [61] (see the previous sections). It turns out that to give an adequate description of viscous fluids in this language requires the involvement of stochastic processes (see, e.g., [106] and [107]). In particular, Newton's second law on the groups of diffeomorphisms, used in the case of perfect fluids, is replaced by its special stochastic analog in terms of Nelson's mean derivatives.

Here we use the material from Chapters 5 and 10. We shall deal with Itô type equations on $\mathcal{D}^s(\mathcal{T}^n)$. Since the connection on $\mathcal{D}^s(\mathcal{T}^n)$ is generated by the flat connection on the torus, the corresponding exponential map is like that on a linear space. So, without loss of generality we use a notation that is typical for Itô equations in linear spaces (see Remark 10.4). Below we consider a certain equation on the manifold $\bar{\beta}$ (a sub-bundle of $T\mathcal{D}^s(\mathcal{T}^n)$ introduced in Section 16.2) in general Belopolskaya-Daletskii form with respect to the exponential map of some special connection. The equations on $\mathcal{D}^s_\mu(\mathcal{T}^n)$ are considered as those on $\mathcal{D}^s(\mathcal{T}^n)$ subjected to a constraint $\bar{\beta}$ as described in Remark 10.4.

The main idea of the description of viscous hydrodynamics in the language of mean derivatives is as follows.

Let a random flow $\xi(t,m)$ with initial data $\xi(0,m) = m \in \mathcal{T}^n$ be given on a flat n-dimensional torus $\mathcal{T}^n$. Suppose that it is a general solution of a stochastic differential equation of the type

$$\xi(t,m) = m + \int_0^t a(s,\xi(s,m))\mathrm{d}s + \sigma w(t) \tag{16.14}$$

modulo factorization with respect to the integral lattice, with real constant $\sigma > 0$. Let $D_*\xi(t,m) = v(t,\xi(t,m))$, where $v(t,m)$ is a C^1-smooth (in t) and a C^2-smooth (in $m \in \mathcal{T}^n$) vector field on $\mathcal{T}^n$. Let $\xi(t,m)$ satisfy the relation

$$D_*D_*\xi(t,m) = F(t,m), \tag{16.15}$$

where $F(t,m)$ is a vector field on $\mathcal{T}^n$. Taking into account formula (8.25), we obtain

$$D_*D_*\xi(t,m) = \left(\frac{\partial}{\partial t}v + (v,\nabla)v - \frac{\sigma^2}{2}\nabla^2 v\right) = \frac{\partial}{\partial t}v + (v,\nabla)v - \frac{\sigma^2}{2}\nabla^2 v. \tag{16.16}$$

Thus (16.15) means that $v(t,m)$ satisfies the equality

$$\frac{\partial}{\partial t}v + (v,\nabla)v - \frac{\sigma^2}{2}\nabla^2 v = F(t,m), \tag{16.17}$$

which is the *Burgers equation* with viscosity $\frac{\sigma^2}{2}$ and external force $F(t,m)$. We interpret (16.15) as a stochastic analog of Newton's second law on the group of Sobolev diffeomorphisms $\mathcal{D}^s(\mathcal{T}^n)$.

The case of viscous incompressible fluids requires some additional constructions. Consider the vector space $T_e\mathcal{D}^s(\mathcal{T}^n)$ of all Sobolev H^s-vector fields ($s > \frac{n}{2}+1$) on $\mathcal{T}^n$ with the L^2-inner product introduced by formula (5.1) where M is replaced by $\mathcal{T}^n$. Consider the subspace $\beta = T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$ consisting of all divergence-free vector fields and the orthogonal projector $P_e : T_e\mathcal{D}^s(\mathcal{T}^n) \to T_e\mathcal{D}^s_\mu(\mathcal{T}^n) = \beta$ introduced in Section 5.1. Recall that by formula (5.7) for any $Y \in T_e\mathcal{D}^s(\mathcal{T}^n)$ we have $P_e(Y) = Y - \mathrm{grad}p$ where p is a certain H^{s+1} function on $\mathcal{T}^n$ that is unique to within an additive constant for given Y.

Let a random flow $\xi(t)$ be given on $\mathcal{T}^n$. Suppose that $\xi(t)$ is a general solution of a stochastic differential equation of the type (16.14). Let $D_*\xi(t) = u(t,\xi(t))$, where $u(t,x)$ is a divergence-free vector field on $\mathcal{T}^n$, C^1-smooth in t and C^2-smooth in $m \in \mathcal{T}^n$, and that $\xi(t,x)$ satisfies the relation

$$P_e D_* D_*\xi(t) = F(t,\xi(t)), \tag{16.18}$$

where $F(t,x)$ is a divergence-free vector field on $\mathcal{T}^n$. Taking into account formulae (8.25) and (5.7), we obtain

$$\begin{aligned} P_e D_* D_*\xi(t,x) &= P_e\left(\frac{\partial}{\partial t}u + (u,\nabla)u - \frac{\sigma^2}{2}\nabla^2 u\right) \\ &= \frac{\partial}{\partial t}u + (u,\nabla)u - \frac{\sigma^2}{2}\nabla^2 u - \mathrm{grad}p. \end{aligned} \tag{16.19}$$

Thus (16.18) means that the divergence-free vector field $u(t,x)$ satisfies the relation

$$\frac{\partial}{\partial t}u + (u,\nabla)u - \frac{\sigma^2}{2}\nabla^2 u - \mathrm{grad}p = F, \tag{16.20}$$

which is the *Navier-Stokes equation* with viscosity $\frac{\sigma^2}{2}$ and external force $F(t,x)$.

We interpret (16.18) as a stochastic analog of Newton's second law on the group $\mathcal{D}^s(\mathcal{T}^n)$, subject to the mechanical constraint $\bar{\beta}$ that is introduced in Section 16.2. In spite of the fact that the constraint is holonomic (i.e., integrable), we do not restrict attention to its integral manifolds. This allows us to apply both the finite and infinite-dimensional language to the investigation more easily.

Now we are in a position to describe this approach in detail. We present it for the case of a viscous incompressible fluid. The compressible case (leading to the Burgers equation as above) can be investigated by a simplification of the incompressible arguments and we leave it for the reader.

For simplicity of presentation, we suppose $s > \frac{n}{2} + 2$. This means that the H^s vector fields on $\mathcal{T}^n$ are at least C^2.

The definition of mean derivatives for processes on $\mathcal{D}^s(\mathcal{T}^n)$ is analogous to that on $\mathbb{R}^n$ and on $\mathcal{T}^n$. In order to distinguish the derivatives on $\mathcal{D}^s(\mathcal{T}^n)$ and on $\mathcal{T}^n$ we denote the former by $\bar{D}$, $\bar{D}_*$ and $\bar{D}_2$ while D, D_* and D_2 remain valid for $\mathcal{T}^n$.

Let $a(t,x)$ be a divergence-free H^s vector field on $\mathcal{T}^n$. Denote by $\bar{a}(t,f)$ the corresponding right-invariant vector field on $\mathcal{D}^s(\mathcal{T}^n)$. Consider also the right-invariant field of the linear operators $\bar{\mathsf{A}}$ introduced in Section 10.2, $\bar{\mathsf{A}}_g : \mathbb{R}^n \to T_g\mathcal{D}^s(\mathcal{T}^n)$. The flow on $\mathcal{T}^n$ generated by equation (16.14) is a solution of the equation

$$\mathrm{d}\xi(t) = \bar{a}(t,\xi(t))\mathrm{d}t + \sigma\bar{\mathsf{A}}(\xi(t))\mathrm{d}w(t) \tag{16.21}$$

on $\mathcal{D}^s(\mathcal{T}^n)$. One can easily see that for the quadratic mean derivative for such $\xi(t)$ we obtain the equality

$$\bar{D}_2\xi(t) = \sigma^2\bar{I} \tag{16.22}$$

where $\bar{I}$ is the field of unit operators in tangent spaces to $\mathcal{T}^n$. It is also evident that if $\xi(t)$ is described by an equation in Itô form and satisfies (16.22), the diffusion term of the equation is $\sigma\bar{\mathsf{A}}(\xi(t))\mathrm{d}w(t)$.

The mechanical interpetation of the sub-bundle $\bar{\beta}$ introduced in Section 16.2 is a constraint. According to the ideology of the geometric description of constraints from Section 11.6, we give the following definition.

Definition 16.26. A stochastic process $\xi(t)$ is said to be *forward admissible* to the constraint $\bar{\beta}$ if $D\xi(t) \in \bar{\beta}_{\xi(t)}$ a.s. for all t.

A stochastic process $\xi(t)$ is called *backward admissible* to the constraint $\bar{\beta}$ if $D_*\xi(t) \in \bar{\beta}_{\xi(t)}$ a.s. for all t.

A vector field X is called *admissible* if $X_f \in \beta_f$ at any $f \in \mathcal{D}^s(\mathcal{T}^n)$.

Following general ideas of mechanics with constraints we can introduce the notions of covariant mean derivatives with respect to a constraint. By $\bar{P}$ we denote the right-invariant field of projectors introduced by formula (5.6).

Definition 16.27. For an admissible vector field X and forward admissible process $\xi(t)$ the expression $\bar{P}\bar{D}X(t,\xi(t))$ is called the *covariant forward mean derivative* with respect to the constraint $\bar{\beta}$.

For an admissible vector field X and backward admissible process $\xi(t)$ the expression $\bar{P}\bar{D}_*X(t,\xi(t))$ is called the *covariant backward mean derivative* with respect to the constraint $\bar{\beta}$.

Let $\eta(t)$ be a backward admissible process. Then, according to Definition 16.27, we can consider the covariant backward mean derivative $\bar{P}\bar{D}_*\bar{D}_*\xi(t)$. Let $F(t,x)$ be a divergence-free H^s-vector field on $\mathcal{T}^n$, i.e., $F(t,x)$ can be considered as a time-dependent vector $F(t) \in \bar{\beta}_e$. Denote by $\bar{F}(t,f)$ the right-invariant vector field on $\mathcal{D}^s(\mathcal{T}^n)$ generated by $F(t)$.

Theorem 16.28 *Let for a process $\xi(t)$ on $\mathcal{D}^s(\mathcal{T}^n)$ the relation $\bar{D}_*\xi(t) = \bar{u}(t,\xi(t))$ holds where $\bar{u}(t,f)$ is a right-invariant vector field on $\mathcal{D}^s(\mathcal{T}^n)$, generated by a divergence-free H^s-vector field $u(t,x)$ on $\mathcal{T}^n$. If $\xi(t)$ satisfies* (16.22) *and the constrained Newton law*

$$\bar{P}\bar{D}_*\bar{D}_*\xi(t) = \bar{F}(t, \xi(t)) \tag{16.23}$$

$u(t, x)$ on $\mathcal{T}^n$ satisfies the Navier-Stokes equation (16.20).

The proof of Theorem 16.28 is reduced to the finite-dimensional arguments used above.

The divergence-free vector field $u(t, x)$ on $\mathcal{T}^n$ from Theorem 16.28, i.e., a time-dependent vector in $\beta_e \subset T_e\mathcal{D}^s(\mathcal{T}^n)$, can be obtained by right translation of the backward velocity $\bar{D}_*\xi(t)$ at e, and so the Navier-Stokes equation (16.20) plays the role of the Euler equation in the "algebra" $T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$ according to the general approach to Euler equations described in Section 11.2. The flow of $u(t, x)$ on $\mathcal{T}^n$, which is a curve on $\mathcal{D}^s(\mathcal{T}^n)$ describing the motion of a viscous incompressible fluid, may be considered as the expectation of the process $\xi(t)$.

So, we need to construct a backward admissible process on $\mathcal{D}^s(\mathcal{T}^n)$ satisfying (16.22) that also satisfies (16.23). It is a difficult problem to find a process with given backward mean derivatives. That is why we shall try to construct $\xi(t)$ by solving first a certain equation of type (16.21) and then changing the time direction in its solution.

Let a process $\eta(t)$ on $\mathcal{D}^s_\mu(\mathcal{T}^n)$ be a solution of the stochastic differential equation of type (16.21) with initial condition $\eta(0) = e$ and suppose it exists for t in a non-random time interval $[0, T]$. Consider the process with inverse time direction $\xi(t) = \eta(T - t)$. Our aim now is to construct an equation for η such that (16.23) is fulfilled for $\xi(t)$, and $\bar{D}_*\xi(t) = \bar{u}(t, \xi(t))$ where $\bar{u}(t, f)$ is an admissible right-invariant vector field with initial condition $u(0, e) = u_0 \in \beta_e$ where $u_0 = u_0(x)$ is a divergence-free H^s-vector field on $\mathcal{T}^n$.

Since the backward mean derivative for $\xi(t)$ is equal to the negative of the forward mean derivative for $\eta(T - t)$, we have $\bar{D}\eta(t) = -\bar{D}_*\xi(T - t) = -u(T - t, \eta(t))$. Hence, taking into account Theorem 8.7 and the fact that $T\pi\mathcal{S}(X) = X$ and $T\pi\tilde{F}^l = 0$, we deduce that $\xi(t)$ satisfies (16.23) if $\eta(t)$ satisfies the equality

$$\mathrm{d}\eta(t) = -\tilde{u}(T - t, \eta(t))\mathrm{d}t + \sigma\bar{\mathsf{A}}(\eta(t))\mathrm{d}w(t) \tag{16.24}$$

and the process $\bar{u}(T - t, \eta(t))$ in $\bar{\beta}$ satisfies the equality

$$D^\eta\tilde{u}(T - t, \eta(t)) = -\mathcal{S}(\tilde{u}(T - t, \eta(t))) - \bar{F}^l(T - t, \tilde{u}(T - t, \eta(t))) \tag{16.25}$$

where $\bar{F}^l(T - t, \tilde{u}(T - t, \eta(t)))$ is the vertical lift of $\bar{F}(T - t, \tilde{u}(T - t, \eta(t)))$.

Denote by $\bar{\mathsf{A}}^T$ the horizontal lift of the field $\bar{\mathsf{A}}$ onto $T\mathcal{D}^s(\mathcal{T}^n)$. Introduce on $\bar{\beta}$ the connection from Section 2.8. Recall that the projections of its geodesics onto $\mathcal{D}^s(\mathcal{T}^n)$ are geodesics of the connection $\bar{\mathsf{H}}$. Denote the exponential map of this connection by $\exp^T$.

Theorem 16.29 *Let the process $u(T - t, \eta(t))$ on $\bar{\beta}$ satisfy the Itô equation in Belopolskaya-Daletskii form*

$$\begin{aligned}\mathrm{d}\bar{u}(T-t,\eta(t)) = \exp^T_{\bar{u}(T-t,\eta(t))}(&-\mathcal{S}(\bar{u}(T-t,\eta(t)))\mathrm{d}t \\ &-F^l(t,\bar{u}(T-t,\eta(t)))\mathrm{d}t + \bar{\mathsf{A}}^T(\tilde{u}(T-t,\eta(t)))\mathrm{d}w(t)).\end{aligned} \tag{16.26}$$

Then the process $\eta(t)$ and the right-invariant admissible vector field $\bar{u}$ on $\mathcal{D}^s(\mathcal{T}^n)$ satisfy (16.24) and (16.25), respectively, and so $\xi(t) = \eta(T-t)$ satisfies (16.23) and (16.22). Hence the divergence-free vector field $u(t,x)$ on $\mathcal{T}^n$ is a solution of (16.20).

Theorem 16.29 follows from the infinite-dimensional version of Lemma 9.3.

The following finite-dimensional interpretation clarifies the construction. The process $\eta(t)$ with initial condition $\eta(0) = e$ on $\mathcal{D}^s(\mathcal{T}^n)$ that satisfies (16.24) is a random flow on $\mathcal{T}^n$. Denote this flow by $\eta(t,x)$ with $\eta(0,x) = x$. The flow $\eta(t,x)$ is the general solution of the Itô stochastic differential equation on $\mathcal{T}^n$

$$\mathrm{d}\eta(t,x) = -u(T-t,\eta(t,x))\mathrm{d}t + \sigma \mathrm{d}w(t) \tag{16.27}$$

with $\mathrm{div}u(t,x) = 0$, the finite-dimensional version of (16.24). By direct calculation of the forward mean derivatives for the finite dimensional process $\eta(t,x)$ we show that $D\eta(t,x) = -u(T-t,\eta(t,x))$ and

$$\begin{aligned}PDD\eta(t,x) = &\frac{\partial}{\partial t}u(T-t,\eta(t,x)) + (u(T-t,\eta(t,x)),\nabla)u(T-t,\eta(t,x)) \\ &-\frac{\sigma^2}{2}\nabla^2 u(T-t,\eta(t,x)) - \mathrm{grad}p.\end{aligned}$$

The latter equality is turned into (16.19) under the change of variables $\eta(t,x) = \xi(T-t)$. Thus equation (16.26) guarantees that for the process $\eta(t)$ satisfying (16.27) the relation $PDD\eta(t,x) = F(t,x)$ holds.

For a stochastic differential equation with respect to a process $\zeta(t)$ on $\mathcal{D}^s(\mathcal{T}^n)$ denote by $\zeta_t(s)$ its solution with initial condition $\zeta_t(t) = e$. Consider the following system on $\mathcal{D}^s(\mathcal{T}^n)$:

$$\begin{aligned}\mathrm{d}\eta(t) &= -\tilde{u}(T-t,\eta(t))\mathrm{d}t + \sigma\bar{\mathsf{A}}(\eta(t))\mathrm{d}w(t) \\ u(t) &= PE\mathsf{Q}_e(u_0 \circ \eta_{T-t}(T)) + \int_0^t F(s)\mathrm{d}s\end{aligned} \tag{16.28}$$

where Q_e is introduced in Definition 5.16 and $u_0 = u(0) \in \bar{\beta}_e$ is the initial value for $u(t)$ (see above). Notice that the first equation of (16.28) is (16.24).

Theorem 16.30 *If the process $\eta(t)$ and the vector $u(t)$ satisfy (16.28), then $u(t)$, considered as a divergence-free vector field on $\mathcal{T}^n$, satisfies (16.20).*

Indeed, taking into account the routine stochastic presentation of solutions of PDEs one can easily derive from the second equation of (16.28) that $PDD\eta(t,x) = F(t,x)$.

The system (16.28) is similar to the system considered by Ya. Belopolskaya in [22] (see also [24]). Equation (16.27) is a part of another system of stochastic differential equations, connected with the Navier-Stokes equation, that was also studied by B. Busnello [36].

We should also mention [42] where the theory of forward-backward stochastic differential equations on the diffeomorphism group of a torus is used to describe viscous incompressible hydrodynamics, and [39] where the infinite-dimensional Wiener process in the stochastic differential equations on that group is used.

16.4.2 Solutions of Burgers, Reynolds and Navier-Stokes equations via stochastic perturbations of inviscid flows

If the backward mean derivative of a process satisfying the stochastic Newton law of the previous section is not generated by a right-invariant vector field, in the "algebra", after passing to the Euler approach, some other types of hydrodynamical equations may arise.

In this section we introduce a special stochastic perturbation of a flow of diffuse matter such that the perturbed flow satisfies the stochastic Newton law of type (16.15), and show that the corresponding curve in the tangent space at the unit satisfies Burgers' equation. The same perturbation of a flow of a perfect incompressible flow without external force satisfies the stochastic Newton law (16.23) with $\bar{F} = 0$, but yields a curve in the tangent space at the unit that is a solution of a Reynolds type equation. Nevertheless, under the action of a certain special external force on the flow, this curve becomes a solution of a Navier-Stokes equation without external force. As above, we consider a fluid motion on the flat n-dimensional torus $\mathcal{T}^n$.

In this section we take $s > \frac{n}{2} + 2$ so that the diffeomorphisms from $\mathcal{D}^s(\mathcal{T}^n)$ and $\mathcal{D}^s_\mu(\mathcal{T}^n)$ are C^2-smooth and $T_e\mathcal{D}^s(\mathcal{T}^n)$ consists of C^2-smooth vector fields. We often use the operator Q introduced in Definition 5.16(iii).

Everywhere below we use the same process $W^{(\sigma)}(t)$ constructed from a Wiener process $w(t)$ in $\mathbb{R}^n$ by formula (10.5) (see also Theorem 10.5). If, in the formula, several random elements appear with subscript ω, this means that they all are taken at "the same" $\omega \in \Omega$, i.e., sometimes the formula may be considered as a description of a non-random element depending on the parameter $\omega \in \Omega$.

Let $g(t)$ be a solution of (16.1) with initial conditions $g(0) = e$ and $\dot{g}(0) = v_0 \in T_e\mathcal{D}^s(\mathcal{T}^n)$. Recall (see Proposition 16.2) that such a solution exists on some time interval $t \in [0, T]$ (for the sake of convenience we take a closed interval inside the domain of $g(t)$). Recall also that $g(t)$ is a flow of diffuse matter without external forces. Consider $v(t) = R^{-1}_{g(t)}\dot{g}(t) = \dot{g}(t) \circ g^{-1}(t) \in$

$T_e\mathcal{D}^s(\mathcal{T}^n)$. This infinite-dimensional vector, considered as a vector field on $\mathcal{T}^n$, will also be denoted by $v(t,m)$. Recall that this vector field satisfies the Hopf equation (16.4).

Consider a process on $\mathcal{D}^s(\mathcal{T}^n)$ of the form $\eta(t) = W^{(\sigma)}(t)\circ g(t)$, $t\in[0,T]$, where $W^{(\sigma)}(t)$ is defined by (10.5). In the finite-dimensional terminology, $\eta(t)$ is a random diffeomorphism of $\mathcal{T}^n$ of the form $\eta(t,m) = g(t,m)+\sigma w(t)$ modulo the factorization with respect to the integral lattice. Introduce the process $\xi(t) = \eta(T-t)$, i.e., in the finite-dimensional notation, $\xi(t,m) = g(T-t,m)+\sigma w(T-t)$.

Since $w(t)$ is a martingale with respect to its own "past", one can easily derive from the properties of conditional expectation that $D_*\xi(t) = \dot g(T-t,m) = v(T-t, g(T-t,m))$ and so $\bar P D_* D_*\xi(t) = \frac{\bar{\mathrm{D}}}{\mathrm{d}s}\dot g(s)_{|s=T-t} = 0$.

Consider the random process

$$\xi_t(s) = \xi(s)\circ\xi^{-1}(t) = W^{(\sigma)}(T-s)\circ g(T-s)\circ g^{-1}(T-t)\circ(W^{(\sigma)}(T-t))^{-1}.$$

Notice that the random diffeomorphism $(W^{(\sigma)}(t))^{-1}$ acts by the rule

$$(W^{(\sigma)}(t))^{-1}(m) = m-\sigma w(t).$$

Obviously $\xi_t(t) = e$. A finite-dimensional description of this process can be given as follows.

By construction $m = \xi(t,\xi^{-1}(t,m)) = g(T-t,\xi^{-1}(t,m))+\sigma w(T-t)$. Then $g(T-t,\xi^{-1}(t,m)) = m-\sigma w(T-t)$ and so $\xi^{-1}(t,m) = g^{-1}(T-t, m-\sigma w(T-t))$. Thus,

$$\begin{aligned}\xi_t(s,m) &= \xi(s, g^{-1}(T-t, m-\sigma w(T-t))\\ &= g\left(T-s, g^{-1}(T-t, m-\sigma w(T-t))\right)+\sigma w(T-s).\end{aligned}$$

We have $\xi_t(t,m) = m-\sigma w(T-t)+\sigma w(T-t) = m$, indeed, $\xi_t(t) = e$ on $\mathcal{D}^s(\mathcal{T}^n)$.

Since $\xi_t(t) = e$, the "present" σ-algebra $\mathcal{N}_t^{\xi_t}$ is trivial and so the conditional expectation with respect to it coincides with the ordinary mathematical expectation. Hence, using the relation between $v(t)$ and $g(t)$ and the definition of D_*, one can easily derive that

$$\begin{aligned}D_*\xi_t(s)_{|s=t} &= E\left(v(T-t, m-\sigma w(T-t))\right)\\ &= E\left(\mathsf{Q}_e TR^{-1}_{W^{(\sigma)}(T-t)} v(T-t)\right)\end{aligned} \tag{16.29}$$

(here t is fixed and the derivative is taken with respect to s).

Introduce on $\mathcal{T}^n$ the vector field $V(t,m) = E(v(t,m-\sigma w(t)))$. As an infinite dimensional vector, we denote it also by $V(t) = E(\mathsf{Q}_e TR^{-1}_{W^{(\sigma)}(t)} v(t))$. Formula (16.29) means that $D_*\xi_t(s)_{|s=t} = V(T-t)$.

Theorem 16.31 *The vector field $V(T-t,m)$ satisfies the Burgers equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}V(T-t,m)+(V(T-t,m)\cdot\nabla)V(T-t,m)-\frac{\sigma^2}{2}\nabla^2 V(T-t,m)=0, \quad (16.30)$$

where ∇^2 is the Laplace-Beltrami operator which on the flat torus coincides with the ordinary Laplacian.

Proof. For $t\in[0,T]$ and $\omega\in\Omega$, define the curve $\zeta_{t,\omega}(s)$ in $s\in[0,T]$ depending on the parameter ω by the formula

$$\zeta_{t,\omega}(s)=R^{-1}_{W^{(\sigma)}_\omega(T-t)}g(T-s,g^{-1}(T-t))=g(T-s,g^{-1}(T-t,m-\sigma w(T))).$$

Note the only difference between $\zeta_{t,\omega}(s)$ and $\xi_{t,\omega}(s)$: the stochastic summand $\sigma w(T-s)$ appears in the expression for $\xi_{t,\omega}(s)$ while it is absent in $\zeta_{t,\omega}(s)$. This means that $\zeta_{t,\omega}(s)$ is an a.s. smooth curve with random initial condition $\zeta_{t,\omega}(t)=(W^{(\sigma)}_\omega(T-t))^{-1}$.

We have $\frac{\mathrm{d}}{\mathrm{d}s}\zeta_{t,\omega}(s)_{|s=t}=-TR^{-1}_{W^{(\sigma)}_\omega(T-t)}v(T-s)$. Since $g(T-s)$ is a geodesic, from Proposition 5.7 it follows that "for almost all ω" (i.e., a.s. for $\omega\in\Omega$) the curve $\zeta_{t,\omega}(s)$ is also a geodesic, i.e., $\frac{\bar{\mathrm{D}}}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}s}\zeta_{t,\omega}(s)=0$. From the construction of the operator L_x in Section 5.2 it follows that the action of the diffeomorphism $W^{(\sigma)}_\omega(t)$ coincides with that of $L_{\sigma w_\omega(t)}$. Hence, by Theorem 5.19 the curve $(W^{(\sigma)}_\omega(T-t))\zeta_{t,\omega}(s)=L_{(\sigma w_\omega(T-t))}\zeta_{t,\omega}(s)$ is a.s. geodesic as well, i.e., $\frac{\bar{\mathrm{D}}}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}s}l_{(\sigma w_\omega(T-t))}\zeta_{t,\omega}(s)=0$. Note that

$$\frac{\mathrm{d}}{\mathrm{d}s}l_{(\sigma w_\omega(T-t))}\zeta_{t,\omega}(s)_{|s=t}=\mathsf{Q}_e\frac{\mathrm{d}}{\mathrm{d}s}\zeta_{t,\omega}(s)_{|s=t}=-\mathsf{Q}_eTR^{-1}_{W^{(\sigma)}_\omega(T-t)}v(T-t).$$

Recall that $E\mathsf{Q}_eTR^{-1}_{W^{(\sigma)}_\omega(T-t)}v(T-t)=V(T-t)$ and $D_*\xi_t(s)_{|s=t}=V(T-t)$ (see above). Then from the above arguments and constructions we derive that

$$D_*D_*\xi_t(s)_{|s=t}=D_*V(T-t,\xi_t(s))_{|s=t}=-E\left(\frac{\bar{D}}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}s}l_{(\sigma w_\omega(t))}\zeta_{t,\omega}(s)_{|s=t}\right)=0.$$

But since $D_*\xi_t(s)_{|s=t}=V(T-t)$, by formula (8.25) the backward derivative $D_*V(T-t,\xi_t(s))$ coincides with the left-hand side of (16.30). Hence (16.30) is satisfied. □

Now, let us turn to the case of viscous incompressible fluids. Let $g(t)$ be a solution of (16.9) on $\mathcal{D}^s_\mu(\mathcal{T}^n)$ with initial conditions $g(0)=e$ and $\dot g(0)=u_0\in T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$. Theorem 16.11 tells us that such a solution exists in some time interval $t\in[0,T]$ (for the sake of convenience we again take a closed interval inside the domain of $g(t)$). Recall that $g(t)$ is a flow of a perfect incompressible fluid without external forces. Consider $u(t)=\dot g(t)\circ g^{-1}(t)\in T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$. This infinite-dimensional vector, considered as a divergence-free vector field on $\mathcal{T}^n$, will be denoted $u(t,m)$. Recall that this vector field satisfies the Euler equation (16.11) without external forces (see Section 16.2).

Since $W^{(\sigma)}(t)$ takes values in $\mathcal{D}^s_\mu(\mathcal{T}^n)$ (see Theorem 10.5), we can repeat on $\mathcal{D}^s_\mu(\mathcal{T}^n)$ the above constructions for $\mathcal{D}^s(\mathcal{T}^n)$, i.e., define $\eta(t) = W^{(\sigma)}(t) \circ g(t)$, where $t \in [0,T]$, and $\xi(t) = \eta(T-t)$ (i.e., in the finite-dimensional notation $\xi(t,m) = g(T-t,m) + \sigma w(T-t))$. It is easy to see that $D_*\xi(t) = \dot{g}(T-t,m) = u(T-t, g(T-t,m))$ and so $\bar{P}D_*D_*\xi(t) = \frac{\tilde{\mathrm{D}}}{\mathrm{d}s}\dot{g}(s)_{|s=T-t} = 0$ on $\mathcal{D}^s_\mu(\mathcal{T}^n)$.

As above, the process $\xi_t(s) = \xi(s) \circ \xi^{-1}(t)$ has the property $\xi_t(t) = e$. Its finite-dimensional description is quite analogous to the case of $\mathcal{D}^s(\mathcal{T}^n)$.

Introduce on $\mathcal{T}^n$ the vector field $U(t,m) = E(u(t, m - \sigma w(t)))$ (a direct analog of $V(t,m)$). We also denote this field as an infinite dimensional vector by $U(t) = E(\mathsf{Q}_e T R^{-1}_{W^{(\sigma)}(t)} u(t))$.

Lemma 16.32 *The vector field $U(t,m)$ is divergence-free.*

Proof. By construction, for an elementary event $\omega \in \Omega$, the diffeomorphism $(W^{(\sigma)}(t)_\omega)^{-1}$ is a shift of the entire torus by a constant vector. Hence, Q_e applied to $TR^{-1}_{W^{(\sigma)}(t)_\omega} u(t)$ is the parallel translation on the torus of the entire divergence-free vector field $u(t)$ by the same constant vector into the opposite direction. Thus $\mathsf{Q}_e T R^{-1}_{W^{(\sigma)}(t)} u(t)$ is a random divergence-free vector field on the torus. Hence its expectation is divergence-free. □

So, $U(t) \in T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$. In particular, we have proved above that

$$D_*\xi_t(s)_{|s=t} = U(T-t). \tag{16.31}$$

Since nothing like Theorem 5.19 holds on $\mathcal{D}^s_\mu(\mathcal{T}^n)$, we have $PD_*D_*\xi_t(s)_{|s=t} = D_*U(T-t, \xi_t(s))_{|s=t} \neq 0$ (the value of this mean derivative is calculated in Remark 16.34 below). Hence there is no analog of Theorem 16.31. We can prove only the following:

Theorem 16.33 *The vector field $U(t,m)$ satisfies the following Reynolds type equation (see, e.g., [207]):*

$$\frac{\partial}{\partial t}U + E\left[\left((u\cdot\nabla)u\right)(t, m - \sigma w(t))\right] - \frac{\sigma^2}{2}\nabla^2 U - \operatorname{grad} p = 0. \tag{16.32}$$

Proof. It follows from the Itô formula that

$$\begin{aligned} &\mathrm{d}u(t, m - \sigma w(t)) \\ &= \frac{\partial u}{\partial t}(t, m - \sigma w(t))\mathrm{d}t + \frac{\sigma^2}{2}\nabla^2 u(t, m - \sigma w(t))\mathrm{d}t - \sigma\, u'\, \mathrm{d}w(t), \end{aligned}$$

where ∇^2, as above, is the Laplace-Beltrami operator and u' is the linear operator of the derivative of u in $m \in \mathcal{T}^n$.

Recall that $u(t,m)$ satisfies the Euler equation without external force, i.e., $\frac{\partial u}{\partial t} = -P((u\cdot\nabla)u)$. Since

$$E\left(\frac{\mathrm{d}}{\mathrm{d}t}u(t, m - \sigma w(t))\right) = \frac{\partial}{\partial t}Eu(t, m - \sigma w(t)) = \frac{\partial}{\partial t}U(t)$$

and $E(\sigma(\nabla u)\mathrm{d}w(t)) = 0$, we derive that

$$\begin{aligned}\frac{\partial}{\partial t}U &= E\left(\frac{\mathrm{d}}{\mathrm{d}t}u(t, m - \sigma w(t))\right)\\ &= E\left[-P\left((u\cdot\nabla)u\right)(t, m-\sigma w(t)) + \frac{\sigma^2}{2}\nabla^2 u(t, m - \sigma w(t))\right]\\ &= -E\left[\left((u\cdot\nabla)u\right)(t, m - \sigma w(t))\right] + \frac{\sigma^2}{2}\nabla^2 U + \operatorname{grad} p.\end{aligned}$$

□

There are standard methods for transforming (16.32) into the standard Reynolds form (see [207]). For a divergence-free vector field $X(m)$ on $\mathcal{T}^n$ (i.e., for a vector $X \in T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$) introduce the random divergence-free vector field $\breve{U}_X(t,m) = X(m - \sigma w(t)) - E(X(m - \sigma w(t))$ (i.e., the vector $\breve{U}_X(t) = \mathsf{Q}_e TR^{-1}_{W^{(\sigma)}(t)}X - E(\mathsf{Q}_e TR^{-1}_{W^{(\sigma)}(t)}X)$ in $T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$). For $X = u(t)$, we obtain $\breve{U}_{u(t)}(t,m) = u(t, m - \sigma w(t)) - U(t,m)$ and so $u(t, m-\sigma w(t)) = U(t,m) + \breve{U}_{u(t)}(t,m)$ and $E\breve{U}_{u(t)}(t,m) = 0$. Then one can easily see that $E([(u\cdot\nabla)u](t, m - \sigma w(t))) = (U\cdot\nabla)U + E[(\breve{U}_{u(t)}\cdot\nabla)\breve{U}_{u(t)}]$. Thus, (16.32) transforms into

$$\frac{\partial}{\partial t}U + (U\cdot\nabla)U - \frac{\sigma^2}{2}\nabla^2 U - \operatorname{grad} p = -E[(\breve{U}_{u(t)}\cdot\nabla)\breve{U}_{u(t)}], \qquad (16.33)$$

which is the standard form of the Reynolds equation. It differs from the Navier-Stokes type relation with viscosity $\frac{\sigma^2}{2}$ by the external force $-E[(\breve{U}_{u(t)}\cdot\nabla)\breve{U}_{u(t)}]$ which depends on $u(t,m)$, not on $U(t,m)$. Recall that (16.33) describes the turbulent motion of a fluid if the dependence of $E[(\breve{U}_{u(t)}\cdot\nabla)\breve{U}_{u(t)}]$ on U is given (say, derived from experimental data, see [207]).

Remark 16.34. For $\xi_t(s)$ as introduced above, formula (16.31) tells us that $\bar{D}_*\xi_t(s)_{|s=t} = U(T-t)$. Then, taking into account formula (8.25), one can easily derive that

$$\bar{P}D_*D_*\xi_t(s)_{s=t} = \bar{P}D_*U(T-s, \xi_t(s))_{s=t} = \frac{\partial}{\partial t}U + (U\cdot\nabla)U - \frac{\sigma^2}{2}\nabla^2 U - \operatorname{grad} p.$$

Thus, (16.33) implies that $\bar{P}D_*D_*\xi_t(s)_{s=t} = -PE[(\breve{U}_{u(T-t)}\cdot\nabla)\breve{U}_{u(T-t)}]$.

Our next aim is to show that a slight modification of the above argument allows us to annihilate the external force in (16.33) by introducing a special random force field on $\mathcal{D}^s_\mu(\mathcal{T}^n)$ into (16.12).

For a random divergence-free a.s. H^{s+1}-vector field $X_\omega(m)$ on $\mathcal{T}^n$ (i.e., for a random vector $X_\omega \in T_e\mathcal{D}^{s+1}_\mu(\mathcal{T}^n) \subset T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$), construct the random vector field $\breve{U}_{X_\omega}(t,m)$ which, for any $\omega \in \Omega$, is given by the formula

$$\breve{U}_{X_\omega}(t,m) = X_\omega(m - \sigma w_\omega(t)) - E(X_\omega(m - \sigma w_\omega(t)).$$

Introduce the non-random H^s vector field $PE[(\breve{U}_{X_\omega} \cdot \nabla)\breve{U}_{X_\omega}]$ and then construct the random vector $\mathfrak{F}_\omega(t, X_\omega)$ in $T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$ by the formula

$$\mathfrak{F}_\omega(t, X_\omega) = \mathsf{Q}_e TR_{W^{(\sigma)}_\omega(t)} PE[(\breve{U}_{X_\omega} \cdot \nabla)\breve{U}_{X_\omega}].$$

Note that $PE[(\breve{U}_{X_\omega} \cdot \nabla)\breve{U}_{X_\omega}]$ and hence $\mathfrak{F}_\omega(t, X_\omega)$ lose their derivatives, i.e., they are H^s-vector fields only since X_ω (and so $\breve{U}_{X_\omega}$) is H^{s+1}. Thus $\mathfrak{F}_\omega(t, X_\omega)$ is well-defined only on an everywhere dense subset $T_e\mathcal{D}^{s+1}_\mu(\mathcal{T}^n)$ of $T_e\mathcal{D}^s_\mu(\mathcal{T}^n)$.

Now introduce the right-invariant vector force field $\bar{\mathfrak{F}}_\omega(t, g, Y_\omega)$ on $\mathcal{D}^s_\mu(\mathcal{T}^n)$, where $Y_\omega \in T_g\mathcal{D}^s_\mu(\mathcal{T}^n)$, which for $g \in \mathcal{D}^{s+1}_\mu(\mathcal{T}^n)$ and $\omega \in \Omega$ is determined by the formula

$$\bar{\mathfrak{F}}_\omega(t, g, Y_\omega) = TR_g \mathfrak{F}_\omega(t, TR^{-1}_g Y_\omega),$$

where $TR^{-1}_g Y_\omega$ is a divergence-free a.s. H^{s+1}-vector field.

Consider the equation

$$\frac{\tilde{\mathrm{D}}}{\mathrm{d}t}\dot{g}_\omega(t) = \bar{\mathfrak{F}}_\omega(t, g_\omega(t), \dot{g}_\omega(t)) \tag{16.34}$$

on $\mathcal{D}^s_\mu(\mathcal{T}^n)$ whose right-hand side is well-defined on the everywhere dense subset $\mathcal{D}^{s+1}_\mu(\mathcal{T}^n)$ in $\mathcal{D}^s_\mu(\mathcal{T}^n)$. Equation (16.34) has no diffusion term and so it is an ordinary differential equation with parameter $\omega \in \Omega$. Here we do not investigate the solvability of (16.34) but suppose that for the initial conditions $g_\omega(0) = e$ and $\dot{g}_\omega(0) = u_0 \in T_e\mathcal{D}^{s+1}_\mu(\mathcal{T}^n)$ it a.s. has a unique H^{s+1}-solution $g_\omega(t)$ which is a.s. well-defined on a non-random time interval $t \in [0, T]$ for some $T > 0$. Consider the divergence-free a.s. H^{s+1}-vector field $u_\omega(t, m)$ on $\mathcal{T}^n$ given by the relation $\dot{g}_\omega(t) = u_\omega(t, g_\omega(t))$. The analog of the above-mentioned vector U now takes the form

$$\mathbb{U}(t, m) = E\big(u_\omega(t, m - \sigma w_\omega(t))\big) = E\mathsf{Q}_e TR^{-1}_{W^{(\sigma)}_\omega(t)} u_\omega(t). \tag{16.35}$$

As in Lemma 16.32 it is easy to see that the vector field (16.35) is divergence-free.

Theorem 16.35 *The divergence-free vector field* $\mathbb{U}$ *given by* (16.35) *satisfies the Navier-Stokes equation without external force and with viscosity* $\frac{\sigma^2}{2}$*:*

$$\frac{\partial}{\partial t}\mathbb{U} + (\mathbb{U} \cdot \nabla)\mathbb{U} - \frac{\sigma^2}{2}\nabla^2\mathbb{U} - \operatorname{grad} p_1 = 0. \tag{16.36}$$

Proof. For the random field $u'_\omega(t, m)$ of linear operators and the random field $u''_\omega(t, m)$ of bilinear operators (here the primes denote derivatives of u in $m \in \mathcal{T}^n$) the stochastic integrals $\int_0^t u'_\omega(t, m)\mathrm{d}w_\omega(t)$ and

$$\int_0^t u''_\omega(t,m)(\mathrm{d}w_\omega(t),\mathrm{d}w_\omega(t)) = \int_0^t \operatorname{tr} u''_\omega \mathrm{d}t = \int_0^t \nabla^2 u_\omega \,\mathrm{d}t$$

are well-defined. Then by applying standard arguments to the Taylor series expansion of u_ω, one can easily see that the Itô formula is well-defined for $u_\omega(t, m - \sigma w_\omega(t))$ and so

$$\begin{aligned} &E\left(\mathrm{d}u_\omega(t, m-\sigma w_\omega(t))\right) \\ &= E\left(\frac{\partial}{\partial t}u_\omega(t,m-\sigma w_\omega(t))\mathrm{d}t + \frac{\sigma^2}{2}\nabla^2 u_\omega(t,m-\sigma w_\omega(t))\mathrm{d}t\right). \end{aligned}$$

From (16.34) it follows (see (16.12) and (16.13)) that $\frac{\partial}{\partial t}u_\omega = -P[(u_\omega\cdot\nabla)u_\omega] + \mathfrak{F}_\omega(t, u_\omega(t))$. Thus, in the same manner as in the proof of Theorem 16.33 and the derivation of (16.33), we obtain

$$\begin{aligned} \frac{\partial}{\partial t}\mathbb{U}(t,m) &= E\left(\frac{\mathrm{d}}{\mathrm{d}t}u_\omega(t,m-\sigma w_\omega(t))\right) = -E\Big[\Big((u_\omega\cdot\nabla)u_\omega\Big)(t,m-\sigma w_\omega(t))\Big] \\ &\quad + \frac{\sigma^2}{2}\nabla^2\mathbb{U} + \operatorname{grad} p + E\mathsf{Q}_e TR^{-1}_{W^{(\sigma)}(t)}\mathfrak{F}_\omega(t,u_\omega(t)) \\ &= -(\mathbb{U}\cdot\nabla)\mathbb{U} + \frac{\sigma^2}{2}\nabla^2\mathbb{U} + \operatorname{grad}p - E[(\breve{U}_{u_\omega(t)}\cdot\nabla)\breve{U}_{u_\omega(t)}] \\ &\quad + E\mathsf{Q}_e TR^{-1}_{W^{(\sigma)}(t)}\mathfrak{F}_\omega(t,u_\omega(t)). \end{aligned} \tag{16.37}$$

But by construction and by formulae (5.13) and (5.14) we get

$$\begin{aligned} &E\mathsf{Q}_e TR^{-1}_{W^{(\sigma)}_\omega(t)}\mathfrak{F}_\omega(t,u_\omega(t)) \\ &= E\mathsf{Q}_e TR^{-1}_{W^{(\sigma)}_\omega(t)}\mathsf{Q}_e TR_{W^{(\sigma)}_\omega(t)} PE\left[(\breve{U}_{u_\omega(t)}\cdot\nabla)\breve{U}_{u_\omega(t)}\right] \\ &= E\mathsf{Q}_e TR^{-1}_{W^{(\sigma)}_\omega(t)} TR_{W^{(\sigma)}_\omega(t)}\mathsf{Q}_{W^{(\sigma)}_\omega(t)^{-1}} PE\left[(\breve{U}_{u_\omega(t)}\cdot\nabla)\breve{U}_{u_\omega(t)}\right] \\ &= PE\left[(\breve{U}_{u_\omega(t)}\cdot\nabla)\breve{U}_{u_\omega(t)}\right]. \end{aligned} \tag{16.38}$$

Since $\mathbb{U}$ is divergence-free, the vector fields $\frac{\partial}{\partial t}\mathbb{U}$ and $\nabla^2\mathbb{U}$ are divergence-free as well. Hence, $\operatorname{grad} p$ in (16.37) is taken from relation (5.7) for $Y = E([(u_\omega\cdot\nabla)u_\omega](t,m-\sigma w_\omega(t)))$, i.e.,

$$PE\left([(u_\omega\cdot\nabla)u_\omega](t,m-\sigma w_\omega(t))\right) = E\left([(u_\omega\cdot\nabla)u_\omega](t,m-\sigma w_\omega(t))\right) - \operatorname{grad} p.$$

Define $\operatorname{grad} p_1$ and $\operatorname{grad} p_2$ by the relations $P(\mathbb{U}\cdot\nabla)\mathbb{U} = (\mathbb{U}\cdot\nabla)\mathbb{U} - \operatorname{grad} p_1$ and $PE[(\breve{U}_{u_\omega(t)}\cdot\nabla)\breve{U}_{u_\omega(t)}] = E[(\breve{U}_{u_\omega(t)}\cdot\nabla)\breve{U}_{u_\omega(t)}] - \operatorname{grad} p_2$. Clearly, $\operatorname{grad} p = \operatorname{grad} p_1 + \operatorname{grad} p_2$ (i.e., to within an additive constant $p = p_1 + p_2$). Thus (16.36) follows from (16.37) and (16.38) in the natural form $\frac{\partial}{\partial t}\mathbb{U} + (\mathbb{U}\cdot\nabla)\mathbb{U} - \frac{\sigma^2}{2}\nabla^2\mathbb{U} - \operatorname{grad} p_1 = 0$. □

References

1. Arnold, L.: Random dynamical systems. Springer-Verlag, Berlin et al. (1998)
2. Arnold, V.: Sur la géometrie différentielle des groupes de Lie de dimension infinie et ses applications a l'hydrodynamique des fluides parfaits. Ann. Inst. Fourier **16**, 319-361 (1966)
3. Arnold, V.I.: Mathematical methods of classical mechanics. Springer, New York (1989)
4. Akhmerov, R.R., Kamenskiĭ, M.I., Rodkina, A.E., Sadovskiĭ, B.N.: The theory of neutral type equations. In: Itogi nauki i tekhniki. Matematicheskij Analiz., vol. 19, pp 55-126. VINITI, Moscow (1982) (Russian)
5. Aubin, J.-P., Cellina, A.: Differential Inclusions. Set-valued maps and viability theory. Springer-Verlag, Berlin et al. (1984)
6. Azarina, S.V., Gliklikh, Yu.E.: Differential equations and inclusions with forward mean derivatives in $\mathbb{R}^n$. Proceedings of Voronezh State University **2006** (2), 138-146 (2006)
7. Azarina, S.V., Gliklikh, Yu.E.: Differential inclusions with mean derivatives. Dynamic Systems and Applications. **16** (1), 49-71 (2007)
8. Azarina, S.V., Gliklikh, Yu.E.: On differential equations and inclusions with mean derivatives on a compact manifold. Discussiones Mathematicae. DICO. **27** (2), 385-397 (2007)
9. Azarina, S.V., Gliklikh, Yu.E.: Stochastic differential inclusions in terms of infinitesimal generators and mean derivatives. Applicable analysis **88** (1), 89-105 (2009)
10. Azarina, S.V., Gliklikh, Yu.E.: Stochastic differential equations and inclusions with mean derivatives relative to the past. International Journal of Difference Equations **4** (1), 27-41 (2009)
11. Azarina, S.V., Gliklikh, Yu.E., Obukhovskii, A.V.: Solvability of Langevin differential inclusions with set-valued diffusion terms on Riemannian manifolds. Applicable Analysis **86** (9), 1105-1116 (2007)
12. Azencott, R.: Behavior of diffusion semi-groups at infinity. Bull. Soc. Math. France. **102**, 193-240 (1974)
13. Baranov, Yu.S.: The two-point end-value problem for equations of perfect incompressible fluid. In: Strygin V.V. (ed.) Priblizhennye metody issledovaniya differentsial'nykh uravnenij i ikh prilozheniya. pp. 3-8. Kuibyshev University (1983) (Russian)
14. Baranov, Yu.S.: The two-point end-value problem for equations of viscous incompressible fluid. In: Borisovich Yu.G. (ed.) Analiz na mnogoobraziyakh i differentsial'nye uravneniya, pp. 134-138. Voronezh University (1986) (Russian)
15. Baranov, Yu.S., Gliklikh, Yu.E.: A remark on the regularity of solutions of Euler's equation of hydrodynamics. Russian Math. Surveys **36** (5), 163-164 (1981)

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics, DOI 10.1007/978-0-85729-163-9,

16. Baranov, Yu.S., Gliklikh, Yu.E.: On the regularity of flows of a perfect incompressible fluid on manifolds with boundary. In: Strygin V.V. (ed.) Priblizhennye metody issledovaniya differentsial'nykh uravnenij i ikh prilozheniya, pp. 24-31. Kuybyshev University (1982) (Russian)
17. Baranov, Yu.S., Gliklikh, Yu.E.: On the smoothness of some mappings of the group of diffeomorphisms. Voronezh University. Deposited in VINITI. July 5, 1983, N 3664 (Russian)
18. Baranov, Yu.S., Gliklikh, Yu.E.: On a non-integrable distribution on the infinite-dimensional manifold of diffeomorphisms. Differentsial'naya geometriya mnogoobrazij figur. Kaliningrad University, **17**, 11-17 (1986) (Russian)
19. Baranov, Yu.S., Gliklikh, Yu.E.: A certain mechanical constraint on the group of volume preserving diffeomorphisms. Funct. Anal. Appl. **22** (2), 61-62 (1988)
20. Bates, L.: You can't get there from here. Differential geometry and its applications. **8**, 273-274 (1998)
21. Baxendale, P.: Measures and Markov processes on function spaces. Bull. Soc. Math. France, **46**, 131-141 (1976)
22. Belopolskaya, Ya.I.: Probabilistic presentation for solutions of boundary-value problems for hydrodynamical equations. Trudy POMI **249**, 71-102 (1997)
23. Belopolskaya, Ya.I., Dalecky, Yu.L.: Stochastic processes and differential geometry. Kluwer Academic Publishers, Dordrecht (1989)
24. Belopolskaya, Ya.I., Gliklikh, Yu.E.: Diffusion processes on the groups of diffeomorphisms and hydrodynamics of viscous incompressible fluid. Transactions of RANS, Series MMMIC. **3** (2), 27-35 (1999)
25. Billingsley P.: Convergence of Probability Measures. Wiley, New York (1969)
26. Bishop, R.L., Crittenden, R.J.: Geometry of Manifolds. Academic Press, N.Y.-L. (1964)
27. Blanchard, P., Combe, P., Zheng, W.: Mathematical and physical aspects of stochastic mechanics. / Lecture Notes in Physics. **281** (1987)
28. Bogacheva E.V., Gliklikh Yu.E.: Topological index for condensing maps on Finsler manifolds with applications to functional-differential equations of neutral type. Topological Methods in Nonlinear Analysis. **26** (2), 287-305 (2005)
29. Borisov, Yu.F.: Parallel translation along Hölder curves in a Riemannian space. Dokl. AN SSSR **197** (5), 995-998 (1971) (Russian)
30. Borisovich, Yu.G., Bliznyakov, N.M., Izrailevich, Ya.A., and Fomenko, T.N.: Introduction to Topology. Fizmatlit, Moscow (1995)
31. Borisovich, Yu.G., Gel'man, B.D., Myshkis, A.D., Obukhovskiĭ, V.V.: Introduction to the theory of set-valued mappings and differential inclusions. KomKniga, Moscow (2005) (Russian)
32. Borisovich, Yu.G., Gliklikh, Yu.E.: On the Lefschetz number for a certain class of set-valued maps. In: Trokhimchuk Yu.Yu. (ed.) 7-th Summer Mathematical School, pp. 283-294. Kiev (1970) (Russian)
33. Borisovič, Yu.G., Gliklih, Yu.E.: Fixed points of mappings of Banach manifolds and some applications. Nonlinear Analysis: TMA. **4** (1), 165-192 (1980)
34. Borisovich, A.Yu., Zvyagin, V.G.: On global convertibility of nonlinear operators, generated by boundary problems. In: Strygin V.V. (ed.) Approksimatsionnye metody issledovaniya differentsial'nykh uravnenii i ikh prilozheniya, pp 27-33. Kuibyshev University (1983)
35. Brzezniak, Z., Elworthy, K.D.: Stochastic differential equations on Banach manifolds. Methods of functional analysis and topology. **6** (1), 43-84 (2000)
36. Busnello, B.: A probabilistic approach to the two-dimensional Navier-Stokes equation. The Annals of Probability. **27** (4), 1750-1788 (1999)
37. Cardin, F., Favretti, M.: On nonholonomic and vakonomic dynamics of mechanical systems with nonintegrable constraints. J. of Geometry and Physics. **18**, 295-325 (1996)

38. Carlen, E.A.: Conservative diffusion. Commun. Math. Physics. **94**, 293-315 (1984)
39. Cipriano, F., Cruzeiro, A.B.: Navier-Stokes Equations and Diffusions on the Group of Homeomorphisms of the Torus. Commun. Math. Phys. **275**, 255-269 (2007)
40. Clark, J.M.C.: An introduction to stochastic differential equations on manifolds. In: Mayne D.Q., Brockett R.W. (eds) Geometric methods in system theory, pp. 131-149. Reidel, Dordrecht-Boston (1973)
41. Conway, E.D.: Stochastic equations with discontinuous drift. Trans. Amer. Math. Soc. **157** (1), 235-245 (1971)
42. Cruzeiro A.B., Shamarova, E.: Navier-Stokes equations and forward-backward SDEs on the group of diffeomorphisms of a torus. Stochastic processes and their applications **119**, 4034-4060 (2009)
43. Dalecky, Yu.L.: Stochastic differential geometry. Russian Mathematical Surveys **38** (3), 87-111 (1983)
44. Daletskii, Yu.L., Shnaiderman (Belopolskaya), Ya.I.: Diffusion and quasi-invariant measures on infinite-dimensional Lie groups. Funct. Anal. Appl. **9** (2), 88-90 (1969)
45. Dankel, T.G. (Jr): Mechanics on manifolds and the incorporation of spin into Nelson's stochastic mechanics. Arch. Rat. Mech. Anal. **37** (3), 192-222 (1970)
46. Darling, R.W.R.: On the convergence of Gangolli processes to Brownian motion on a manifold. Stochastics **12**, 277-301 (1984)
47. De Angelist, G.F., Jona-Lasino, G., Servo, M., Zanghi M.: Stochastic mechanics of Dirac particle in two spacetime dimensions. Jour. Phys., Ser. A. **19**, 865-871 (1986)
48. Deimling, K.: Multivalued differential equations. Walter de Gruyter, Berlin-New York (1992)
49. Derdzinski, A.: Geometry of the standard model of elementary particles. Springer-Verlag, Berlin et al. (1992)
50. DeWitt-Morette, C., Elworthy, K.D.: A stepping stone to stochastic analysis. New stochastic methods in physics. Physics Reports. **77** (3), 121-382 (1981)
51. Dirac, P.A.M.: General Theory of Relativity. John Wiley and Sons, New-York-London-Sydney-Toronto (1975)
52. Dmitriev, V.P.: Stochastic Mechanics. Vysshaya Shkola, Moscow (1990) (Russian)
53. Dohrn, D., Guerra, F.: Nelson's stochastic mechanics on Riemannian manifolds. Lettere al Nuovo Cimento **22** (4), 121-127 (1978)
54. Dohrn, D., Guerra, F.: Geodesic correction to stochastic parallel displacement of tensors. Lecture Notes in Physics. **93**, 165-181 (1979)
55. Dohrn, D., Guerra, F., Ruggiero, P.: Spinning particles and relativistic particles in framework of Nelson's stochastic mechanics. Lect. Notes Phys. **106**, 165-181 (1979)
56. Dombrowski, P.: On the geometry of tangent bundle. J. für die reine und adgewandte Mathematik. **210** (1/2), 73-88 (1962)
57. Dubrovin, B.A., Fomenko, A.T., Novikov, S.P.: Modern geometry – methods and applications, vol. 1. Springer, New York Berlin Heidelberg (1984)
58. Dynkin, D.G.: Diffusion of tensors. Dokl. AN SSSR **179** (6), 179-182 (1968)
59. Ebin, D.G.: The motion of slightly compressible fluids viewed as a motion with strong constraining force. Ann. Math. **105**, 141-200 (1977)
60. Ebin, D.G.: Viscous fluids in a domain with frictionless boundary. In: Rassias Th.M. (ed.) Global analysis – Analysis on manifolds. / Teubner Text zur Mathematik, **57**, pp. 93-110, Teubner, Leipzig (1983)
61. Ebin, D.G., Marsden, J.: Groups of diffeomorphisms and the motion of an incompressible fluid. Annals of Math. **92** (1), 102–163 (1970)
62. Egorov, Yu.V.: Linear differential equations of principal type. Nauka, Moscow (1984) (Russian)
63. Eisenhart, L.P.: Riemannian geometry. Princeton University Press, Princeton, 1926.
64. Eliasson, H.I.: Geometry of manifolds of maps. J. Diff. Geometry. **1**, 169-194 (1967)
65. Elliott, R.J., Anderson, B.D.O.: Reverse time diffusions. Stochastic Process. Appl. **19**, 327-339 (1985)

66. Elworthy, K.D.: Stochastic differential equations on manifolds. / Lect. Notes of London Math. Soc. **70**. Cambridge University Press, Cambridge (1982)
67. Elworthy, K.D.: Stochastic flows and the C_0-diffusion property. Stochastics. **6**, 233-238 (1982)
68. Elworthy, K.D., Le Jan, Y., Li, X.-M.: On the geometry of diffusion operators and stochastic flows / Lect. Notes Math. **1720** (1999)
69. Emery, M.: Stochastic calculus on manifolds. Springer, Berlin et al. (1989)
70. Emery, M.: On two transfer principles in stochastic differential geometry. Lecture Notes in Mathematics. **1426**, 407-441 (1989)
71. Faddeev, L.D., Vershik, A.M.: Differential geometry and Lagrangian mechanics with constraints. Dokl. AN SSSR **202** (3), 555-557 (1972)
72. Fedorov, F.I.: The Lorentz Group. Nauka, Moscow (1979)
73. Fenyes I.: Eine wahrscheinlichkeitstheoretische Begrüdung und Interpretation der Quantenmechanik. Zeitschrift für Physik. **132**, 81-106 (1952)
74. Filippov, A.F.: Differential equations with discontinuous right-hand sides. Kluwer Academic Publishers, Dordrecht (1988)
75. Freidlin, M.I.: On the factorization of positive semi-definite matrices. Probability Theory and Its Applications. **13** (2), 375-378 (1968)
76. Freidlin, M.I., Wentzel, A.D.: Random perturbations of dynamical systems. Springer, New York Berlin Heidelberg (1984)
77. Gangolli, R.: On the construction of a certain diffusion on differentiable manifolds. Z.F.W. **2**, 406-419 (1964)
78. Gantmakher, F.R.: The Theory of Matrices. Nauka, Moscow (1967) (Russian)
79. Gel'man, B.D.: Continuous approximations of set-valued mappings and fixed points. Mathematical Notices. **78** (2), 212-222 (2005)
80. Gel'man, B.D., Gliklikh, Yu.E.: The two-point boundary-value problem in geometric mechanics with discontinuous forces. Applied Mathematics and Mechanics. **44** (3), 565-569 (1980) (Russian)
81. Gikhman, I.I., Skorokhod, A.V.: Introduction to the Theory of Stochastic Processes. Nauka, Moscow (1977) (Russian)
82. Gihman, I.I., Skorohod, A.V.: The Theory of Stochastic Processes. Vol. 1. Springer-Verlag, New York (1974)
83. Gihman, I.I., Skorohod, A.V.: Theory of Stochastic Processes. Vol. 3. Springer-Verlag, New York (1979)
84. Gikhman, I.I., Skorokhod, A.V.: Stochastic differential equations and their applications. Naukova Dumka, Kiev (1982) (Russian)
85. Ginzburg, V.L.: Accessible Points and closed trajectories of mechanical systems. Appendix F in [107], pp. 192-201 (1997)
86. Gliklikh, Yu.E.: Integral operators on a manifold. Trudy matem. fakul'teta Voronezhskogo Universiteta. **4**, 29-35 (1971) (Russian)
87. Gliklikh, Yu.E.: Fixed points of multivalued mappings with nonconvex images and the rotation of multivalued vector fields. In: Borisovich Yu.G. (ed.) Sbornik Trudov Aspirantov Matematicheskogo Fakul'teta, pp. 30-38, Voronezh University, Voronezh (1972) (Russian)
88. Gliklikh, Yu.E.: On a certain generalization of the Hopf-Rinow theorem on geodesics. Russian Math. Surveys. **29** (6), 161-162 (1974)
89. Gliklikh, Yu.E.: On non-local prolongation conditions for integral curves of vector fields. Diff. Equations **12** (4), 743-744 (1977)
90. Gliklikh, Yu.E.: The two-point boundary-value problem in geometric mechanics of systems with a bounded force field. Voronezh University. Deposited in VINITI. June 6, 1977, N 2217-77 (Russian)
91. Gliklikh, Yu.E.: On a certain analogue of differential equations with discrete delay on a Riemanian manifold. Voronezh University. Deposited in VINITI. March 20, 1980, N 1089-80 (Russian)

92. Gliklikh, Yu.E.: On systems of geometric mechanics with a delayed control force. Voronezh University. Deposited in VINITI. October 14, 1980, N 4384-80 (Russian)
93. Gliklikh, Yu.E.: On equations with discrete delay on non-linear spaces. In: Perov A.I. (ed.) Operatornye metody v nelinejnom analize, pp. 29-33. Voronezh University Press, Voronezh (1982) (Russian)
94. Gliklikh, Yu.E.: Riemannian parallel translation in non-linear mechanics. Lect. Notes Math. **1108**, 128-151 (1984)
95. Gliklikh, Yu.E.: On Riemannian metrics possessing a uniform Riemannian atlas. Differentsial'naya geometriya mnogoobrazij figur (Kalinigrad). **16**, 17-19 (1985) (Russian)
96. Gliklikh, Yu.E.: The Itô line integral with Riemannian parallel displacement and stochastic equations on manifolds. Theory of Probability and Mathematical Statistics (Translations of Amer. Math. Soc.) **35**, 31-36 (1986)
97. Gliklikh, Yu.E.: Riemannian parallel translation, the Itô integral and stochastic equations on manifolds. Lecture Notes in Mathematics. **1214**, 171-195 (1986)
98. Gliklikh, Yu.E.: Operators of integral type and differential inclusions on manifolds subordinated to non-involutory distributions. In: Korneichuk N.P. (ed.) Nekotorye voprosy analiza i differentsial'noj topologii, pp. 22-28. Kiev (1988) (Russian)
99. Gliklikh, Yu.E.: Analysis of Riemannian Manifolds and Problems of Mathematical Physics. Voronezh University Press, Voronezh (1989) (Russian)
100. Gliklikh, Yu.E.: Stochastic differential geometry of the group of diffeomorphisms and the motion of viscous incompressible fluid. In: Fifth international Vilnius conference on probability theory and mathematical statistics. Abstracts, Vol. 1, pp. 173-174 (1989)
101. Gliklikh, Yu.E.: Velocity hodograph equation in mechanics on Riemannian manifolds. In: Janyska, J., Krupka, D. (eds.) Differential geometry and its applications, pp. 308-312. World Scientific, Singapore Teaneck (1990)
102. Gliklikh, Yu.E.: Calculation of Nelson's mean derivatives and the existence of solutions in stochastic mechanics. In: Borisovich Yu.G. (ed.) Problems of Geometry, Topology and Mathematical Physics (New Developments in Global Analysis series), pp. 17-36. Voronezh University Press, Voronezh (1992)
103. Gliklikh, Yu.E.: On existence theorems in ordinary and relativistic stochastic mechanics. In: Borisovich Yu. G. (ed.) Methods and Applications of Global Analysis (New Developments in Global Analysis series), pp. 61-80. Voronezh University Press, Voronezh (1993)
104. Gliklikh, Yu.E.: A new version of the Lagrangian approach to hydrodynamics of viscous incompressible fluid. Mathematical Notices. **55** (4), 15-24 (1994)
105. Gliklikh, Yu.E.: Calculation of Nelson's mean derivatives for Ito processes and some applications. In: Borisovich Yu. G. (ed.) Global and Stochastic Analysis (New Developments in Global Analysis series), pp. 29-39. Voronezh University Press (1995)
106. Gliklikh, Yu.E.: Ordinary and Stochastic Differential Geometry as a Tool for Mathematical Physics. Kluwer, Dordrecht (1996)
107. Gliklikh, Yu.E.: Global Analysis in Mathematical Physics. Geometric and Stochastic Methods. Springer-Verlag, New York (1997)
108. Gliklikh, Yu.E.: Stochastic equations in mean derivatives and their applications. I. Transactions of RANS, Series MMMIC. **1** (4), 26-52 (1997) (Russian)
109. Gliklikh, Yu.E.: Total and local topological indices for maps of Hilbert and Banach manifolds. Topological Methods in Nonlinear Analysis. **15** (1), 17-31 (2000)
110. Gliklikh, Yu.E.: Parallel translation and geodesics on the groups of diffeomerphisms of flat torus. In: Zvyagin V.G. (ed.) Topological methods of non-linear analysis, pp. 31-40. Voronezh State University, Voronezh (2000) (Russian)
111. Gliklikh, Yu.E.: Stochastic equations in mean derivatives and their applications. II. Transactions of RANS, Series MMMIC. **4** (4), 17-36 (2000) (Russian)

112. Gliklikh, Yu.E.: Deterministic viscous hydrodynamics via stochastic analysis on groups of diffeomorphisms. Methods of Functional Analysis and Topology. **9** (2), 146-153 (2003)
113. Gliklikh, Yu.E.: Deterministic viscous hydrodynamics via stochastic processes on groups of diffeomorphisms. In: Davis I.M. et al. (eds) Probabilistic Metohds in Fluids, pp. 179-190. World Scientific, Singapore (2003)
114. Gliklikh, Yu.E.: A necessary and sufficient condition for completeness of stochastic flows continuous at infinity. Warwick preprint: August 2004. No 8 (2004)
115. Gliklikh, Yu.E.: Global and stochastic analysis in problems of mathematical physics. KomKniga, Moscow (2005) (Russian)
116. Gliklikh, Yu.E.: Necessary and sufficient conditions for global in time existence of solutions of ordinary, stochastic, and parabolic differential equations. Abstract and Applied Analysis. **2006** (Article ID 39786), 1-17 (2006)
117. Gliklikh, Yu.E.: Solutions of Burgers, Reynolds and Navier-Stokes equations via stochastic perturbations of inviscid flows. Journal of nonlinear mathematical physics. **17** (Suppl. 1), 15-29 (2010)
118. Gliklikh, Yu.E., Fedorenko, I.V.: On the geometrization of a certain class of mechanical systems with random perturbations of the force. Voronezh University. Deposited in VINITI. October 21, 1980, N 4481 (Russian)
119. Gliklikh, Yu.E., Fedorenko, I.V.: Equations of geometric mechanics with random force fields. In: Strygin V.V. (ed.) Priblizhennye metody issledovaniya differentsial'nykh uravneniĭ i ikh prilozheniya, pp. 64-72. Kuĭbyshev University, Kuĭbyshev (1981)
120. Gliklikh, Yu.E., Morozova, L.A.: On the notion of L^1-completeness of a stochastic flow on a manifold. Abstract and Applied Analysis. **7** (12), 627-635 (2002)
121. Gliklikh, Yu.E., Morozova, L.A.: Conditions for global existence of solutions of ordinary differential, stochastic differential and parabolic equations. Internat. J. Math. and Math. Sciences. **2004** (17), 901-912 (2004)
122. Gliklikh, Yu.E., Obukhovskiĭ, A.V.: Stochastic differential inclusions of Langevin type on Riemannian manifolds. Discussiones Mathematicae. DICO. **21** (2), 173-190 (2001)
123. Gliklikh, Yu.E., Obukhovskiĭ, A.V.: A viable solution of a two-point boundary value problem for a second order differential inclusion on a Riemannian manifold. Proceedings of Voronezh State University, series Physics, Mathematics. **2003** (2), 144-149 (2003) (Russian)
124. Gliklikh, Yu.E., Obukhovskiĭ, A.V.: On a two point boundary value problem for second-order differential inclusions on Riemannian manifolds. Abstract and Applied Analysis. **2003** (10), 591-600 (2003)
125. Gliklikh, Yu.E., Obukhovskiĭ, A.V.: On differential inclusions of velocity hodograph type with Carathéodory conditions on Riemannian manifolds. Discussiones Mathematicae. DICO. **24**, 41-48 (2004)
126. Gliklikh, Yu.E., Ratiner, P.S.: On a certain type of second order differential equations on total spaces of fiber bundles with connections. In: Rassias Th.M. (ed.) Nonlinear Analysis in Geometry and Topology, pp. 99-106. Hadronic Press, Palm Harbor, Fl. (2000)
127. Gliklikh, Yu., Sinelnikov, A.: On global continuability of solutions of infinite-dimensional ordinary differential equations. Applicable Analysis. **87** (2), 221-232 (2008)
128. Gliklikh, Yu.E., Zastawniak, T.J.: Solution of the Cauchy problem for the Nelson-Newton equation. In: Borisovich Yu.G. (ed.) Global and Stochastic Analysis (New Developments in Global Analysis series), pp. 41-50. Voronezh University Press, Voronezh (1995)
129. Gliklikh, Yu.E, Zastawniak, T.J.: Solution of the Newton-Nelson equation with random initial data. Appendix to [106], pp. 166-174 (1996)
130. Gliklikh, Yu.E., Zastawniak, T.J.: Existence theorems in Nelson's stochastic mechanics on Riemannian manifolds and space-times of general relativity. In: Rassias Th.M.

(ed.) Nonlinear Mathematical Analysis and Its Applications, pp. 39-58. Hadronic Press, Palm Harbor, Fl. (1998)

131. Gliklikh, Yu.E., Zykov, P.S.: On the two-point boundary value problem for quadratic second order differential equations and inclusions on manifolds. Abstract and Applied Analysis. **2006** (Article ID 30395), 1-9 (2006)
132. Gliklikh, Yu.E., Zykov, P.S.: A two-point boundary value problem on a Lorentz manifold arising in A. Poltorak's concept of reference frame. Journal of nonlinear mathematical physics. **14** (3), 381-389 (2007)
133. Gliklikh, Yu.E., Zykov, P.S.: A fixed point approach to some two-point boundary value problems for differential inclusions on manifolds. Fixed point theory. **8** (2), 263-272 (2007)
134. Godbillon, C.: Géométrie différentielle et mécanique analitique. Hermann, Paris (1969)
135. Goldberg, S.I.: Curvature and homology. Academic Press, N.Y. (1962)
136. Golin, S.: Uncertainty relations in stochastic mechanics. Jour. Math. Physics. **26** (11), 2781-2783 (1985)
137. Grabert H., Hänggi P., Talkner P.: Is quantum mechanics equivalent to a classical stochastic process? Phys. Rev. A. **19** (6), 2440-2445 (1979)
138. Grigor'yan, A.A.: On stochastically complete manifolds. Dokl. AN SSSR **290** (3), 534-537 (1986)
139. Grigor'yan, A.: Analytic and Geometric Background of Recurrence and Non-explosion of Brownian Motion on Riemannian Manifolds. Bulletin (New Series) of the American Mathematical Society. **36** (2), 135-249 (1999)
140. Gromoll, D., Klingenberg, W., Meyer W.: Riemannsche Geometrie im Großen. Springer-Verlag, Berlin et al. (1968)
141. Guerra, F., Morato, L.M.: Quantization of dynamical systems and stochastic control theory. Phys. Reviews, Ser. D., 1774-1786 (1983)
142. Guerra, F., Ruggiero, P.: A note on relativistic Markov processes. Lettere al Nuovo Cimento. **23** (16), 529-534 (1978)
143. Guillemin, V., Sternberg, S.: On the equation of motion of a classical particle in a Yang-Mills field and the principle of general covariance. Hadronic J. **1**, 1-32 (1978)
144. Hartman, P.: Ordinary Differential Equations. John Wiley and Sons, N.Y.-L.-Sydney (1964)
145. He X.: A probabilistic method for Navier-Stokes vorticies. J. Appl. Probab. (2001).- P. 1059-1066.
146. Helgason, S.: Differential geometry and symmetric spaces. Academic Press, New York – London (1962)
147. Hsu, E.P.: Stochastic Analysis on Manifolds (Graduate Studies in Mathematics, Volume 38). AMS, Providens, R.I. (2002)
148. Ikeda, N., Watanabe S.: Stochastic differential equations and diffusion processes. North Holland, Amsterdam-Oxford-New York (1981)
149. Itô, K.: Stochastic differential equations in a differentiable manifold, 1. Nagoya Math. J. **1** (1), 35-47 (1950)
150. Itô K.: Stochastic differential equations in a differentiable manifold, 2. Mem. Coll. Sci. Univ. Kioto Math. **28**, 81-85 (1953)
151. Itô, K.: Stochastic parallel displacement. Lect. Notes Math. **451**, 1-7 (1975)
152. Itô, K.: Extensions of stochastic integral. In: Proc. Intern. Symp. SDE (Kyoto, 1976), pp. 95-109. New York (1978)
153. Jeulin, T., Yor, M.: Filtration des ponts Browniens et equations différentielles stochstiques leneares. Seminaire de Probabilités 24. Lect. Notes Math. **1426**, 227-265 (1990)
154. Jorgensen, E.: Construction of Brownian motion and Ornstein-Uhlenbeck process in a Riemannian manifold on basis of the Gangolli-McKean injection scheme. Z.F.W. **44**, 71-87 (1978)

155. Kamenskiĭ, M., Obukhovskiĭ, V., Zecca, P.: Condensing Multivalued Maps and Semilinear Differential Inclusions in Banach Spaces. Walter de Gruyter, Berlin-New York (2001)
156. Kantorovich, L.V., Akilov, G.P.: Functional analysis. Pergamon Press, Oxford (1982)
157. Kato T. On classical solutions of the two dimensional non-stationary Euler equation. Arch. for Rat. Mech. and Analysis. **25** (3), 188-200 (1967)
158. Kisielewicz, M.: Some remarks on boundary value problem for differential inclusions. Discussiones Mathematicae. DICO. **17** (1,2), 43-50 (1997)
159. Klein, F.: Vorlesungen über höhere Geometrie (dritte Auflage bearbeitet und herausgegaben von W. Blaschke). Julius Springer, Berlin (1926)
160. Knight, F.B.: Essentials of Brownian motion and diffusion. (Mathematical Surveys 18). AMS, Providence, R.I. (1981)
161. Kobayashi, S., Nomizu, K.: Foundations of Differential Geometry. Vol. 1. Vol. 2. Interscience Publishers, New York – London (1963)
162. Korolyuk, V.S., Portenko, N.I., Skorokhod, A.V., et al.: Handbook in probability theory and mathematical statistics. Moscow, Nauka (1985) (Russian)
163. Krein, S.G.: (Ed.) Functional analysis. Wolters-Noordhoff, Groningen (1972)
164. Krein, S.G.: Linear differential equations in Banach space (Transl. Math. Monogr., 29). AMS, Providence, R.I. (1971)
165. Krylov, N.V.: Controlled diffusion processes. Springer, N.Y. (1980)
166. Kryszewski, W.: Homotopy properties of set-valued mappings. Torun University, Torun (1997)
167. Kunita, H.: Stochastic flows and stochastic differential equations. Cambridge Ubiversity Press, Cambridge et al. (1990)
168. Kuratowski, K.: Topology. Vol. 1. Academic Press, N.Y.-L. (1966)
169. Kozlov, V.V.: Global calculus of variations and classical mechanics. Russian Math. Surveys. **40** (2), 61-120 (1985)
170. Ladyzhenskaya, O.A.: The mathematical theory of viscous incompressible flow. Gordon and Breach, New York (1969)
171. Landau, L.D., Lifshitz, E.M.: Mechanics. Butterworth-Heinemann (1976)
172. Lang, S.: Differential manifolds. Springer-Verlag, N.Y. (1985)
173. Leandre, R., Mohammed, S.E.A.: Stochastic functional differential equations on manifolds. Probab. Theory Related Fields. **121** (1), 117-135 (2001)
174. Li, X.-M.: Properties at infinity of diffusion semigroups and stochastic flows via weak uniform covers. J. Potential Analysis. **3**, 339-357 (1994)
175. Liptser, R.S., Shiryaev, A.N.: Statistics of random processes. Springer, N.Y. (1984)
176. Liptser, R.S., Shiryaev, A.N.: The theory of martingales. Kluwer, Dordrecht (1989)
177. Malliavin, P.: Géométrie différentielle stochastique. Montreal University Press, Montreal (1978)
178. Marsden, J., Ebin, D., Fisher, A.: Diffeomorphism groups, hydrodynamics and relativity. In: Proc. 13th Biennal sem. of Canad. Math. Soc., pp. 135-279. Montreal (1970)
179. Meyer, P.A.: Géométrie stochastique sans larmes. Lecture Notes in Mathematics. **850**, 44-102 (1981)
180. Meyer, P.A.: A differential geometric formalism for the Itô calculus. Lecture Notes in Mathematics. **851**, 256-270 (1981)
181. Milnor, J.: Morse Theory. Princeton University Press, Princeton, NJ (1963)
182. Misner, C., Thorne, K., Wheeler, J.: Gravitation. Freeman, N.Y. (1973)
183. Motyl, J.: On the Solution of Stochastic Differential Inclusion. Journal of mathematical analysis and applications. **192**, 117-132 (1995)
184. Myshkis, A.D.: Generalization of the theorem on the stationary point of the dynamical system inside a closed trajectory. Mat. sbornik. **34** (3), 525-540 (1954)
185. Nagasawa, N., Maruyama, T.: An application of time reversal of Markov processes to a problem of population genetics. Advances Appl. Prob. **11**, 457-478 (1979)

186. Nash, J.: The Imbedding Problem for Riemannian manifolds. The Annals of Mathematics. **63** (1), 20-63 (1956)
187. Nelson, E.: Derivation of the Schrödinger equation from Newtonian mechanics. Phys. Reviews. **150** (4), 1079-1085 (1966)
188. Nelson, E.: Dynamical theory of Brownian motion. Princeton University Press, Princeton (1967)
189. Nelson, E.: "Le mouvement brownien relativiste" by J.-P. Caubet (book review). Bull. Amer. Math. Soc. **84** (10), 121-124 (1978)
190. Nelson, E.: Quantum fluctuations. Princeton University Press, Princeton (1985)
191. Nomizu K., Ozeki H.: The existence of complete Riemannian metrics. Proc. Amer. Math. Soc. **12**, 889-891 (1961)
192. Norris, J.R.: Construction of diffusions with a given density. In: Stochastic calculus in applications (Proceedings of the Cambridge Symposium 1987), pp. 69-77. Longman, Harlow (1988)
193. Novikov, S.P.: Hamiltonian formalism and multi-valued analog of Morse theory. Russian Math. Surveys. **37** (5), 3-49 (1982)
194. Parthasarathy, K.R.: Introduction to Probability and Measure. Springer-Verlag, New York (1978)
195. Penot, J.-P.: Weak topology on functional manifolds. In: Global analysis and its applications. Vol. 3, pp. 75-84. International atomic energy agency, Vienna (1974)
196. Poltorak, A.: On the covariant theory of gravitation. In: 9th Interna-tional Conference on General Relativity and Gravitation. Abstracts. Vol. 2, pp. 516. Jena (1980) (Available at arXiv:gr-qc http://arxiv.org/abs/gr-qc/0403050v3 24 March 2004)
197. Poltorak, A.: On the Energy Problem in General Relativity. In: Bertotti B., de Felice F., Pascolini A. (eds.) 10-th International Conference on General Relativity and Gravitation, Padova, Contributed Papers, Vol.1, pp. 609 (1983) (Enlarged version is available at arXiv:gr-qc http://arxiv.org/abs/gr-qc/0403107v1 25 March 2004)
198. Poltorak, A.: Gravity as nonmetricity. General relativity in metric-afine space (L_n, g). (Available at arXive:gr-qc http://arxiv.org/abs/gr-qc/0407060v2 30 July 2004)
199. Rodichev, V.I.: Evolution of the notion of reference frame and Einstein's program. In: Einsteininan miscellany, pp. 286–334. Nauka, Moscow (1974)
200. Sachs, R.K., Wu, H.: General Relativity for Mathematicans. Springer-Verlag, N.Y. (1977)
201. Salomonson, P., Skagerstam, B.-C. and Winnberg, J.- O.: Equations of motion of a Yang-Mills particle. Physical Review D. **16** (8), 2581-2585 (1977)
202. Schutz, B.F.: Geometrical methods of mathematical physics. Cambridge University Press, Cambridge (1982)
203. Schwartz, L.: Processus de Markov et desingration regulieres. Annales de l'Institut Fourier de l'Université de Grenoble. **XXVII** (3), 211-277 (1977)
204. Schwartz, L.: Géométrie différentielle du 2ème ordre, semimartingales et équations différentielles stochastique sur une variété différentielle. Lecture Notes in Mathematics. **921**, 1-150 (1982)
205. Schwartz, L.: Semimartingales and their stochastic calculus on manifolds. Montreal University Press, Montreal (1984)
206. Schwartz, L.: Le semigroupe d'une diffusion en liaison avec les trajectories. Séminair de Probabilités XXIII. Lect. Notes Math. **1372**, 326-342 (1989)
207. L. I. Sedov: Mechanics of continuous medium. World Scientific, River Edge, NJ, **2**, 1997
208. Shiryaev, A.N.: Probability. Springer, New York Berlin Heidelberg (1984)
209. Shnirel'man, A.I.: On the geometry of the diffeomorphism group and the dynamics of a perfect incompressible fluid. Matem. Sbornik **128** (1)(9), 82-109 (1985)
210. Smolentsev, N.K.: On the Mopertiu principle. Sibirian Math. J. **20** (5), 1092-1098 (1979)
211. Smolentsev, N.K.: Integrals of flows of the ideal barotropic fluid. Sibirian Math. Journal. **23** (1), 205-208 (1982)

212. Sternberg, S.: Lectures on Differential Geometry. Prentice Hall, Inc. Englewood Cliffs, N.J. (1964)
213. Sternberg, S.: Minimal coupling and symplectic mechanics of a classical particle in the presence of a Yang-Mills field. Proc. Nat. Acad. Sci. USA. **74** (12), 5253-5254 (1977)
214. Sternberg, S.: On the role of field theories in our physical conception of geometry. Differential Geometric Methods in Mathematical Physics II. Lecture Notes in Mathematics. **676**, 1-80 (1978)
215. Sternberg, S., Ungar, T.: Classical and prequantized mechanics without Lagrangians or Hamiltonians. Hadronic Journal. **1**, 33-76 (1978)
216. Stratonovich, R.L.: Conditional Markov processes and their application to optimal control theory. Moscow University Press, Moscow (1966) (Russian)
217. Synge, J.L.: Hodographs of general dynamical systems. Trans. Royal Soc. Canada. Ser. 3. **25**, 121-136 (1931)
218. Synge, J.L.: Tensorial methods in dynamics. University of Toronto, Toronto (1936)
219. Temam, R.: On the Euler equations of incompressible perfect fluids. J. Funct. Anal. **20**, 32-43 (1975)
220. Temam, R.: Local existence of C^∞-solutions of the Euler equations of incompressible perfect fluids. Lecture Notes in Mathematics. **565**, 184-194 (1976)
221. Temam, R.: Remarks on the Euler equations. In: Proc. Sym. Pure Math. Vol. 45, Part 2, pp. 429-430. AMS, Providence R.I. (1986)
222. Tolstonogov, A.A.: Differential inclusions in Banach space. Nauka, Novosibirsk (1986) (Russian)
223. Trenogin, V.A.: Analyse Fonctionelle. Mir, Moscow (1985).
224. Vershik, A.M.: Classical and non-classical dynamics with constraints. Lecture Notes in Mathematics. **1108**, 278-301 (1984)
225. Vershik, A.M., Gershkovich, V.Ya.: Non-holonomic problems and the theory of distributions. Acta Applicandae, **12**, 181-209 (1987)
226. Vershik, A.M., Gershkovich, V.Ya.: Non-holonomic dynamical systems: geometry of distributions and variational problems. In: Encylopedia of Mathematical Sciences, vol. 16, pp. 1-81. Springer, New York Berlin Heidelberg (1994)
227. Vujičič, V.A.: The absolute integral of a tensor. Publ. Inst. Math. Beograd. **10** (24), 199-202 (1970) (Russian).
228. Vujičič, V.A.: Covariant integrals in mechanics. Applied Mathematics and Mechanics. **40**, 346-351 (1976) (Russian).
229. Vujičić, V.A.: Kovarijantna dinamika. Matematički Institut, Beograd (1981)
230. Wolibner, W.: Un theoreme sur l'existence du mouvement plan d'un fluide parfait, homogene, incompressible, pendant un temps infinie longue. Math.Z. **37**, 698-726 (1933)
231. Wong, S.K.: Field and particle equations for classical Yang-Mills field and particles with isotopic spin. Il Nuovo Cimento. **65A** (4), 689-694 (1970)
232. Yakovlev, E.I.: On the solvability of two-end problems for some ordinary second order differential equations on manifolds. In: Baku International Topological Conference. Abstracts. Part 2, pp. 361. Baku (1987) (Russian)
233. Yasue, K.: Quantum mechanics and stochastic control theory. J. Math. Phys. **22** (5), 1010-1020 (1981)
234. Yau, S.-T.: On the heat kernel of a complete Riemannian manifold. J. Math. Pures et Appl. **57** (2), 191-201 (1978)
235. Yosida, Y.: Functional Analysis. Springer-Verlag, Berlin et al. (1965)
236. Yudovich, V.I.: The two-dimensional non-stationary problem on flowing ideal incompressible fluid through a given domain. Mat. Sbornik. **64** (4), 562-568 (1964)
237. Zastawniak, T.: A relativistic version of Nelson's stochastic mechanics. Europhys. Lett. **13**, 13-17 (1990)

238. Zastawniak, T.: Markov diffusion in relativistic stochastic mechanics. In: Proceedings of Swansea Conference on Stochastic Mechanics, 1990, pp. 280-297. World Scientific, Singapore (1992)
239. Zhelobenko, D.P.: Compact Lie groups and their representations. Amer. Math. Soc., Providence, RI (1973)
240. Zheng, W.: Tightness results for laws of diffusion processes. Application to stochastic mechanics. Ann. Inst. Henri Poincare. **B21**, 103-124 (1985)
241. Zykov, P.S.: A condition for solvability of the two-point boundary value problem for second order differential equations and inclusions on manifolds. Applicable Analysis **88** (6), 895-902 (2009)

Index

Y.E. Gliklikh, *Global and Stochastic Analysis with Applications to Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-0-85729-163-9,