Theoretical and Mathematical Physics

Gerd Rudolph
Matthias Schmidt

Differential Geometry and Mathematical Physics

Part I. Manifolds, Lie Groups and Hamiltonian Systems

Springer

Differential Geometry and 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 further volumes:
www.springer.com/series/720

Gerd Rudolph • Matthias Schmidt

Differential Geometry and Mathematical Physics

Part I. Manifolds, Lie Groups and Hamiltonian Systems

Springer

Gerd Rudolph
Institute for Theoretical Physics
University of Leipzig
Leipzig, Germany

Matthias Schmidt
Institute for Theoretical Physics
University of Leipzig
Leipzig, Germany

ISSN 1864-5879 ISSN 1864-5887 (electronic)
Theoretical and Mathematical Physics
ISBN 978-94-007-5344-0 ISBN 978-94-007-5345-7 (eBook)
DOI 10.1007/978-94-007-5345-7
Springer Dordrecht Heidelberg New York London

Library of Congress Control Number: 2012951155

Printed on acid-free paper

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

Introduction

This book is the first of two volumes on differential geometry and mathematical physics. It is the result of our teaching these subjects at the University of Leipzig over the last few decades to students of physics and of mathematics. The present volume is devoted to manifolds, Lie groups and the theory of Hamiltonian systems. The second volume will deal with fibre bundles, topology and gauge field theory, including aspects of the theory of gravity.

While the aim and scope of differential geometry are somewhat well defined, it is, perhaps, less clear what we possibly mean by mathematical physics. Historically, this term was rather imprecise and so is still nowadays. Indeed, its interpretation depends on culture and context. In our understanding, mathematical physics is the area where theoretical physics and pure mathematics meet, stimulate and fertilize each other. On the one hand, this interplay leads to a deeper structural understanding of theoretical physics and to new results obtained with new mathematical methods. On the other hand, it stimulates the development of old and new branches in mathematics. Thus, in our understanding, it is impossible to draw a precise borderline between theoretical physics and pure mathematics. Over the last decades it sometimes happened that the solution of a problem posed by physicists had an even larger impact on the development of mathematics than on the field of physics from where it arose. There is a number of texts where the status and the role of mathematical physics is discussed, see e.g. the papers of Greenberg [112], Faddeev [87] and Jaffe and Quinn [151],[1] as well as the classical contributions of Poincaré [241] and Hilbert [128].

There is no doubt that, over the last few centuries, the interrelation between physics and geometry has been especially tight and fruitful. In particular, this interaction has stimulated the development of modern differential geometry. In this complex process, which we cannot describe here,[2] the development of the notion of manifold was of great importance. The conceptual definition of this notion was

[1] This very interesting and somewhat provocative article stimulated a lot of responses by leading scientists, see [27].

[2] See [264–266] and [281, 282] for a detailed discussion of these historical aspects.

presented by Riemann in his famous Habilitationsvortrag in Göttingen in the year 1854, see [251].[3] Riemann developed a general geometry (nowadays called Riemannian geometry), which included Euclidean as well as non-Euclidean geometry as originated by Bolyai, Gauß and Lobachevsky. From Riemann's presentation it is clear that a deeper understanding of the nature of physical space was one of the main motivations for his studies. In his understanding, space works on bodies and bodies have an influence on space. This idea made Riemann depart from the metaphysical attitude towards space as a given unchangeable entity and pass to a modern field theoretical point of view. He even suggested that the metric might be determined by the physical masses. Thus, on a rather philosophical level, he made a step towards the conceptual foundations of Einstein's theory of gravity, which came 60 years later. At the same time, he created the mathematical framework for this theory.

In the following years, a number of great mathematicians contributed to the field, but it was Poincaré who brought the concept of manifold to its modern form.[4] As he said himself, he was led to this concept by his previous investigations on the theory of differential equations and its applications to dynamics, in particular, of the n-body problem. On the one hand, on the basis of this abstract manifold concept, he laid the foundations of modern algebraic topology. On the other hand, he continued his studies on dynamical systems with emphasis on their global, qualitative behaviour, on the way creating a lot of tools which nowadays still play an important role. At the same time, he provided a geometrization of the formalism of analytical mechanics as developed by Lagrange,[5] Hamilton,[6] Jacobi, Liouville and Poisson. Instead of formulating dynamics in terms of local coordinates in Euclidean space, he viewed it as a global system described by a Hamiltonian vector field on the phase space manifold. Thus, modern symplectic geometry was born.[7]

In the twentieth century, the interaction between physics and geometry continued to be strong and successful. Of course, first of all, we should mention Einstein's theory of gravity, which is based on the discovery that gravity is a geometric property of spacetime and that spacetime is curved by matter. Starting from the nineteen-fifties, all other fundamental forces were geometrized in a similar spirit leading to modern gauge theory. The necessary mathematical foundations, including the general theory

---

[3] There were, of course, precursors. In particular, Cauchy and Gauß should be mentioned. Gauß even used the term manifold, but in his understanding it was restricted to affine subspaces of an n-dimensional vector space.

[4] See the famous Analysis situs [239] from the year 1895, together with five subsequent complements within the following nine years. Actually, Poincaré gave two definitions of a manifold: a manifold as a level set and a manifold as given by an atlas of local charts.

[5] The origin of symplectic geometry dates back to Lagrange's early work on celestial mechanics, see [177] and [308] for a detailed discussion by Weinstein.

[6] See [120–122]. Hamilton was led to the formulation of dynamics in terms of a system of first order differential equations for a general mechanical system through his studies in optics.

[7] The word symplectic was invented by Weyl [312] to give a name to the group of linear transformations, preserving a non-degenerate, skew-symmetric bilinear form, and the term symplectic geometry was proposed by Siegel, see [272].

of fibre bundles and connections, were developed by É. Cartan, Koszul, Ehresmann and Chern. Once these geometrical formulations of the fundamental forces had been found, another fascinating interaction of geometry and physics took place. In the mid-seventies physicists insisted on classifying the solutions (of a certain type) of the Yang-Mills field equations of classical gauge theory. This problem was finally solved by leading mathematicians and the techniques developed on the way, in turn, led to fundamentally new and deep insights in topology. In particular, exotic topological structures on Euclidean four-space were found. We should also add that, starting from the work of Kaluza, Klein, Einstein and Weyl in the nineteen-twenties up until the present, there has been much effort devoted to searching for an ultimate geometrical model unifying all fundamental forces. This gave another strong impetus to the development of modern differential geometry and related fields. Some of the aspects just mentioned will be discussed in volume 2 of this book.

To finish this brief historical introduction, we should make two further remarks. Firstly, it should be mentioned that in the process described above the concept of symmetry played a fundamental role. The creation of the mathematical foundations of this concept dates back to Lie, who in the eighteen-seventies developed a general theory of transformations.[8] Lie was influenced by the work of Galois on symmetries of polynomial equations, by the work of Jacobi on partial differential equations, and by Klein, whose aim was to unify geometry and group theory. The theory was essentially pushed forward by Killing and É. Cartan, who classified semisimple Lie algebras and developed their representation theory. The early period closes with the contributions by Weyl, who created the representation theory of semisimple Lie groups. It was also Weyl who first applied concepts of group theory to quantum mechanics. It goes without saying that the general theory of fibre bundles and connections and, consequently, also the theory of gauge fields, heavily rests on Lie group theory.

Secondly, over the last few decades, it has become more and more clear that symplectic geometry plays a special role. This is not only due to the fact that there is merely a lot of applications of symplectic techniques in many areas of mathematics and in physics. There is something more: a phenomenon which Arnold called symplectization, see [22], [111] and also [308]. Indeed, there seems to be growing evidence that many concepts, constructions and results from different branches of mathematics and mathematical physics (like the theory of partial differential equations, the calculus of variations or the theory of group representations) can be recast in symplectic terms, finding in this way their ultimate ground. The theory of Hamiltonian systems in its modern form is of course still one of the most prominent examples. Via Hamilton-Jacobi theory there is a close link to the theory of linear partial differential equations. Here, representing a differential operator on a manifold by its symbol on the cotangent bundle and seeking solutions in terms of Lagrangian immersions and geometrical objects living on them, one arrives at a symplectized

[8] See [182–184]. For a historical overview on Lie group theory we refer to [50] and [126].

theory of first order partial differential equations.[9] In a similar spirit and in close relation, large parts of the theory of singularities have been symplectized. Another beautiful example is provided by the orbit method of Kirillov, Kostant and Souriau, which constitutes one of the cornerstones of geometric quantization. For a large class of Lie groups, this method yields a bijective correspondence between irreducible unitary representations of a Lie group and transitive symplectic actions of this group.

We conclude with a few remarks on the structure and the contents of this volume. It contains three building blocks, each consisting of four chapters. In the first four chapters, we present the calculus on manifolds. The next four chapters are devoted to the theory of Lie groups and Lie group actions and to an introduction to linear symplectic algebra and symplectic geometry. These chapters constitute the link between the abstract calculus and the theory of finite-dimensional Hamiltonian systems, which we develop in the final four chapters. There, we had to make a reasonable selection of the topics to be presented. It is probably fair to say that our choice of material was made more from a physicist's point of view,[10] thus, putting emphasis on the concepts of symmetry and integrability and on Hamilton-Jacobi theory. At the same time, this means that we had to exclude a lot of interesting topics like, for instance, equivariant Hamiltonian dynamics or variational methods. Since each chapter has its own introduction, here we omit a detailed description of the contents.

We assume that the reader is familiar with elementary algebra and calculus, as well as with the basics of classical mechanics. Some knowledge in classical electrodynamics and in thermodynamics as well as in elementary set topology will be helpful. The book is self-contained, that is, starting with the theory of differentiable manifolds, it guides the reader to a number of advanced topics in the theory of Hamiltonian systems. At some points, we touch on current research. It is our strong belief that without detailed case studies a deep understanding of the abstract material can be hardly achieved. Thus, we have included many worked examples, some of them are taken up repeatedly. Moreover, at the end of almost every section the reader will find a number of exercises.

---

[9] This is due to Maslov [197] and Hörmander [141], see Chap. 12 of this volume. In this context, the symplectization of Morse theory plays a basic role, see Sects. 8.9 and 12.4.

[10] In particular, this means that we do not go into advanced topics related to symplectic topology. However, at some points we touch on it. For a thorough presentation of symplectic topology we refer to the textbooks of Hofer and Zehnder [139] and of McDuff and Salamon [206].

Acknowledgements

It is a pleasure to thank Szymon Charzyński, Bernd Crell, Jochen Dittmann, Elisabeth Fischer, Christian Fleischhack, Johannes Huebschmann, Peter Jarvis, Jerzy Kijowski, Rainer Matthes, Olaf Richter†, Matthias Schwarz, Torsten Tok, Igor P. Volobuev and Raimar Wulkenhaar for fruitful discussions. We are especially grateful to Christian Fleischhack, Peter Jarvis and Rainer Matthes for reading parts of the manuscript. We also thank our students Michael Kath and Sandro Wenzel for their support in writing the lecture notes from which this volume evolved. Over the years, we received numerous valuable remarks on these notes from students, notably from Jörn Boehnke, Alexander Hertsch, André Jäschke, Adam Reichold and Axel Schüler.

Contents

Chapter 1
Differentiable Manifolds

In this chapter, we introduce the basic notions of the theory of manifolds. We start with the very notion of a manifold, illustrate it by a number of examples and discuss the class of examples provided by level sets in some detail. Next, we carry over the notions of differentiability, the concept of tangent space and the notion of derivative of a mapping from classical calculus to the theory of manifolds. In this context, we also generalize the basic theorems of classical calculus to the case of manifolds. Finally, we discuss some more advanced topics needed later on: the notion of submanifold and the concept of transversality. Since the notion of a submanifold is quite subtle, we treat this subject in detail. In our terminology, a submanifold is defined by an injective immersion. There are two important special classes of submanifolds showing up in various applications. They are called embedded and initial, respectively. Throughout the text, the reader will find a large number of illustrative examples.

1.1 Basic Notions and Examples

Manifolds are topological spaces which locally look like $\mathbb{R}^n$. Therefore, they allow for an extension of the notions of classical calculus. For the topological notions used in this section we refer the reader to the standard literature, e.g., [53], [55], [199] or [267].

In the sequel, we will use the following notation. An element $\mathbf{x} \in \mathbb{R}^n$ is an n-tuple written as $\mathbf{x} = (x_1, \dots, x_n)$. The Euclidean scalar product on $\mathbb{R}^n$ is denoted by

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i$$

and the corresponding norm is denoted by $\| \cdot \|$. For the standard basis in $\mathbb{R}^n$ we write $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$. The dual basis $\{\mathbf{e}^{*1}, \dots, \mathbf{e}^{*n}\}$ in $\mathbb{R}^{n*}$ is defined by $\mathbf{e}^{*i}(\mathbf{e}_j) = \delta^i{}_j$. For $\mathbf{x} \in \mathbb{R}^n$, its coefficients in the standard basis are given by

$$x^i = \mathbf{e}^{*i}(\mathbf{x}).$$

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*,
Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7_1,

Thus, we have $\mathbf{x} = \sum_{i=1}^{n} x^i \mathbf{e}_i$. Of course, numerically, the numbers x^i and x_i coincide. If there is no danger of confusion, we will use the Einstein summation convention, that is, we will also write $\mathbf{x} = x^i \mathbf{e}_i$.

Definition 1.1.1 (Topological manifold) A topological space M is called a topological manifold if it is Hausdorff, second countable and locally homeomorphic to $\mathbb{R}^n$ for some fixed $n \in \mathbb{N}$. This means that for every $m \in M$, there exists an open neighbourhood U of m in M and a mapping $\kappa : U \to \mathbb{R}^n$ such that $\kappa(U)$ is open in $\mathbb{R}^n$ and κ is a homeomorphism onto its image. The pair (U, κ) is called a local chart for M. A family $\mathscr{A} = \{(U_\alpha, \kappa_\alpha) : \alpha \in A\}$ of local charts satisfying $\bigcup_{\alpha \in A} U_\alpha = M$ is called an atlas for M. The number n is called the dimension of M.

Let (U, κ) be a local chart on M. If $m \in U$, we say that (U, κ) is a chart at m. The functions

$$\kappa^i := \mathbf{e}^{*i} \circ \kappa : U \to \mathbb{R}, \quad 1 \leq i \leq n,$$

define a system of local coordinates on U. The numbers $\kappa^i(m)$ are called the local coordinates of m in the chart (U, κ). In particular, $(\mathbb{R}^n, \mathrm{id}_{\mathbb{R}^n})$ is a global chart for $\mathbb{R}^n$, that is, the identical mapping endows $\mathbb{R}^n$ with the structure of a topological manifold. The corresponding coordinates are $\kappa^i(\mathbf{x}) = x^i$.

Remark 1.1.2

1. Since $\mathbb{R}^n$ is not homeomorphic to $\mathbb{R}^m$ for $n \neq m$, the dimension of a topological manifold is unique.
2. As a consequence of being second countable, topological manifolds can have at most countably many connected components. Moreover, every point possesses a countable neighbourhood basis, so that one can use sequences to test topological properties like continuity of mappings or closedness of subsets.
3. As a consequence of being locally homeomorphic to $\mathbb{R}^n$, topological manifolds inherit all the local properties of $\mathbb{R}^n$. That is, they are locally compact (every point has a compact neighbourhood), locally connected (every point admits a neighbourhood basis of connected open sets), etc.

Let (U_1, κ_1) and (U_2, κ_2) be local charts on M. If $U_1 \cap U_2 \neq \varnothing$, the mapping

$$\kappa_2 \circ \kappa_1^{-1} : \mathbb{R}^n \supset \kappa_1(U_1 \cap U_2) \to \kappa_2(U_1 \cap U_2) \subset \mathbb{R}^n,$$

pictured in Fig. 1.1, and its inverse $\kappa_1 \circ \kappa_2^{-1}$ are called the transition mappings of (U_1, κ_1) and (U_2, κ_2). The transition mappings are homeomorphisms. Since $\kappa_1(U_1 \cap U_2)$ and $\kappa_2(U_1 \cap U_2)$ are open subsets of $\mathbb{R}^n$ it makes sense to ask whether the transition mappings are differentiable.

Definition 1.1.3 Let M be a topological manifold and let k be a nonnegative integer or ∞.

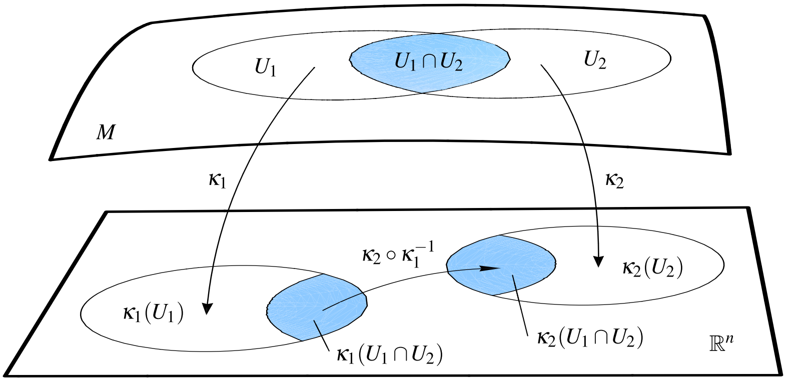


Fig. 1.1 Two local charts and one of their transition mappings

1. Two local charts (U_1, κ_1) and (U_2, κ_2) are compatible of class C^k if either $U_1 \cap U_2$ is empty or their transition mappings $\kappa_1 \circ \kappa_2^{-1}$ and $\kappa_2 \circ \kappa_1^{-1}$ are of class C^k.
2. An atlas $\mathscr{A}$ for M is of class C^k if any two charts are compatible of class C^k.
3. Two atlases $\mathscr{A}_1$ and $\mathscr{A}_2$ of class C^k are equivalent, if any chart of $\mathscr{A}_1$ is compatible of class C^k with any chart of $\mathscr{A}_2$; equivalently, if their union is an atlas of class C^k.

Point 3 defines an equivalence relation on the class of all C^k-atlases for M.

Definition 1.1.4 (Differentiable manifold) An equivalence class of atlases of class C^k for a topological manifold M is called a differentiable structure of class C^k, or just a C^k-structure, on M. A topological manifold together with a differentiable structure of class C^k is called a differentiable manifold of class C^k, or just a C^k-manifold.

By definition, a C^0-manifold is just a topological manifold. The property to be of class C^∞ will be referred to as smooth. While in Chaps. 1 and 2 the general C^k-case is treated, starting from Chap. 3, everything will be assumed to be smooth, with occasional exceptions if necessary.

Remark 1.1.5

1. Let $l < k$. It is obvious that if two local charts are compatible of class C^k, they are also compatible of class C^l. Consequently, an atlas of class C^k is also an atlas of class C^l. The reader may convince himself that a C^k-manifold can be viewed as a C^l-manifold if necessary.
2. By analogy, one defines the notion of real and complex analytic structure and manifold. For a real analytic structure, the transition mappings are required to be real analytic. For a complex analytic structure, the local charts are assumed to take values in $\mathbb{C}^n$ for some n and the transition mappings are assumed to be holomorphic.

To define a differentiable structure it suffices to construct an atlas. Since the union of arbitrarily many equivalent C^k-atlases is a C^k-atlas, in each equivalence class there is a unique maximal atlas. Since for any chart (U, κ) and any open subset $V \subset U$, $(V, \kappa_{\restriction V})$ is a local chart and since all these induced local charts are compatible of class C^k with one another, all of them are contained in the maximal atlas.

Example 1.1.6 (Linear spaces) For $\mathbb{R}^n$ itself, the single chart $(\mathbb{R}^n, \mathrm{id}_{\mathbb{R}^n})$ provides an atlas of class C^∞. The differentiable structure defined by this atlas is called the standard smooth structure of $\mathbb{R}^n$ and the corresponding coordinates are called the standard coordinates on $\mathbb{R}^n$. Similarly, for an arbitrary open subset $U \subset \mathbb{R}^n$, the natural inclusion mapping $\iota_U : U \to \mathbb{R}^n$ provides a single chart and hence defines a smooth structure. It is clear that all of this carries over to arbitrary finite-dimensional real vector spaces.

Example 1.1.7 (Chart domains) Let M be a manifold of class C^k and let (U, κ) be a local chart on M. Then (U, κ) is a global chart on U, thus providing a smooth atlas on U. Hence, independently of the differentiability class of M, U together with the atlas $\{(U, \kappa)\}$ is a smooth manifold. It has the same dimension as M.

Example 1.1.8 (Open subsets) Let M be a manifold of class C^k with atlas $\mathscr{A}$ and let $W \subset M$ be an open subset. If (U, κ) is a local chart on M with $U \cap W \neq \varnothing$, then $U \cap W$ is open in W, $\kappa(U \cap W)$ is open in $\mathbb{R}^n$ and $\kappa_{\restriction U \cap W} : U \cap W \to \kappa(U \cap W)$ is a homeomorphism. Therefore,

$$\mathscr{A}_W := \big\{(U \cap W, \kappa_{\restriction U \cap W}) : (U, \kappa) \in \mathscr{A}, U \cap W \neq \varnothing\big\}$$

is an atlas for W. Since the transition mappings of this atlas are obtained from those of $\mathscr{A}$ by restriction to open subsets, they are of class C^k again. Thus, the atlas $\mathscr{A}_W$ defines a differentiable structure of class C^k on W and hence W is a differentiable manifold of the same class and the same dimension as M.

Example 1.1.9 (Spheres) The sphere S^n is the set of solutions of the equation $\|\mathbf{x}\|^2 = 1$ in the variable $\mathbf{x} \in \mathbb{R}^{n+1}$. In the relative topology induced from $\mathbb{R}^{n+1}$, S^n is Hausdorff and second countable. We construct an atlas by means of stereographic projection. Fix an arbitrary point $\mathbf{e}_0 \in \mathrm{S}^n$, put $U_\pm := \mathrm{S}^n \setminus \{\pm\mathbf{e}_0\}$ and define mappings $\kappa_\pm : U_\pm \to \mathbb{R}^n$ by

$$\kappa_\pm(\mathbf{x}) := \frac{\mathbf{x} - (\mathbf{x} \cdot \mathbf{e}_0)\mathbf{e}_0}{1 \mp \mathbf{x} \cdot \mathbf{e}_0}, \quad \mathbf{x} \in U_\pm,$$

where we have identified $\mathbb{R}^n$ with the hyperplane of $\mathbb{R}^{n+1}$ orthogonal to $\mathbf{e}_0$, see Fig. 1.2. The mappings $\kappa_\pm$ are obviously bijective, continuous and open, hence homeomorphisms. Since $U_+ \cup U_- = \mathrm{S}^n$, the sphere S^n is a topological manifold. The transition mapping $\kappa_- \circ \kappa_+^{-1}$ maps $\kappa_+(U_+ \cap U_-) = \mathbb{R}^n \setminus \{0\}$ to $\kappa_-(U_+ \cap U_-) =$

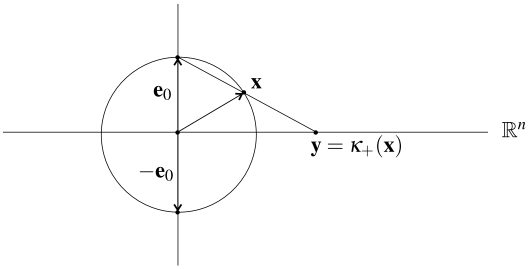


Fig. 1.2 Stereographic projection

$\mathbb{R}^n \setminus \{0\}$. There holds

$$\kappa_-(\mathbf{x}) = \kappa_+(\mathbf{x}) \frac{1 - \mathbf{x} \cdot \mathbf{e}_0}{1 + \mathbf{x} \cdot \mathbf{e}_0} = \frac{\kappa_+(\mathbf{x})}{\|\kappa_+(\mathbf{x})\|^2},$$

and hence

$$\kappa_- \circ \kappa_+^{-1}(\mathbf{y}) = \frac{\mathbf{y}}{\|\mathbf{y}\|^2}.$$

This mapping is of class C^∞. Since $\kappa_+ \circ \kappa_-^{-1} = \kappa_- \circ \kappa_+^{-1}$, the local charts (U_+, κ_+) and (U_-, κ_-) define a smooth structure on S^n.

Remark 1.1.10 (Differentiable structure induced by a family of mappings) For the topology of S^n one may also choose the initial topology[1] induced by the mappings $\kappa_\pm : \mathrm{S}^n \to \mathbb{R}^n$. This has the advantage that κ_+ and κ_- are automatically homeomorphisms then, because the transition mappings $\kappa_+ \circ \kappa_-^{-1}$ and $\kappa_- \circ \kappa_+^{-1}$ are of class C^∞ and hence in particular continuous. It is not hard to show that the initial topology induced by $\kappa^\pm$ coincides with the relative topology[2] induced from $\mathbb{R}^{n+1}$ (Exercise 1.1.1).

More generally, let M be a set and let $n \in \mathbb{N}$. Assume that we are given a countable covering $\{U_\alpha\}$ of M such that for every α there exists an injective mapping $\kappa_\alpha : U_\alpha \to \mathbb{R}^n$ whose image is open. Assume further that the transition mappings are differentiable of class C^k. (This makes sense, because the assumptions imply that their domains are open subsets of $\mathbb{R}^n$.) Equip M with the initial topology defined by the family of mappings $\{(U_\alpha, \kappa_\alpha)\}$. This topology is second countable, because so is $\mathbb{R}^n$ and the family is countable. This topology is also Hausdorff: let $m_a, m_b \in M$, $m_a \neq m_b$. If there is an α such that $m_a, m_b \in U_\alpha$, the assertion is obvious. Otherwise, let $m_a \in U_\alpha$ and $m_b \in U_\beta$. Since U_α is open, there exists a neighbourhood V_a of m_a in M whose closure $\overline{V_a}$ in M is contained in U_α. Since $m_b \notin U_\alpha$, $U_\beta \setminus \overline{V_a}$ is a neighbourhood of m_b, and it does not intersect V_a. This yields the assertion.

[1]The coarsest topology such that both κ_+ and κ_- are continuous.

[2]The coarsest topology such that the natural inclusion mapping $\mathrm{S}^n \to \mathbb{R}^{n+1}$ is continuous.

Continuity of the transition mappings ensures that the $(U_\alpha, \kappa_\alpha)$ are local charts on M. Since the U_α cover M and since the transition mappings are of class C^k, these charts establish a C^k-atlas on M. The corresponding C^k-structure is said to be induced by the family of mappings $\{(U_\alpha, \kappa_\alpha)\}$.

Example 1.1.11 (Polar and spherical coordinates) The polar coordinates r, ϕ in the plane and the spherical coordinates r, ϑ, ϕ in $\mathbb{R}^3$ are curvilinear coordinates which are often used in physical problems with rotational symmetry in order to simplify calculations. The polar coordinate chart is defined as the inverse of the mapping

$$\mathbb{R}_+ \times (-\pi, \pi) \to \mathbb{R}^2, \qquad (r, \phi) \mapsto (r\cos\phi, r\sin\phi).$$

Hence, the chart domain is $U = \mathbb{R}^2 \setminus \{\lambda \mathbf{e}_1 : \lambda \leq 0\}$. This chart is smoothly compatible with the standard smooth structure on $\mathbb{R}^2$. By restriction to $\mathrm{S}^1 \subset \mathbb{R}^2$, the coordinate function ϕ yields a local chart which is smoothly compatible with the smooth structure discussed in Example 1.1.9. The spherical coordinate chart is defined as the inverse of the mapping

$$\mathbb{R}_+ \times (0, \pi) \times (-\pi, \pi) \to \mathbb{R}^2, \qquad (r, \vartheta, \phi) \mapsto (r\sin\vartheta\cos\phi, r\sin\vartheta\sin\phi, r\cos\vartheta).$$

Hence, the chart domain is $U = \mathbb{R}^3 \setminus \{\lambda \mathbf{e}_1 + \mu \mathbf{e}_3 : \lambda \leq 0, \mu \in \mathbb{R}\}$. This chart is smoothly compatible with the standard smooth structure on $\mathbb{R}^3$. By restriction to $\mathrm{S}^2 \subset \mathbb{R}^3$, the two coordinate functions ϕ and ϑ yield a local chart on the sphere S^2 which is smoothly compatible with the smooth structure discussed in Example 1.1.9.

Example 1.1.12 (Möbius strip) Let M be the topological quotient of the open subset $\mathbb{R} \times (-1, 1) \subset \mathbb{R}^2$ by the equivalence relation

$$(s_1, t_1) \sim (s_2, t_2) \quad \text{iff} \quad (s_2, t_2) = \big(s_1 + 2\pi k, (-1)^k t_1\big) \quad \text{for some } k \in \mathbb{Z}.$$

M is called the Möbius strip. Let $p : \mathbb{R} \times (-1, 1) \to M$ denote the natural projection. As the quotient of a second countable space, M is second countable. It is Hausdorff: for $m_1, m_2 \in M$, define

$$d(m_1, m_2) = \inf\Big\{\sqrt{(s_2 - s_1)^2 + (t_2 - t_1)^2} : (s_i, t_i) \in p^{-1}(m_i), i = 1, 2\Big\} \tag{1.1.1}$$

and show that d is a metric on M, compatible with the quotient topology (Exercise 1.1.3). To construct an atlas, we show that p is open. For every $k \in \mathbb{Z}$, the mapping

$$\varphi_k : \mathbb{R} \times (-1, 1) \to \mathbb{R} \times (-1, 1), \qquad \varphi_k(s, t) := \big(s + 2\pi k, (-1)^k t\big),$$

is a homeomorphism. If $O \subset \mathbb{R} \times (-1, 1)$ is open, then $\varphi_k(O)$ and hence

$$p^{-1}\big(p(O)\big) = \bigcup_{k \in \mathbb{Z}} \varphi_k(O)$$

is open. Therefore, $p(O)$ is open. In particular, the open subsets

$$V_1 := (-\pi, \pi) \times (-1, 1), \qquad V_2 := (0, 2\pi) \times (-1, 1)$$

of $\mathbb{R} \times (-1, 1)$ project to open subsets $U_i := p(V_i)$ of M. Since the restrictions of p to V_1 and V_2 are injective, the mappings $p_{\restriction V_i} : V_i \to U_i$ are homeomorphisms. Then, (U_i, κ_i) with $\kappa_i := (p_{\restriction V_i})^{-1}$ are local charts on M. Due to $M = U_1 \cup U_2$, they form an atlas, thus turning M into a topological manifold.

Finally, we check differentiability of the transition mapping $\kappa_2 \circ \kappa_1^{-1}$. It maps a point (s, t) in $\kappa_1(U_1 \cap U_2) = \{(s, t) \in V_1 : s \neq 0\}$ to the unique representative of the class of (s, t) in $\kappa_2(U_1 \cap U_2) = \{(s, t) \in V_2 : s \neq \pi\}$. Thus, on the connected component $s < 0$ of $\kappa_1(U_1 \cap U_2)$, the transition mapping $\kappa_2 \circ \kappa_1^{-1}$ maps (s, t) to $(s + 2\pi, -t)$, whereas on the connected component $s > 0$ it is given by the identical mapping. On both components it is of class C^∞. Therefore, the atlas just constructed turns M into a smooth manifold.

In the course of this book, the Möbius strip will turn up again at several places, notably as an example of a vector bundle (Chap. 2) and as an example of the quotient of a group action (Chap. 6).

Remark 1.1.13 (Quaternions) As a preparation for the examples to follow, let us recall the definition of quaternions. Let $\mathbb{H}$ be the real vector space spanned by the basis elements $\mathbf{1}$, $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$. $\mathbb{H}$ carries an associative multiplication which is defined by the relations

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -\mathbf{1}, \qquad \mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \qquad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \qquad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}.$$

This way, $\mathbb{H}$ becomes a real associative unital algebra with unit element $\mathbf{1}$. As such, it is generated by any two out of the three quaternionic units $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$. Since any nonzero element has an inverse, $\mathbb{H}$ is a division algebra (or, when viewed as a ring, a division ring or skew field). The assignment

$$\mathbf{1} \mapsto \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{i} \mapsto \begin{bmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{bmatrix}, \quad \mathbf{j} \mapsto \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \mathbf{k} \mapsto \begin{bmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix} \tag{1.1.2}$$

extends to an injective homomorphism of real algebras from $\mathbb{H}$ to $\mathrm{M}_2(\mathbb{C})$, the algebra of complex 2×2-matrices. The mapping

$$\mathbf{x} = x_0\mathbf{1} + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} \mapsto \overline{\mathbf{x}} := x_0\mathbf{1} - x_1\mathbf{i} - x_2\mathbf{j} - x_3\mathbf{k},$$

where the x_i are real, is an algebra involution of $\mathbb{H}$, called quaternionic conjugation.

Example 1.1.14 (General linear group) Let $\mathrm{M}_n(\mathbb{K})$ denote the algebra of $n \times n$-matrices with entries from $\mathbb{K} = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$ and let $\mathrm{GL}(n, \mathbb{K})$ denote the subset of invertible matrices. $\mathrm{GL}(n, \mathbb{K})$ is a group and $\mathrm{M}_n(\mathbb{K})$ is a real vector space of dimension n^2 for $\mathbb{K} = \mathbb{R}$, $2n^2$ for $\mathbb{K} = \mathbb{C}$ and $4n^2$ for $\mathbb{K} = \mathbb{H}$. Using the fact that the operator norm on $\mathrm{M}_n(\mathbb{K})$ satisfies $\|AB\| \leq \|A\|\|B\|$ for all $A, B \in \mathrm{M}_n(\mathbb{K})$, one can show that $\mathrm{GL}(n, \mathbb{K})$ is open in $\mathrm{M}_n(\mathbb{K})$, see Exercise 1.1.5. In case $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, openness also follows from the fact that, here, $\mathrm{GL}(n,\mathbb{K})$ is the preimage of the open subset $\mathbb{K}\setminus\{0\}\subset\mathbb{K}$ under the determinant mapping

$$\det : \mathrm{M}_n(\mathbb{K}) \to \mathbb{K},$$

because the determinant is n-linear and hence continuous. As a consequence of Examples 1.1.6 and 1.1.8, $\mathrm{GL}(n,\mathbb{K})$ is a differentiable manifold of dimension $n^2\cdot\dim_{\mathbb{R}}\mathbb{K}$. According to Example 1.1.8, global charts are given by the mappings

$$\mathrm{GL}(n,\mathbb{R}) \to \mathbb{R}^{n^2}, \qquad a \mapsto \left(a^i{}_j\right)$$
$$\mathrm{GL}(n,\mathbb{C}) \to \mathbb{R}^{2n^2}, \qquad a_1 + \mathrm{i}a_2 \mapsto \left((a_1)^i{}_j, (a_2)^i{}_j\right)$$
$$\mathrm{GL}(n,\mathbb{H}) \to \mathbb{R}^{4n^2}, \qquad a_1\mathbf{1} + a_2\mathbf{i} + a_3\mathbf{j} + a_4\mathbf{k} \mapsto \left((a_1)^i{}_j, (a_2)^i{}_j, (a_3)^i{}_j, (a_4)^i{}_j\right),$$

where $a_l \in \mathrm{GL}(n,\mathbb{R})$ and $a^i{}_j$ are the matrix entries of a.[3] The mapping

$$\mathrm{GL}(n,\mathbb{K}) \times \mathrm{GL}(n,\mathbb{K}) \to \mathrm{GL}(n,\mathbb{K}), \qquad (a,b) \mapsto ab^{-1}, \tag{1.1.3}$$

is smooth, because with respect to the above global chart, it is given by a system of rational functions with nonvanishing denominators. This means that $\mathrm{GL}(n,\mathbb{K})$ is a Lie group, cf. Chap. 5.

Let us finish this example with a remark on the topological structure of $\mathrm{GL}(n,\mathbb{R})$. Since $\det : \mathrm{GL}(n,\mathbb{R}) \to \mathbb{R}$ is continuous and $\mathbb{R}\setminus\{0\}$ consists of the connected components $x > 0$ and $x < 0$, $\mathrm{GL}(n,\mathbb{R})$ decomposes into the open subsets

$$\mathrm{GL}(n,\mathbb{R})_\pm = \left\{a \in \mathrm{GL}(n,\mathbb{R}) : \pm\det(a) > 0\right\}. \tag{1.1.4}$$

Both are manifolds of dimension n^2. Due to

$$\mathrm{GL}(n,\mathbb{R})_- = \begin{bmatrix} -1 & 0 \\ 0 & \mathbb{1}_{n-1} \end{bmatrix} \cdot \mathrm{GL}(n,\mathbb{R})_+,$$

they are homeomorphic. One can show that $\mathrm{GL}(n,\mathbb{R})_+$, and hence also $\mathrm{GL}(n,\mathbb{R})_-$, is connected. Thus, $\mathrm{GL}(n,\mathbb{R})_+$ and $\mathrm{GL}(n,\mathbb{R})_-$ are the connected components of $\mathrm{GL}(n,\mathbb{R})$. In contrast to that, $\mathrm{GL}(n,\mathbb{C})$ is connected. For proofs, see Exercise 5.1.9 in Chap. 5 or, for example, [129].

Example 1.1.15 (Projective space) Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$. We use the notation

$$\mathbb{K}^n_* = \left\{\mathbf{x} \in \mathbb{K}^n : \mathbf{x} \neq 0\right\}, \qquad \mathbb{K}^n_1 := \left\{\mathbf{x} \in \mathbb{K}^n : \|\mathbf{x}\| = 1\right\},$$

where $\|\cdot\|$ denotes the norm of the natural scalar product $\mathbf{x}^\dagger\mathbf{y} = \sum_{i=1}^n \overline{x_i}\, y_i$. Here $\overline{x_i}$ denotes the natural involution of $\mathbb{K}$, that is, the identical mapping for $\mathbb{K} = \mathbb{R}$, complex conjugation for $\mathbb{K} = \mathbb{C}$ and quaternionic conjugation for $\mathbb{K} = \mathbb{H}$. The projective

[3] Usually, we will denote elements of an algebra by capital $A, B, C, \ldots$ and elements of a group by small $a, b, c, \ldots$.

space $\mathbb{K}\mathrm{P}^n$ is defined as the topological quotient of $\mathbb{K}_*^{n+1}$ by the equivalence relation $\mathbf{x} \sim \mathbf{y}$ iff $\mathbf{y} = \mathbf{x}\lambda$ for some $\lambda \in \mathbb{K}_*$.[4] Its elements correspond to one-dimensional $\mathbb{K}$-subspaces of $\mathbb{K}^{n+1}$. Let $p : \mathbb{K}_*^{n+1} \to \mathbb{K}\mathrm{P}^n$ denote the natural projection to equivalence classes. First, we show that $\mathbb{K}\mathrm{P}^n$ is Hausdorff. Let $m_1 \neq m_2$ be elements of $\mathbb{K}\mathrm{P}^n$. Choose representatives $\mathbf{x}, \mathbf{y} \in \mathbb{K}_*^{n+1}$ with $\|\mathbf{x}\| = 1 = \|\mathbf{y}\|$, denote

$$l(\mathbf{x}, \mathbf{y}) := \min_{\lambda \in \mathbb{K}_1} \|\mathbf{x}\lambda - \mathbf{y}\|$$

and consider the subsets

$$K_{\mathbf{x}} := \left\{ \mathbf{z} \in \mathbb{K}_*^{n+1} : \max_{\lambda \in \mathbb{K}_1} \left\| \frac{\mathbf{z}}{\|\mathbf{z}\|}\lambda - \mathbf{x} \right\| < \frac{1}{2} l(\mathbf{x}, \mathbf{y}) \right\} \subset \mathbb{K}_*^{n+1}$$

and $K_{\mathbf{y}}$, defined analogously. Since $\mathbb{K}_1$ is compact, the maxima and the minimum exist. Due to $m_1 \neq m_2$, the minimum is nonzero. The subsets $K_{\mathbf{x}}$ and $K_{\mathbf{y}}$ are disjoint and open unions of equivalence classes. Therefore, $p(K_{\mathbf{x}})$ and $p(K_{\mathbf{y}})$ are disjoint neighbourhoods of m_1 and m_2 in $\mathbb{K}\mathrm{P}^n$, respectively. Next, we construct local charts. To this end, we observe that if an equivalence class possesses a representative with $x_i \neq 0$ for a given i, then it possesses a representative with $x_i = 1$ and the latter is unique. Hence, if we put

$$V_i := \left\{ \mathbf{x} \in \mathbb{K}_*^{n+1} : x_i = 1 \right\}, \qquad U_i := p(V_i),$$

then each element of U_i has a unique representative in V_i and, by restriction, p induces bijections $V_i \to U_i$. By inverting these mappings and by identifying V_i in the obvious way with $\mathbb{K}^n$, we obtain bijective mappings $\kappa_i : U_i \to \mathbb{K}^n$. We show that (U_i, κ_i) are local charts. Since their inverses are given by p and hence are continuous, we only have to show that U_i is open and that κ_i is continuous. To this end we put $W_i := \{\mathbf{x} \in \mathbb{K}_*^{n+1} : x_i \neq 0\}$ and consider the mapping

$$\chi_i : W_i \to \mathbb{K}^n \times \mathbb{K}_*, \qquad \chi_i(\mathbf{x}) := \left(\left(x_1 x_i^{-1}, \ldots, \widehat{x_i}, \ldots, x_{n+1} x_i^{-1} \right), x_i \right).$$

Let $A \subset \mathbb{K}^n$ be open. There holds $p^{-1}(\kappa_i^{-1}(A)) = \chi_i^{-1}(A \times \mathbb{K}_*)$. Since χ_i is continuous, $p^{-1}(\kappa_i^{-1}(A))$ is open in W_i and hence open in $\mathbb{K}_*^{n+1}$. Then, $\kappa_i^{-1}(A)$ is open in $\mathbb{K}\mathrm{P}^n$. This implies, first, that U_i is open and, second, that κ_i is continuous. Thus, for each i, (U_i, κ_i) is a local chart on $\mathbb{K}\mathrm{P}^n$, indeed. The corresponding coordinates are called homogeneous. Since the U_i cover $\mathbb{K}\mathrm{P}^n$, the local charts (U_i, κ_i) provide an atlas of $\mathbb{K}\mathrm{P}^n$ and thus equip $\mathbb{K}\mathrm{P}^n$ with the structure of a topological manifold. We leave it to the reader to check that the transition mappings are smooth (Exercise 1.1.6).

Remark 1.1.16 The groups $\mathrm{GL}(n, \mathbb{K})$ are examples of Lie groups. The projective spaces $\mathbb{K}\mathrm{P}^n$ are examples of homogeneous spaces. Both Lie groups and homogeneous spaces will be treated in detail in Chap. 5.

---

[4] We adopt the convention that scalars multiply from the right; this is of course relevant for $\mathbb{K} = \mathbb{H}$ only.

To conclude this section, we discuss the operations of direct product and direct sum. Let M_1 and M_2 be differentiable manifolds of class C^k and dimensions n_1 and n_2. Let $\mathscr{A}_i$ be atlases for M_i. The direct product of topological spaces $M_1 \times M_2$ is Hausdorff and second countable. Any pair of local charts $(U_1, \kappa_1) \in \mathscr{A}_1$ and $(U_2, \kappa_2) \in \mathscr{A}_2$ defines a local chart on $M_1 \times M_2$ by $(U_1 \times U_2, \kappa_1 \times \kappa_2)$, where

$$\kappa_1 \times \kappa_2 : U_1 \times U_2 \to \mathbb{R}^{n_1+n_2}, \qquad (\kappa_1 \times \kappa_2)(m_1, m_2) := \big(\kappa_1(m_1), \kappa_2(m_2)\big).$$

This way, the direct product of atlases $\mathscr{A}_1 \times \mathscr{A}_2$ provides an atlas for $M_1 \times M_2$. Since for the transition mappings there holds

$$(\kappa_1 \times \kappa_2) \circ (\rho_1 \times \rho_2)^{-1} = \big(\kappa_1 \circ \rho_1^{-1}\big) \times \big(\kappa_2 \circ \rho_2^{-1}\big),$$

the local charts are compatible of class C^k. Hence, $M_1 \times M_2$, together with the atlas $\mathscr{A}_1 \times \mathscr{A}_2$, is a differentiable manifold of class C^k and dimension $n_1 + n_2$. It is called the direct product of M_1 and M_2.

Now let $n_1 = n_2 = n$. The direct sum of topological spaces $M_1 \sqcup M_2$ is Hausdorff and second countable. The disjoint union of atlases $\mathscr{A}_1 \sqcup \mathscr{A}_2$ yields an atlas for $M_1 \sqcup M_2$. Since the domains of charts from $\mathscr{A}_1$ do not intersect the domains of charts from $\mathscr{A}_2$, this atlas is of class C^k. Hence, $M_1 \sqcup M_2$ together with the atlas $\mathscr{A}_1 \sqcup \mathscr{A}_2$ is a differentiable manifold of class C^k and dimension n. It is called the direct sum of M_1 and M_2.

Example 1.1.17

1. The n-torus $\mathrm{T}^n = \mathrm{S}^1 \times \overset{n}{\cdots} \times \mathrm{S}^1$ is a smooth manifold of dimension n. The cylinder $\mathrm{S}^1 \times \mathbb{R}$ is a smooth manifold of dimension 2.
2. As a manifold, the general linear group $\mathrm{GL}(n, \mathbb{R})$ is the direct sum of the manifolds $\mathrm{GL}(n, \mathbb{R})_+$ and $\mathrm{GL}(n, \mathbb{R})_-$, see Example 1.1.14.

Exercises

1.1.1 Show that the initial topology of S^n induced by stereographic projection with respect to an arbitrary point and its antipode coincides with the relative topology induced from $\mathbb{R}^{n+1}$.

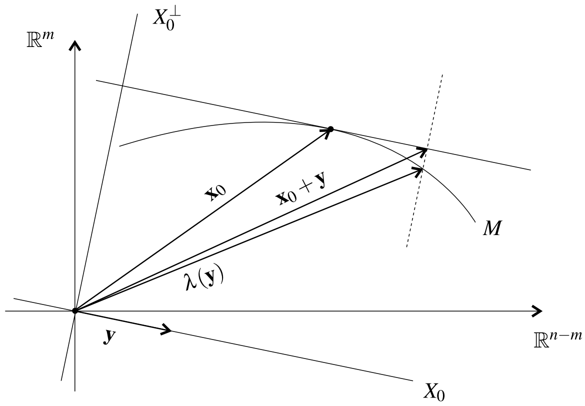
1.1.2 Prove the statements of Example 1.1.11.

1.1.3 Show that (1.1.1) defines a metric on the Möbius strip which is compatible with the quotient topology, see Example 1.1.12.

1.1.4 A model of a Möbius strip can be produced by gluing the ends of a long narrow paper strip, with one end twisted by an angle of 180 degrees. Consider the strip one obtains when one end is twisted by 360 degrees instead. A strip with this twisting is also obtained by cutting a Möbius strip along the middle line. Show that this strip is homeomorphic to the untwisted strip. Why can't the strip be untwisted though?

1.1.5 Prove that $\mathrm{GL}(n, \mathbb{K})$ is open in $\mathrm{M}_n(\mathbb{K})$, see Example 1.1.14.
Hint. For $A \in \mathrm{GL}(n, \mathbb{K})$ define $U_A := \{B \in \mathrm{M}_n(\mathbb{K}) : \|A - B\| < \|A^{-1}\|^{-1}\}$. Show that U_A is a neighbourhood of A in $\mathrm{M}_n(\mathbb{K})$ and that for

Fig. 1.3 Level Set Theorem



every $B \in U_A$, the series $\sum_{k=0}^{\infty}(\mathbb{1} - A^{-1}B)^k A^{-1}$ converges absolutely and the limit is inverse to B.

1.1.6 Verify that the local charts of Example 1.1.15 are smoothly compatible.

1.1.7 Show the following.

(a) $\mathbb{K}\mathrm{P}^1$ is homeomorphic to $\mathrm{S}^{\dim_{\mathbb{R}}\mathbb{K}}$.

(b) $\mathbb{R}\mathrm{P}^2$ is homeomorphic to the space obtained by contracting the boundary of a Möbius strip to a point.
Hint. Study how the points of the subset $\{\mathbf{x} \in \mathbb{R}^3_* : x_1^2 + x_2^2 = 1, |x_3| < a\}$ of $\mathbb{R}^3_*$ get identified via the defining equivalence relation of $\mathbb{R}\mathrm{P}^2$. Then, let $a \to \infty$.

1.2 Level Sets

Level sets at regular values of differentiable mappings of $\mathbb{R}^n$ provide a great variety of examples for manifolds. Let us start this section with recalling some terminology. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable mapping of class C^k, $k \geq 1$. An element $\mathbf{x} \in \mathbb{R}^n$ is called a regular point of f if $f'(\mathbf{x})$ has rank m; otherwise $\mathbf{x}$ is called a critical or singular point. An element $\mathbf{c} \in \mathbb{R}^m$ is called a regular value of f if all points $\mathbf{x} \in f^{-1}(\mathbf{c})$ are regular; otherwise $\mathbf{c}$ is called a critical or singular value. If $\mathbf{c}$ is a regular value, then necessarily $n \geq m$.

Theorem 1.2.1 (Level Set Theorem) *Let $U \subset \mathbb{R}^n$ be open, let $f : U \to \mathbb{R}^m$ be a differentiable mapping of class C^k and let $\mathbf{c} \in \mathbb{R}^m$ be a regular value of f such that the level set $M := f^{-1}(\mathbf{c})$ is nonempty. Then, M is a differentiable manifold of class C^k and dimension $n - m$.*

Proof Let $\mathbf{x}_0 \in M$. We define

$$X_0 := \ker f'(\mathbf{x}_0) \equiv \left\{\mathbf{x} \in \mathbb{R}^n : f'(\mathbf{x}_0)(\mathbf{x}) = 0\right\} \subset \mathbb{R}^n.$$

Since $f'(\mathbf{x}_0)$ has rank m, X_0 is an $(n-m)$-dimensional subspace of $\mathbb{R}^n$. Let $X_0^\perp$ denote the orthogonal complement of X_0 with respect to the Euclidean metric and let $p_0 : \mathbb{R}^n \to X_0$ be the orthogonal projection to X_0, see Fig. 1.3. We show that p_0 defines a local chart on M. The mapping

$$h : X_0 \times X_0^\perp \to \mathbb{R}^m, \qquad h(\mathbf{y}, \mathbf{z}) := f(\mathbf{x}_0 + \mathbf{y} + \mathbf{z}) - \mathbf{c},$$

is of class C^k and satisfies $h(0,0) = 0$. Since for any $\mathbf{w} \in X_0^\perp$,

$$h'(0,0)(0,\mathbf{w}) = f'(\mathbf{x}_0)(\mathbf{w}),$$

the equality $h'(0,0)(0,\mathbf{w}) = 0$ implies $\mathbf{w} \in X_0$ and hence $\mathbf{w} = 0$. According to the Implicit Function Theorem, there exists an open neighbourhood V of 0 in X_0 and a mapping $F : V \to X_0^\perp$ of class C^k such that $h(\mathbf{y}, F(\mathbf{y})) = 0$ for all $\mathbf{y} \in V$. Then, $\mathbf{x}_0 + \mathbf{y} + F(\mathbf{y}) \in M$ for all $\mathbf{y} \in V$, so that we obtain a continuous mapping

$$\lambda : V \to M, \qquad \lambda(\mathbf{y}) := \mathbf{x}_0 + \mathbf{y} + F(\mathbf{y}),$$

see Fig. 1.3. Let $U := \lambda(V)$ and define the mapping

$$\kappa : U \to V, \qquad \kappa(\mathbf{x}) := p_0(\mathbf{x} - \mathbf{x}_0).$$

Check that $U = M \cap (\mathbf{x}_0 + p_0^{-1}(V))$, $\kappa \circ \lambda = \mathrm{id}_V$ and $\lambda \circ \kappa = \mathrm{id}_U$. It follows that U is open in M and κ is a homeomorphism, that is, (U, κ) is a local chart on M. Application of this construction to each $\mathbf{x} \in M$ yields an atlas for M. We check that the local charts are compatible of class C^k: let $(\tilde{U}, \tilde{\kappa})$ be the chart associated with $\tilde{\mathbf{x}}_0 \in M$ and assume $U \cap \tilde{U} \neq \varnothing$. The transition mapping is given by

$$X_0 \supset \kappa(U \cap \tilde{U}) \xrightarrow{\tilde{\kappa} \circ \kappa^{-1}} \tilde{\kappa}(U \cap \tilde{U}) \subset \tilde{X}_0, \qquad y \mapsto \tilde{p}_0\big(\mathbf{x}_0 + \mathbf{y} + F(\mathbf{y}) - \tilde{\mathbf{x}}_0\big).$$

As a composition of F with affine mappings it is of class C^k. □

Remark 1.2.2

1. The subspace $X_0 = \ker f'(\mathbf{x}_0)$ is spanned by the tangent vectors of differentiable curves in M passing through $\mathbf{x}_0$. It is, therefore, called the tangent space of M at $\mathbf{x}_0$. By shifting this subspace to $\mathbf{x}_0$ one obtains an affine subspace, the tangent hyperplane, which is tangent to M in $\mathbf{x}_0$ and whose vector space of translations is given by the tangent space X_0, see Fig. 1.3.
2. In the case $m = 1$, $f : U \to \mathbb{R}$ is a C^k-function and $f'(\mathbf{x})$ is the gradient of f at $\mathbf{x}$. Since the rank can be at most 1, a point $\mathbf{x} \in U$ is regular iff $f'(\mathbf{x}) \neq 0$. Thus, $c \in \mathbb{R}$ is a regular value of f and hence the level set $f^{-1}(c)$ is a C^k-manifold iff the gradient does not vanish on $f^{-1}(c)$.
3. Theorem 1.2.1 and the above remarks generalize in an obvious way to the situation where U is an open subset of a finite-dimensional real vector space W and f takes values in a finite-dimensional real vector space V. We also note that, as a consequence of the proof, there exists an atlas of the level set $M \subset W$ whose local charts (U, κ) fulfil the following conditions:

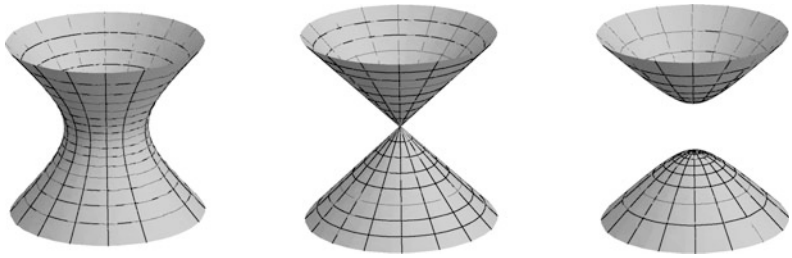


Fig. 1.4 The level sets of Example 1.2.4

(a) κ is the restriction to U of a linear mapping $W \to \mathbb{R}^{\dim M}$,
(b) the mapping $\iota \circ \kappa^{-1} : \kappa(U) \to W$, where $\iota : M \to W$ is the natural inclusion mapping, is of class C^k.

It follows that, given level sets $M_i \subset W_i$, $i = 1, 2$, of C^k-mappings at regular values and a C^k-mapping $\Phi : W_1 \to W_2$ satisfying $\Phi(M_1) \subset M_2$, the induced mapping $M_1 \to M_2$ is of class C^k (Exercise 1.2.1).

Example 1.2.3 (Spheres) The sphere S^n is the level set of the function

$$f : \mathbb{R}^{n+1} \to \mathbb{R}, \qquad f(\mathbf{x}) := \|\mathbf{x}\|^2$$

at the value $c = 1$, cf. Example 1.1.9. Since $\mathbf{x} = 0$ is the only singular point of f, the assumptions of Theorem 1.2.1 are satisfied. Therefore, this theorem yields a smooth structure on S^n. One can show that this structure coincides with the one constructed by means of stereographic projection in Example 1.1.9 (Exercise 1.2.4). For $\mathbf{x} \in S^n$, the subspace $\ker f'(\mathbf{x})$ consists of those vectors which are orthogonal to $\mathbf{x}$. By shifting this subspace to $\mathbf{x}$, one obtains the tangent plane of the sphere at this point, indeed.

Example 1.2.4 (Hyperboloid) We consider the function

$$f : \mathbb{R}^3 \to \mathbb{R}, \qquad f(\mathbf{x}) := x_1^2 + x_2^2 - x_3^2.$$

The only singular point of f is $\mathbf{x} = 0$. Hence, the only singular value of f is $c = f(0) = 0$; all $c \neq 0$ are regular. We determine the level sets $M = f^{-1}(c)$. Since f is invariant under rotations about the x_3-axis, M is the surface of revolution of the curve $x_1^2 - x_3^2 = c$ in the x_1–x_3-plane. This is a hyperboloid which is one-sheeted for $c > 0$ and two-sheeted for $c < 0$, see Fig. 1.4. In both cases, Theorem 1.2.1 yields a smooth structure. As a link between these two cases, for the singular value $c = 0$, M is a double cone. In this case, Theorem 1.2.1 does not apply. In fact, the double cone is *not* a topological manifold (Exercise 1.2.2). In the case $c < 0$, using $\mathbf{x}_0 = (0, \pm\sqrt{c}, 0)$, one obtains a global chart for each sheet. Since the domains of these two charts do not intersect, this atlas is automatically smooth. In the case

$c > 0$, one needs at least three charts of the type used in the proof of the Level Set Theorem. We leave this case to the reader (Exercise 1.2.3).

Remark 1.2.5 In contrast to the double cone $\{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 - x_3^2 = 0\}$ of Example 1.2.4, the single cone $\{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 - x_3^2 = 0, x_3 \geq 0\}$ is a topological manifold, and even a manifold of class C^∞. Indeed, orthogonal projection to the x_1–x_2-plane yields a global chart.

Example 1.2.6 (Classical groups) Let $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{H}$. For $n, m = 0, 1, 2, \ldots$, define

$$\mathbb{1}_{n,m} := \begin{bmatrix} \mathbb{1}_n & 0 \\ 0 & -\mathbb{1}_m \end{bmatrix}, \qquad J_n = \begin{bmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{bmatrix}$$

and recall the following inner products on $\mathbb{K}^n$:

$$\mathbf{x}^{\mathrm{T}}\mathbf{y} = \sum_{i=1}^{n} x_i\, y_i, \qquad \mathbf{x}^{\dagger}\mathbf{y} = \sum_{i=1}^{n} \overline{x_i}\, y_i, \qquad x_i, y_i \in \mathbb{K}.$$

As before, $\overline{x_i}$ denotes the natural involution on $\mathbb{K}$. By a classical group one means the general linear group $\mathrm{GL}(n, \mathbb{K})$ or one of the following subgroups.

1. The *unimodular group* $\mathrm{SL}(n, \mathbb{K})$, $\mathbb{K} = \mathbb{R}, \mathbb{C}$, is the group of linear mappings of unit determinant,
$$\mathrm{SL}(n, \mathbb{K}) = \big\{a \in \mathrm{GL}(n, \mathbb{K}) : \det(a) = 1\big\}.$$
2. The *real orthogonal group* $\mathrm{O}(n, m) \subset \mathrm{GL}(n + m, \mathbb{R})$ is the group of isometries of the symmetric bilinear form $\mathbf{x}^{\mathrm{T}}\mathbb{1}_{n,m}\mathbf{y}$ on $\mathbb{R}^{n+m}$,
$$\mathrm{O}(n, m) = \big\{a \in \mathrm{GL}(n + m, \mathbb{R}) : a^{\mathrm{T}}\mathbb{1}_{n,m}a = \mathbb{1}_{n,m}\big\}.$$
In case $m = 0$ one writes $\mathrm{O}(n) = \mathrm{O}(n, 0)$; this is the group of isometries of the Euclidean scalar product.
3. The *special real orthogonal group* $\mathrm{SO}(n, m)$ is the subgroup of $\mathrm{O}(n, m)$ of isometries with unit determinant,
$$\mathrm{SO}(n, m) = \mathrm{O}(n, m) \cap \mathrm{SL}(n + m, \mathbb{R}).$$
In case $m = 0$, one writes $\mathrm{SO}(n)$.
4. The *complex orthogonal group* $\mathrm{O}(n, \mathbb{C}) \subset \mathrm{GL}(n, \mathbb{C})$ is the group of isometries of the symmetric bilinear form $\mathbf{x}^{\mathrm{T}}\mathbf{y}$ on $\mathbb{C}^n$,
$$\mathrm{O}(n, \mathbb{C}) = \big\{a \in \mathrm{GL}(n, \mathbb{C}) : a^{\mathrm{T}}a = \mathbb{1}\big\}.$$
5. The *special complex orthogonal group* is the subgroup of $\mathrm{O}(n, \mathbb{C})$ of isometries of unit determinant,
$$\mathrm{SO}(n, \mathbb{C}) = \mathrm{O}(n, \mathbb{C}) \cap \mathrm{SL}(n, \mathbb{C}).$$

6. The *unitary group* $\mathrm{U}(n,m) \subset \mathrm{GL}(n+m,\mathbb{C})$ is the group of isometries of the Hermitian form $\mathbf{x}^\dagger \mathbb{1}_{n,m}\mathbf{y}$ on $\mathbb{C}^{n+m}$,
$$\mathrm{U}(n,m) = \left\{a \in \mathrm{GL}(n+m,\mathbb{C}) : a^\dagger \mathbb{1}_{n,m} a = \mathbb{1}_{n,m}\right\}.$$
In case $m=0$ one writes $\mathrm{U}(n)=\mathrm{U}(n,0)$; this is the subgroup of isometries of the natural scalar product on $\mathbb{C}^n$.
7. The *special unitary group* $\mathrm{SU}(n,m)$ is the subgroup of $\mathrm{U}(n,m)$ of isometries of unit determinant,
$$\mathrm{SU}(n,m) = \mathrm{U}(n,m) \cap \mathrm{SL}(n+m,\mathbb{C}).$$
In case $m=0$ one writes $\mathrm{SU}(n)=\mathrm{SU}(n,0)$.
8. The *symplectic group* $\mathrm{Sp}(n,\mathbb{K}) \subset \mathrm{GL}(2n,\mathbb{K})$, $\mathbb{K}=\mathbb{R},\mathbb{C}$, is the group of isometries of the antisymmetric bilinear form $\omega(\mathbf{x},\mathbf{y}) = \mathbf{x}^{\mathrm{T}} J_n \mathbf{y}$ on $\mathbb{K}^{2n}$,
$$\mathrm{Sp}(n,\mathbb{K}) = \left\{a \in \mathrm{GL}(2n,\mathbb{K}) : a^{\mathrm{T}} J_n a = J_n\right\}.$$
9. The *quaternionic symplectic group* $\mathrm{Sp}(n,m) \subset \mathrm{GL}(n+m,\mathbb{H})$ is the group of isometries of the Hermitian form $\mathbf{x}^\dagger \mathbb{1}_{n,m}\mathbf{y}$ on $\mathbb{H}^{n+m}$,
$$\mathrm{Sp}(n,m) = \left\{a \in \mathrm{GL}(n+m,\mathbb{H}) : a^\dagger \mathbb{1}_{n,m} a = \mathbb{1}_{n,m}\right\}.$$
In case $m=0$ one writes $\mathrm{Sp}(n)=\mathrm{Sp}(n,0)$; this is the group of isometries of the natural scalar product on $\mathbb{H}^n$.

By writing down defining relations and applying the Level Set Theorem one can show that all the classical groups are smooth manifolds (Exercise 1.2.6). As an example, we carry out the proof for the groups $\mathrm{O}(n)$ and $\mathrm{SO}(n)$.

Let $\mathrm{S}_n(\mathbb{R})$ denote the linear subspace of $\mathrm{M}_n(\mathbb{R})$ of symmetric matrices. $\mathrm{O}(n)$ is the level set at the value $c=\mathbb{1}$ of the mapping $f : \mathrm{M}_n(\mathbb{R}) \to \mathrm{S}_n(\mathbb{R})$, $f(a) := a^{\mathrm{T}} a$. Hence, for $\mathrm{O}(n)$ it suffices to show that $\mathbb{1}$ is a regular value of f. To do so, we calculate the derivative f' of f in $a \in \mathrm{M}_n(\mathbb{R})$:
$$f'(a)(X) = a^{\mathrm{T}} X + X^{\mathrm{T}} a, \qquad X \in \mathrm{M}_n(\mathbb{R}).$$
If $f(a)=\mathbb{1}$, then for any $B \in \mathrm{S}_n(\mathbb{R})$ there holds $f'(a)(aB) = 2B$. This shows that $f'(a)$ is surjective for all $a \in f^{-1}(\mathbb{1})$, that is, $\mathbb{1}$ is a regular value of f, indeed. Then, for $\mathrm{SO}(n)$ it suffices to show that it is open in $\mathrm{O}(n)$. Since for any $a \in \mathrm{O}(n)$ there holds $\det(a^2) = \det(a^{\mathrm{T}} a) = 1$, we have $\det(a) = \pm 1$. Hence,
$$\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{GL}(n,\mathbb{R})_+,$$
cf. Example 1.1.14. Now, the assertion follows from the fact that $\mathrm{GL}(n,\mathbb{R})_+$ is open in $\mathrm{M}_n(\mathbb{R})$.

Finally, the Level Set Theorem also yields the dimensions of the classical groups. For example, since $\mathrm{S}_n(\mathbb{R})$ has dimension $\frac{1}{2}n(n+1)$, $\mathrm{O}(n)$ has dimension
$$n^2 - \frac{1}{2}n(n+1) = \frac{1}{2}n(n-1).$$

The same is true for the open subset $\mathrm{SO}(n)$. It turns out that the groups $\mathrm{O}(n,m)$ and $\mathrm{SO}(n,m)$ have the same dimension as $\mathrm{O}(n+m)$. The dimensions of the other classical groups are

$$\begin{aligned} \mathrm{SL}(n,\mathbb{R}) &: \quad n^2-1, & \mathrm{SL}(n,\mathbb{C}) &: \quad 2n^2-2,\\ \mathrm{O}(n,\mathbb{C}), \mathrm{SO}(n,\mathbb{C}) &: \quad n(n-1), & \mathrm{U}(n,m) &: \quad (n+m)^2,\\ \mathrm{SU}(n,m) &: \quad (n+m)^2-1, & \mathrm{Sp}(n,\mathbb{R}) &: \quad n(2n+1),\\ \mathrm{Sp}(n,\mathbb{C}) &: \quad 2n(2n+1), & \mathrm{Sp}(n,m) &: \quad (n+m)\big(2(n+m)+1\big). \end{aligned}$$

Remark 1.2.7

1. Let $G \subset \mathrm{GL}(n,\mathbb{K})$ denote one of the above classical groups. With respect to the smooth structures provided by the Level Set Theorem, the mapping $G \times G \to G$, given by $(a,b) \mapsto ab^{-1}$, is smooth. This follows from Remark 1.2.2/3 and from the smoothness of the mapping (1.1.3). We conclude that the classical groups are Lie groups and, in addition, Lie subgroups of $\mathrm{GL}(n,\mathbb{K})$, cf. Chap. 5.
2. One has $\mathrm{O}(n) \setminus \mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{GL}(n,\mathbb{R})_-$, and this subset is also open in $\mathrm{O}(n)$. Hence, $\mathrm{SO}(n)$ is closed in $\mathrm{O}(n)$. One can show that $\mathrm{SO}(n)$ is connected; this is in fact part of the proof of the connectedness of $\mathrm{GL}(n,\mathbb{R})_+$, see the corresponding remark in Example 1.1.14. Thus, $\mathrm{O}(n)$ consists of the connected components $\mathrm{SO}(n)$ and $\mathrm{O}(n) \setminus \mathrm{SO}(n)$.
3. Being level sets, the classical groups are closed subsets of $\mathrm{GL}(n,\mathbb{K})$. As isometry groups of scalar products, the groups $\mathrm{O}(n)$, $\mathrm{U}(n)$ and $\mathrm{Sp}(n)$ are compact, see Exercise 1.2.7. Since $\mathrm{SO}(n)$ and $\mathrm{SU}(n)$ are closed subsets of $\mathrm{O}(n)$ and $\mathrm{U}(n)$, respectively, they are compact, too. None of the other classical groups is compact.

Exercises

1.2.1 Show that the restriction of a C^k-mapping between vector spaces to level sets is of class C^k, cf. Remark 1.2.2/3.

1.2.2 Show that the double cone $\{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 - x_3^2 = 0\}$ is not a topological manifold, cf. Example 1.2.4.

1.2.3 Using the method of the proof of the Level Set Theorem, construct an atlas for the one-sheeted hyperboloid of Example 1.2.4.

1.2.4 Show that the smooth structure on S^n provided by the Level Set Theorem coincides with the smooth structure constructed by means of stereographic projection in Example 1.1.9.

1.2.5 Show that the following level sets are smooth manifolds and construct atlases:

(a) the paraboloid $M = \{\mathbf{x} \in \mathbb{R}^3 : x_3 = x_1^2 + x_2^2\}$,

(b) the ellipsoid $M = \{\mathbf{x} \in \mathbb{R}^3 : \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 1\}$,

(c) the rotational torus $M = \{\mathbf{x} \in \mathbb{R}^3 : (\sqrt{x_1^2 + x_2^2} - a)^2 + x_3^2 = b^2\}$, where $0 < b < a$.

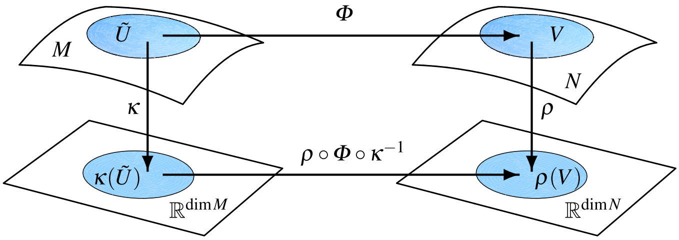


Fig. 1.5 Local representative of a continuous mapping between manifolds. Here, $\tilde{U} = U \cap \Phi^{-1}(V)$

1.2.6 Use the Level Set Theorem 1.2.1 for showing that the classical groups listed in Example 1.2.6 are smooth manifolds.

1.2.7 Show that the isometry group of the natural scalar product on $\mathbb{K}^n$ is compact. *Hint*. Choose an appropriate norm on $\mathrm{M}_n(\mathbb{K})$, $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{H}$, and show that the subset of isometries is bounded.

1.3 Differentiable Mappings

In this section, we carry over the notion of differentiability from classical calculus to the theory of manifolds. Let M and N be C^k-manifolds and let $\Phi : M \to N$ be a continuous mapping.[5] As usual, we will refer to M, to N, to the elements of N and to $\Phi(M)$ as, respectively, the domain, the range, the values and the image of Φ. Let (U, κ) be a local chart on M and let (V, ρ) be a local chart on N such that $U \cap \Phi^{-1}(V) \neq \varnothing$. The mapping

$$\Phi_{\kappa,\rho} : \kappa\big(U \cap \Phi^{-1}(V)\big) \to \rho(V), \qquad \Phi_{\kappa,\rho}(\mathbf{x}) := \rho \circ \Phi \circ \kappa^{-1}(\mathbf{x}),$$

pictured in Fig. 1.5, is called the local representative or the local representation of Φ with respect to the charts (U, κ) and (V, ρ). Since by continuity of Φ, $\Phi^{-1}(V)$ is open, the local representatives are mappings between open subsets of $\mathbb{R}^{\dim M}$ and $\mathbb{R}^{\dim N}$. Hence, it makes sense to ask whether they are differentiable.

Definition 1.3.1 (Differentiable mapping) Let M and N be C^k-manifolds. A mapping $\Phi : M \to N$ is called differentiable of class C^k, or just a C^k-mapping, if it is continuous and if for any pair of local charts (U, κ) on M and (V, ρ) on N such that $U \cap \Phi^{-1}(V) \neq \varnothing$, the local representative $\Phi_{\kappa,\rho}$ is of class C^k. The set of all such mappings is denoted by $C^k(M, N)$.

[5] In this book, mappings between manifolds are usually denoted by capital Greek letters like $\Phi, \Psi, \ldots$ or small Greek letters like $\varphi, \psi, \chi, \ldots$.

This definition yields the notion of C^k-mapping of C^l-manifolds for any $l \geq k$ by viewing the C^l-manifolds as C^k-manifolds according to Remark 1.1.5/1. By analogy, one defines the notions of real analytic mapping between real analytic manifolds and of complex analytic mapping between complex analytic manifolds by requiring the local representatives to be real or complex analytic, respectively. Continuing previous terminology, C^∞-mappings will be referred to as smooth mappings.

Remark 1.3.2

1. An equivalent definition of differentiability is the following. A mapping $\Phi : M \to N$ is differentiable of class C^k iff there exist atlases $\{(U_i, \kappa_i) : i \in I\}$ on M and $\{(V_j, \rho_j) : j \in J\}$ on N such that for every $i \in I$ there is $j \in J$ such that $\Phi(U_i) \subset V_j$ and the local representative $\Phi_{\kappa_i,\rho_j} : \kappa(U_i) \to \rho(V_j)$ is of class C^k. Note that, here, continuity of f need not be required. Indeed, for every pair of charts (U_i, κ_i), (V_j, ρ_j) such that $\Phi(U_i) \subset V_j$ we have $\Phi_{\restriction U_i} = \rho_j^{-1} \circ \Phi_{\kappa_i,\rho_j} \circ \kappa_i$, which is continuous as a composition of continuous mappings.
2. A mapping between open subsets of finite-dimensional vector spaces, endowed with the standard smooth structure (Example 1.1.6), is of class C^k in the sense of Definition 1.3.1 iff it is of class C^k in the sense of classical calculus. To see this, choose global charts corresponding to two chosen bases. In particular, multilinear mappings between finite-dimensional real vector spaces are smooth.
3. Let (U_i, κ_i) and (V_i, ρ_i), $i = 1, 2$, be local charts on M and N, respectively, such that
$$W := U_1 \cap U_2 \cap \Phi^{-1}(V_1 \cap V_2) \neq \varnothing.$$
Then, for any $m \in W$,
$$\Phi_{\kappa_2,\rho_2}(m) = \left(\rho_2 \circ \rho_1^{-1}\right) \circ \Phi_{\kappa_1,\rho_1} \circ \left(\kappa_1 \circ \kappa_2^{-1}\right)(m).$$
This shows that the local representatives with respect to two different pairs of charts are related via the transition mappings between these charts. In particular, in order to decide whether a continuous mapping $\Phi : M \to N$ is differentiable it suffices to test the local representatives with respect to arbitrary but fixed atlases.
4. Consider the case $N = \mathbb{R}$ with the standard smooth structure. A C^k-mapping $f : M \to \mathbb{R}$ is called a C^k-function. The space of C^k-functions is denoted by $C^k(M)$. It carries the structure of a real associative algebra with operations
$$(\alpha f + g)(m) := \alpha f(m) + g(m), \qquad (f \cdot g)(m) := f(m)g(m),$$
where $f, g \in C^k(M)$, $\alpha \in \mathbb{R}$, and $m \in M$.
5. Consider the case where $M = I \subset \mathbb{R}$ is an open interval with the standard smooth structure. A C^k-mapping $\gamma : I \to N$ is called a C^k-curve. We say that γ is a curve through $p \in N$ if $0 \in I$ and $\gamma(0) = p$.
6. Let (U, κ) be a local chart on the C^k-manifold M. Consider the smooth manifold U with global chart (U, κ). Viewed as a mapping from U to $\mathbb{R}^n$ (with its standard smooth structure), κ is smooth. Similarly, viewed as functions on U, the coordinate functions κ^i are smooth.

Definition 1.3.3 (Diffeomorphism) Let M and N be differentiable manifolds of class C^k. A differentiable mapping $\Phi : M \to N$ of class C^k is called a diffeomorphism of class C^k, or just a C^k-diffeomorphism, if it is bijective and the inverse mapping is of class C^k, too. If a diffeomorphism of class C^k exists, one says that M and N are diffeomorphic of class C^k, or just C^k-diffeomorphic.

As in the case of C^k-mappings in Definition 1.3.1, Definition 1.3.3 yields the notion of C^k-diffeomorphism of C^l-manifolds for all $l \geq k$ as well. If the differentiability class of M and N is fixed, we will usually just speak of diffeomorphisms. In addition, we will use the following terminology. By a local diffeomorphism from M to N we mean a diffeomorphism from an open subset of M onto an open subset of N. By saying that a C^k-mapping $\Phi : M \to N$ is locally a diffeomorphism we mean that for every point of M there exists an open neighbourhood U such that $\Phi(U)$ is open in N and the mapping $U \to \Phi(U)$ induced by Φ is a diffeomorphism.

Example 1.3.4 (Diffeomorphic differentiable structures) Define mappings $\kappa_i : \mathbb{R} \to \mathbb{R}$, $i = 1, 2, 3$, by

$$\kappa_1(x) := x, \qquad \kappa_2(x) := \begin{cases} 2x | x \geq 0 \\ x | x < 0, \end{cases} \qquad \kappa_3(x) := x^3.$$

Each κ_i is a homeomorphism and hence a global chart on $\mathbb{R}$. Therefore, each κ_i defines a smooth atlas $\mathscr{A}_i$ and hence a smooth structure on $\mathbb{R}$. However, no two of the charts κ_i are compatible, because none of the mappings $\kappa_2 \circ \kappa_1^{-1}$, $\kappa_2 \circ \kappa_3^{-1}$ or $\kappa_1 \circ \kappa_3^{-1}$ is differentiable at the origin. Hence, the atlases $\mathscr{A}_i$ are pairwise incompatible. The corresponding smooth structures are diffeomorphic though: consider the mappings $\Phi_i : (\mathbb{R}, \mathscr{A}_i) \to (\mathbb{R}, \mathscr{A}_1)$, $\Phi_i := \kappa_i$, $i = 2, 3$. Since they are homeomorphisms, they are bijective and continuous and their inverses are also continuous. The (global) representative of Φ_i with respect to the charts κ_1 and κ_i is

$$(\Phi_i)_{\kappa_i,\kappa_1} = \kappa_1 \circ \Phi_i \circ \kappa_i^{-1} = \mathrm{id}_{\mathbb{R}}.$$

Analogously, $(\Phi_i^{-1})_{\kappa_1,\kappa_i} = \mathrm{id}_{\mathbb{R}}$. Hence, Φ_i is a diffeomorphism. In contrast, the identical mapping $\Phi = \mathrm{id}_{\mathbb{R}}$ is not a diffeomorphism. Indeed, as a mapping $(\mathbb{R}, \mathscr{A}_1) \to (\mathbb{R}, \mathscr{A}_2)$, $\mathrm{id}_{\mathbb{R}}$ has the global representative $(\mathrm{id}_{\mathbb{R}})_{\kappa_1,\kappa_2} = \kappa_2$ and is thus not differentiable at the origin. As a mapping $(\mathbb{R}, \mathscr{A}_1) \to (\mathbb{R}, \mathscr{A}_3)$, $\mathrm{id}_{\mathbb{R}}$ has the global representative $(\mathrm{id}_{\mathbb{R}})_{\kappa_1,\kappa_3} = \kappa_3$ and is, therefore, of class C^∞. Its inverse, however, is not differentiable at the origin. Finally, as a mapping $(\mathbb{R}, \mathscr{A}_2) \to (\mathbb{R}, \mathscr{A}_3)$, $\mathrm{id}_{\mathbb{R}}$ has the global representative

$$(\mathrm{id}_{\mathbb{R}})_{\kappa_2,\kappa_3} = \begin{cases} (\frac{1}{2}x)^3 | x \geq 0 \\ x^3 | x < 0 \end{cases}$$

and is, therefore, of class C^2. Again, the inverse mapping is not differentiable at the origin.

Remark 1.3.5 (Differentiable structures on $\mathbb{R}^n$) One can show that on the topological space $\mathbb{R}$, all smooth structures are diffeomorphic to the standard smooth structure. This is also true for $\mathbb{R}^n$ except for $n = 4$. On $\mathbb{R}^4$, besides the standard smooth structure, there exist so-called exotic smooth structures [98]. A further example of a topological manifold which possesses non-diffeomorphic smooth structures is the sphere S^7 [159].

Next, we discuss partitions of unity, a tool which will be used for example for integration of differential forms on manifolds. Recall that a family of subsets $\{V_i : i \in I\}$ of M is called locally finite if for any $m \in M$ there is a neighbourhood U_m such that $U_m \cap V_i = \varnothing$ for all but a finite number of $i \in I$.

Definition 1.3.6 (Partition of unity) Let M be a C^k-manifold. A partition of unity of M is a family $\{g_i : i \in I\}$ of C^k-functions on M such that

1. the family of supports $\{\text{supp}(g_i) : i \in I\}$ is locally finite,
2. $g_i \geq 0$ for all $i \in I$,
3. $\sum_{i \in I} g_i(m) = 1$ for all $m \in M$.

As a consequence of property 3, the family of supports of a partition of unity forms a covering of M. Recall that a covering $\{V_\alpha : \alpha \in A\}$ of M is subordinate to a covering $\{W_\beta : \beta \in B\}$ of M if for any $\alpha \in A$ there is an element $\beta \in B$ such that $V_\alpha \subset W_\beta$. One says that a partition of unity is subordinate to a given covering of M if the family of supports of the partition is subordinate to that covering.

Proposition 1.3.7 (Existence) *For any open covering* $\{U_\alpha : \alpha \in A\}$ *of a* C^k-*manifold there exists a partition of unity* $\{g_i : i \in I\}$ *with the following properties*:

1. I *is countable*,
2. $\text{supp}(g_i)$ *is compact for all* $i \in I$,
3. $\{g_i : i \in I\}$ *is subordinate to* $\{U_\alpha : \alpha \in A\}$.

Proof See, for instance, [73, §16.4] or [180, §II.3]. □

Remark 1.3.8 Let us conclude this section with a remark on the relation between the differentiability classes C^k for different k. Let M be a topological manifold and let $k \geq 1$. One can show that every C^k-structure on M is C^k-diffeomorphic to a C^∞-structure and that this C^∞-structure is unique up to C^∞-diffeomorphisms [130, Ch. 3]. One may even replace C^∞ by real analytic here. Another question is how many diffeomorphism classes of differentiable structures (hence of smooth structures) exist on a given topological manifold. On the one hand, according to Remark 1.3.5, there are topological manifolds admitting more than one diffeomorphism class. On the other hand, there exist topological manifolds which are non-smoothable, that is, which do not admit a C^k-structure with $k \geq 1$ at all, see [158, 276].

Exercises

1.3.1 Show that $\mathbb{R}^n \setminus \{0\}$ is diffeomorphic to the cylinder $\mathrm{S}^{n-1} \times \mathbb{R}$ over the sphere S^{n-1}. For $n = 2$ and $n = 3$, how is this fact related to the definition of polar and spherical coordinates?

1.3.2 Show that the mapping $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$, defined by

$$\Phi(\mathbf{x}) := \big(x_1 \sin(x_3) + x_2 \cos(x_3), x_1 \cos(x_3) - x_2 \sin(x_3), x_3\big),$$

maps the unit sphere $\mathrm{S}^2 \subset \mathbb{R}^3$ diffeomorphically onto itself.

1.3.3 Consider the ellipsoid $M = \{\mathbf{x} \in \mathbb{R}^3 : \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 1\}$ of Exercise 1.2.5/(b). Define mappings $\Phi, \Psi : (-\frac{\pi}{2}, \frac{\pi}{2}) \times (0, 2\pi) \to \mathbb{R}^3$ by

$$\Phi(\alpha, \beta) := (a_1 \cos\alpha \cos\beta, a_2 \cos\alpha \sin\beta, a_3 \sin\alpha),$$

$$\Psi(\alpha, \beta) := (-a_1 \cos\alpha \cos\beta, a_2 \sin\alpha, a_3 \cos\alpha \sin\beta).$$

Show that

(a) the images of Φ and Ψ are open subsets of M and hence smooth manifolds,

(b) Φ and Ψ are diffeomorphisms onto their images.

1.3.4 Consider the rotational torus $M = \{\mathbf{x} \in \mathbb{R}^3 : (\sqrt{x_1^2 + x_2^2} - a)^2 + x_3^2 = b^2\}$, $0 < b < a$, of Exercise 1.2.5/(c). Show that

(a) M is diffeomorphic to the direct product $\mathrm{T}^2 = \mathrm{S}^1 \times \mathrm{S}^1$,

(b) the mapping $\Phi : M \to \mathrm{S}^2 \subset \mathbb{R}^3$, defined by $\Phi(\mathbf{x}) := \frac{\mathbf{x}}{\|\mathbf{x}\|}$, is differentiable. What is the image of Φ?

1.3.5 Identify S^1 with the unit circle in $\mathbb{C}$ and define an equivalence relation on $\mathrm{S}^1 \times (-1, 1)$ by $(\alpha, t) \sim (\beta, s)$ if $\beta = \alpha$, $s = t$ or $\beta = -\alpha$, $s = -t$. By analogy with Example 1.1.12, construct a differentiable structure on the corresponding topological quotient. Show that the manifold so obtained is diffeomorphic to the Möbius strip.

1.4 Tangent Space

In this section, we generalize the notion of tangent space, defined in Remark 1.2.2/1 for level sets of C^k-mappings $\mathbb{R}^n \to \mathbb{R}^m$ at regular values, to arbitrary differentiable manifolds. Let $k \geq 1$.

In case M is a level set, the elements of the tangent space at m may be viewed as the tangent vectors of C^k-curves γ in M with $\gamma(0) = m$. To determine the tangent vector of γ one makes use of the fact that γ is also a curve in the ambient linear space $\mathbb{R}^n$, where the derivative $\frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \gamma(t)$ is well defined. Since this model of tangent space is not intrinsic for M, i.e., since it requires more than just the differentiable structure of M, it cannot be carried over to abstract manifolds. The notion of a curve, however, is intrinsic, see Remark 1.3.2/5. Let $K_m(M)$ denote the collection of all

C^k-curves through m. In the situation where M is a level set, two curves through m define the same tangent vector in m if their derivatives at $t = 0$ coincide. In this case, they should be viewed as being equivalent. By means of local charts, this equivalence relation will be carried over to abstract manifolds. For that purpose, we have to make sure that it does not depend on the choice of chart.

Lemma 1.4.1 *Let M be a C^k-manifold, let $m \in M$ and let γ_1, γ_2 be curves of class C^k through m. If $\frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 (\kappa \circ \gamma_1)(t) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 (\kappa \circ \gamma_2)(t)$ for some local chart (U, κ) at m, then this holds for every such chart.*

Proof The proof is an exercise in applying the chain rule and is, therefore, left to the reader (Exercise 1.4.1). □

Thus, we can define two curves γ_1, γ_2 in $K_m(M)$ to be equivalent if for some (and hence any) local chart (U, κ) at m there holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 (\kappa \circ \gamma_1)(t) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 (\kappa \circ \gamma_2)(t). \tag{1.4.1}$$

Let $\mathrm{T}_m M$ denote the set of equivalence classes. On $\mathrm{T}_m M$, a linear structure can be defined as follows. Let (U, κ) be a local chart at m. Due to Lemma 1.4.1, the mapping

$$F_m^\kappa : \mathrm{T}_m M \to \mathbb{R}^n, \qquad [\gamma] \mapsto F_m^\kappa\big([\gamma]\big) := \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 (\kappa \circ \gamma)(t), \tag{1.4.2}$$

is well defined and injective. Since $\mathbf{x} \in \mathbb{R}^n$ is the image under F_m^κ of the equivalence class of the curve

$$\gamma^{\mathbf{x}}(t) := \kappa^{-1}\big(\kappa(m) + t\mathbf{x}\big), \tag{1.4.3}$$

F_m^κ is also surjective. The inverse mapping is given by

$$\big(F_m^\kappa\big)^{-1}(\mathbf{x}) = \big[\gamma^{\mathbf{x}}\big]. \tag{1.4.4}$$

By means of F_m^κ, we transport the linear structure of $\mathbb{R}^n$ to $\mathrm{T}_m M$, that is, we define

$$\alpha\,[\gamma_1] + \beta\,[\gamma_2] := \big(F_m^\kappa\big)^{-1}\Big(\alpha F_m^\kappa\big([\gamma_1]\big) + \beta F_m^\kappa\big([\gamma_2]\big)\Big).$$

This definition does not depend on the choice of chart, because for a second chart (V, ρ) one has

$$F_m^\rho\big([\gamma]\big) = \big(\rho \circ \kappa^{-1}\big)'\big(\kappa(m)\big) \cdot F_m^\kappa\big([\gamma]\big) \tag{1.4.5}$$

(Exercise 1.4.2). As a result, $\mathrm{T}_m M$ is a real linear space of the same dimension as M and the mappings F_m^κ are vector space isomorphisms.

Definition 1.4.2 (Tangent space) The real linear space $\mathrm{T}_m M$ is called the tangent space of M at m. Its elements are called tangent vectors at m.

Tangent vectors will usually be denoted by X_m, Y_m etc. The local representative of a tangent vector $X_m \in \mathrm{T}_m M$ with respect to the chart (U, κ) is defined to be

$$\mathbf{X}_m^\kappa = F_m^\kappa(X_m). \tag{1.4.6}$$

The relation between $\mathbf{X}_m^\kappa$ and the local representative $\mathbf{X}_m^\rho$ with respect to another chart (V, ρ) can be read off from (1.4.5):

$$\mathbf{X}_m^\rho = \left[\left(\rho \circ \kappa^{-1}\right)'\left(\kappa(m)\right)\right] \cdot \mathbf{X}_m^\kappa. \tag{1.4.7}$$

The coefficients of the local representative $\mathbf{X}_m^\kappa$ with respect to the standard basis of $\mathbb{R}^n$ are[6]

$$X_m^{\kappa,i} = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0}\left(\kappa^i \circ \gamma\right)(t). \tag{1.4.8}$$

For these coefficients, (1.4.7) yields the transformation law (summation convention)

$$X_m^{\rho,i} = \left[\left(\rho \circ \kappa^{-1}\right)'\left(\kappa(m)\right)\right]^i_j X_m^{\kappa,j}, \tag{1.4.9}$$

that is, the transformation is given by the Jacobi matrix of the coordinate transformation.

Example 1.4.3

1. Let M be an open subset of a finite-dimensional real vector space V. For every $v \in M$, the assignment of the velocity vector $\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0}\gamma(t) \in V$ to the class $[\gamma] \in \mathrm{T}_v M$ defines an isomorphism of the tangent space $\mathrm{T}_v M$ with the ambient vector space V. The inverse of this isomorphism is obtained by assigning to $u \in V$ the class of the curve $t \mapsto v + tu$. This isomorphism will be referred to as the natural identification of $\mathrm{T}_v M$ with V.
2. Let V and W be finite-dimensional real vector spaces and let $M \subset V$ be the level set of a regular value of a C^k-mapping $f : V \to W$. For every $v \in M$, the assignment of the velocity vector $\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0}\gamma(t) \in V$ to the class $[\gamma] \in \mathrm{T}_v M$ defines a natural isomorphism of the tangent space $\mathrm{T}_v M$ with the subspace $\ker f'(v)$ of V. The proof is left to the reader (Exercise 1.4.4). This shows that Definition 1.4.2 formalizes the idea of the tangent space at a point of a level set, indeed.

Next, we discuss an equivalent, more algebraic view on tangent vectors. Consider the algebra $C^k(M)$ of C^k-functions on M. Let $X_m \in \mathrm{T}_m M$ and let $\gamma \in K_m(M)$ be a curve representing X_m. By analogy with the directional derivative of analysis in $\mathbb{R}^n$ we define the directional derivative of $f \in C^k(M)$ along the curve γ by $\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0}(f \circ \gamma)(t)$. Due to

$$\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0}(f \circ \gamma)(t) = \left[\left(f \circ \kappa^{-1}\right)'\left(\kappa(m)\right)\right] \cdot \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0}(\kappa \circ \gamma)(t)$$

[6] If there is no danger of confusion, we will usually omit the chart label and just write X_m^i.

and (1.4.1), for another curve $\tilde{\gamma}$ representing X_m there holds

$$\frac{\mathrm{d}}{\mathrm{d}t}{}_{\restriction 0}(f \circ \gamma)(t) = \frac{\mathrm{d}}{\mathrm{d}t}{}_{\restriction 0}(f \circ \tilde{\gamma})(t).$$

That is, the directional derivative depends on X_m only. Therefore, X_m defines a linear mapping, denoted by the same symbol,

$$X_m : C^k(M) \to \mathbb{R}, \qquad X_m(f) := \frac{\mathrm{d}}{\mathrm{d}t}{}_{\restriction 0}(f \circ \gamma)(t), \tag{1.4.10}$$

where γ is some curve representing X_m. For $f, g \in C^k(M)$, the product rule yields

$$X_m(f \cdot g) = X_m(f)g(m) + f(m)X_m(g), \tag{1.4.11}$$

that means, X_m is a derivation at m.

Definition 1.4.4 (Derivation at a point) Let M be a C^k-manifold and let $m \in M$. A derivation at m is a linear mapping $D_m : C^k(M) \to \mathbb{R}$ satisfying

$$D_m(f \cdot g) = D_m(f)g(m) + f(m)D_m(g) \quad \text{for all } f, g \in C^k(M). \tag{1.4.12}$$

The set of all derivations at m will be denoted by $\mathrm{D}_m M$. It carries the structure of a real linear space. Assigning to a tangent vector its directional derivative (1.4.10) defines an injection of $\mathrm{T}_m M$ into $\mathrm{D}_m M$, which preserves the linear structure,

$$(X_m + Y_m)(f) = X_m(f) + Y_m(f) \tag{1.4.13}$$

(Exercise 1.4.7). Thus, $\mathrm{T}_m M$ may be identified with a linear subspace of $\mathrm{D}_m M$. Via the isomorphism F^κ_m, the standard basis $\{\mathbf{e}_i\}$ of $\mathbb{R}^n$ induces a basis in $\mathrm{T}_m M$. We determine the derivations which correspond to the elements of this basis. Due to (1.4.3) and (1.4.4), for $f \in C^k(M)$,

$$\Big(\big(F^\kappa_m\big)^{-1}(\mathbf{e}_i)\Big)(f) = \frac{\mathrm{d}}{\mathrm{d}t}{}_{\restriction 0}\Big(f \circ \kappa^{-1} \circ \big(\kappa(m) + t\mathbf{e}_i\big)\Big) = \frac{\partial(f \circ \kappa^{-1})}{\partial x^i}\big(\kappa(m)\big).$$

Thus, the derivations[7]

$$\partial^\kappa_{i,m} : C^k(M) \to \mathbb{R}, \qquad \partial^\kappa_{i,m}(f) := \frac{\partial(f \circ \kappa^{-1})}{\partial x^i}\big(\kappa(m)\big), \quad i = 1, \dots, n, \tag{1.4.14}$$

form a basis in $\mathrm{T}_m M$. By construction, the coefficients of a tangent vector X_m with respect to this basis are given by the components of the local representative $\mathbf{X}^\kappa_m$,

$$X_m = X^{\kappa,i}_m \, \partial^\kappa_{i,m}. \tag{1.4.15}$$

---

[7] As for the local representative of a tangent vector, we will usually omit the chart label and just write $\partial_{i,m}$.

Moreover, (1.4.8) implies

$$X_m^{\kappa,i} = X_m\big(\kappa^i\big). \tag{1.4.16}$$

If (V, ρ) is another local chart at m, the two bases $\{\partial^{\kappa}_{i,m}\}$ and $\{\partial^{\rho}_{i,m}\}$ of $\mathrm{T}_m M$ are related by (Exercise 1.4.3)

$$\partial^{\rho}_{j,m} = \big[\big(\kappa \circ \rho^{-1}\big)'\big(\rho(m)\big)\big]^i_j \partial^{\kappa}_{i,m}. \tag{1.4.17}$$

Proposition 1.4.5 *Let M be a C^k-manifold, let $m \in M$ and let D_m be a derivation at m. For arbitrary $f, g \in C^k(M)$, the following holds.*

1. *If f is constant, then $D_m(f) = 0$.*
2. *If $f_{\restriction U} = g_{\restriction U}$ for some neighbourhood U of m, then $D_m(f) = D_m(g)$.*

Proof 1. Write $f = \lambda\, 1$, where $\lambda = f(m)$ and 1 denotes the constant function with value 1. Then,

$$\begin{aligned} D_m(f) &= \lambda D_m(1) = \lambda D_m(1 \cdot 1) = \lambda\big(D_m(1) \cdot 1 + 1 \cdot D_m(1)\big) \\ &= 2\lambda D_m(1) = 2 D_m(f). \end{aligned}$$

2. By assumption, $(f - g)_{\restriction U} = 0$. There exists a C^k-function h on M such that $h(m) = 1$ and $h_{\restriction M \setminus U} = 0$. For example, choose a smooth function on $\mathbb{R}^n$ which has the value 1 at the origin and vanishes outside some ε-ball, transport it by an appropriate local chart to U and extend it by 0 to a function on M. Then, $(f - g)h = 0$ and hence

$$0 = D_m\big((f - g)h\big) = D_m(f - g)h(m) + \big(f(m) - g(m)\big)D_m(h) = D_m(f - g).$$

□

Remark 1.4.6 The notion of derivation at a point is related to the notion of derivation in the theory of modules over algebras as follows. Let $\mathfrak{A}$ be an algebra and let $\mathfrak{M}$ be a bimodule over $\mathfrak{A}$. A linear mapping $D : \mathfrak{A} \to \mathfrak{M}$ is called a derivation if it satisfies

$$D(ab) = D(a)b + aD(b), \quad a, b \in \mathfrak{A}.$$

Definition 1.4.4 is obtained from this more general definition by choosing $C^k(M)$ for $\mathfrak{A}$ and $\mathbb{R}$ for $\mathfrak{M}$, with the bimodule structure being defined by $f \cdot x := f(m)x$ and $x \cdot f := x f(m)$, where $f \in C^k(M)$, $x \in \mathbb{R}$.

We now show that in the case $k = \infty$, every derivation at m comes from a tangent vector. That is, for smooth manifolds, the subspace $\mathrm{T}_m M$ in fact coincides with $\mathrm{D}_m M$ and the linear spaces $\mathrm{T}_m M$ and $\mathrm{D}_m M$ may be viewed as different realizations of the same object.

Proposition 1.4.7 *For every point m of a C^∞-manifold M there holds $\mathrm{T}_m M = \mathrm{D}_m M$.*

Proof Let (U, κ) be a local chart at a chosen point m_0. It suffices to show that every $D_{m_0} \in \mathrm{D}_{m_0}M$ can be written as a linear combination of $\partial^\kappa_{1,m_0}, \dots, \partial^\kappa_{n,m_0}$. Let $f \in C^\infty(M)$. According to Taylor's Theorem, in a neighbourhood V of $\mathbf{x}_0 = \kappa(m_0)$, the local representative $\tilde{f} := f \circ \kappa^{-1}$ can be written in the form

$$\tilde{f}(\mathbf{x}) = \tilde{f}(\mathbf{x}_0) + \frac{\partial \tilde{f}}{\partial x^i}(\mathbf{x}_0)\big(x^i - x_0^i\big) + \tilde{f}_{ij}(\mathbf{x})\big(x^i - x_0^i\big)\big(x^j - x_0^j\big), \tag{1.4.18}$$

where $\tilde{f}_{ij}$ are smooth functions on V, given by

$$\tilde{f}_{ij}(\mathbf{x}) = \int_0^1 \mathrm{d}s\,(1-s)\,\frac{\partial^2 \tilde{f}}{\partial x^i \partial x^j}\big((1-s)\mathbf{x}_0 + s\mathbf{x}\big).$$

Since

$$\frac{\partial \tilde{f}}{\partial x^i}(\mathbf{x}_0) = \frac{\partial(f \circ \kappa^{-1})}{\partial x^i}\big(\kappa(m_0)\big) = \partial^\kappa_{i,m_0}(f),$$

Formula (1.4.18) yields the following decomposition on $\kappa^{-1}(V) \subset U$:

$$\begin{aligned} f(m) =& f(m_0) + \partial^\kappa_{i,m_0}(f)\big(\kappa^i(m) - x_0^i\big) \\ &+ \big(\tilde{f}_{ij} \circ \kappa(m)\big)\big(\kappa^i(m) - x_0^i\big)\big(\kappa^j(m) - x_0^j\big). \end{aligned} \tag{1.4.19}$$

There exist open neighbourhoods W_1, W_2 of m_0 and a smooth function h on M such that $\overline{W}_1 \subset W_2$, $\overline{W}_2 \subset \kappa^{-1}(V)$, $h_{\restriction W_1} = 1$ and $h_{\restriction M \setminus W_2} = 0$. Multiply κ^i and $\tilde{f}_{ij} \circ \kappa$ on W_2 by h and extend the resulting functions by 0 to M. With κ^i and $\tilde{f}_{ij} \circ \kappa$ modified in this way, (1.4.19) holds on W_1 and all the functions involved are from $C^\infty(M)$. The latter ensures that we can apply D_{m_0} to both sides of this equation. The second assertion of Proposition 1.4.5 ensures that after application of D_{m_0}, both sides are still equal. Then, a brief calculation using the first assertion of this proposition yields $D_{m_0}(f) = D_{m_0}(\kappa^i)\partial^\kappa_{i,m_0}(f)$. Since this holds for all $f \in C^\infty(M)$, the assertion follows. □

Remark 1.4.8

1. Let M be a smooth manifold, let $m \in M$ and let (U, κ) be a local chart at m. As a by-product of the proof of Proposition 1.4.7 we note that for a given derivation $D_m \in \mathrm{D}_m M$, the tangent vector corresponding to D_m can be represented by the smooth curve $t \mapsto \kappa^{-1}(\kappa(m) + tD_m(\kappa^i)\mathbf{e}_i)$.
2. The proof of Proposition 1.4.7 does not work in the case of finite k. Indeed, for $k = 1$ there is no decomposition of f like (1.4.19) and for $k \geq 2$, the functions $\tilde{f}_{ij} \circ \kappa$ are of class C^{k-2}, hence one cannot apply a derivation of C^k-functions to them. As it turns out, this failure cannot be repaired. In fact, for finite k there is a big difference between $\mathrm{T}_m M$ and $\mathrm{D}_m M$, because $\mathrm{D}_m M$ can be shown to be infinite-dimensional in this case [229]. To summarize, in the case of finite k, $\mathrm{T}_m M$ is identified with a proper subspace of $\mathrm{D}_m M$. This subspace consists of

the directional derivatives defined by C^k-curves through m and is spanned by the derivations $\partial^\kappa_{i,m}$, $i = 1, \dots, n$, associated with an arbitrary local chart (U, κ) at m.

To conclude this section, we briefly discuss the dual vector space of $\mathrm{T}_m M$. Recall that the dual vector space V^* of a vector space V over a field $\mathbb{K}$ consists of all $\mathbb{K}$-linear mappings $\xi : V \to \mathbb{K}$. With respect to the operations

$$(k\xi_1 + \xi_2)(v) := k\xi_1(v) + \xi_2(v), \quad k \in \mathbb{K}, v \in V,$$

V^* is a vector space over $\mathbb{K}$ of the same dimension as V.

Definition 1.4.9 (Cotangent space) Let M be a C^k-manifold and let $m \in M$. The dual vector space of $\mathrm{T}_m M$ is called the cotangent space of M at m and is denoted by $\mathrm{T}^*_m M$. Its elements are called cotangent vectors or covectors in m.

Covectors will be denoted by α_m, β_m etc. Evaluation of covectors on tangent vectors will often be written in the form of a pairing: $\alpha_m(X_m) \equiv \langle \alpha_m, X_m \rangle$. Every $f \in C^k(M)$ defines a covector

$$(\mathrm{d}f)_m(X_m) := X_m(f), \tag{1.4.20}$$

which is called the differential of f at m. For example, for the coordinate functions κ^i of a local chart (U, κ) at m we obtain

$$\left(\mathrm{d}\kappa^i\right)_m(X_m) = X_m\left(\kappa^i\right) = X_m^{\kappa,i}. \tag{1.4.21}$$

As a consequence, the set of differentials $\{(\mathrm{d}\kappa^1)_m, \dots, (\mathrm{d}\kappa^n)_m\}$ yields the basis in $\mathrm{T}^*_m M$ which is dual to the basis $\{\partial^\kappa_{1,m}, \dots \partial^\kappa_{n,m}\}$ in $\mathrm{T}_m M$. In particular, every covector can be written in the form

$$\alpha_m = \alpha^\kappa_{i,m}\left(\mathrm{d}\kappa^i\right)_m, \qquad \alpha^\kappa_{i,m} = \alpha_m\left(\partial^\kappa_{i,m}\right), \tag{1.4.22}$$

and the system of real numbers $(\alpha^\kappa_{1,m}, \dots, \alpha^\kappa_{n,m})$, viewed as an element of $\mathbb{R}^{n*}$, is called the local representative of α_m with respect to the local chart (U, κ). As for tangent vectors, we will usually omit the chart label and just write $\alpha_{i,m}$.

Let (V, ρ) be a second local chart at m. In view of (1.4.21) and (1.4.22), we can read off the transformation laws for the basis and for the coefficients from (1.4.9) and (1.4.17), respectively:

$$\mathrm{d}\rho^i = \left[\left(\rho \circ \kappa^{-1}\right)'\left(\kappa(m)\right)\right]^i_j \mathrm{d}\kappa^j, \qquad \alpha^\rho_{i,m} = \left[\left(\kappa \circ \rho^{-1}\right)'\left(\rho(m)\right)\right]^j_i \alpha^\kappa_{j,m}. \tag{1.4.23}$$

Exercises

1.4.1 Prove Lemma 1.4.1.

1.4.2 Prove Formula (1.4.5).

1.4.3 Prove Formula (1.4.17).
1.4.4 Prove the assertion of Example 1.4.3/2.
1.4.5 For every point of one of the level sets of Exercise 1.2.5, determine the tangent space as an affine subspace of $\mathbb{R}^3$.
1.4.6 For each of the classical groups of Example 1.2.6, determine the tangent space at the unit element as a subspace of the corresponding matrix algebra $\mathrm{M}_n(\mathbb{K})$. Show that all the subspaces so obtained are closed under the operation of taking the commutator of matrices. (Later on, we will see that these subspaces, together with this operation, are realizations of the Lie algebras of the classical groups.)
1.4.7 Prove Formula (1.4.13).
1.4.8 Prove Remark 1.4.8/1.

1.5 Tangent Mapping

In this section, we generalize the notion of the derivative of a mapping from classical calculus to the theory of manifolds. Thereafter, we extend the basic theorems of classical calculus to manifolds.

In the sequel, assume $k \geq 1$. Let M and N be C^k-manifolds and let $\Phi : M \to N$ be a C^k-mapping. Φ maps a C^k-curve γ in M to the curve $\Phi \circ \gamma$ in N, which is of class C^k again. If $\gamma \in K_m(M)$, then $\Phi \circ \gamma \in K_{\Phi(m)}(N)$. If $\gamma_1 \sim \gamma_2$, then $\Phi \circ \gamma_1 \sim \Phi \circ \gamma_2$, because for local charts (U, κ) on M at m and (V, ρ) on N at $\Phi(m)$ there holds

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Big\restriction_0 (\rho \circ \Phi \circ \gamma_1)(t) &= \big(\rho \circ \Phi \circ \kappa^{-1}\big)'\big(\kappa(m)\big) \cdot \frac{\mathrm{d}}{\mathrm{d}t}\Big\restriction_0 (\kappa \circ \gamma_1)(t) \\ &= \big(\rho \circ \Phi \circ \kappa^{-1}\big)'\big(\kappa(m)\big) \cdot \frac{\mathrm{d}}{\mathrm{d}t}\Big\restriction_0 (\kappa \circ \gamma_2)(t) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big\restriction_0 (\rho \circ \Phi \circ \gamma_2)(t).\end{aligned}$$

Thus, Φ induces a mapping of the tangent spaces.

Definition 1.5.1 (Tangent mapping at a point) Let M and N be C^k-manifolds, let $\Phi \in C^k(M, N)$ and let $m \in M$. The mapping $\Phi'_m : \mathrm{T}_m M \to \mathrm{T}_{\Phi(m)} N$, defined by

$$\Phi'_m(X_m) := [\Phi \circ \gamma],$$

where γ is some curve representing X_m, is called the tangent mapping of Φ at m.

The tangent mapping has the following properties.

Proposition 1.5.2 *Let M and N be C^k-manifolds, let $\Phi \in C^k(M, N)$ and let $m \in M$.*

1. *Φ'_m is linear.*
2. $(\mathrm{id}_M)'_m = \mathrm{id}_{\mathrm{T}_m M}$.
3. *If P is another C^k-manifold and $\Psi \in C^k(N, P)$, then $(\Psi \circ \Phi)'_m = \Psi'_{\Phi(m)} \circ \Phi'_m$.*
4. *If Φ is a diffeomorphism, then Φ'_m is bijective and one has $(\Phi'_m)^{-1} = (\Phi^{-1})'_{\Phi(m)}$.*

Assertion 3 is referred to as the chain rule.

Proof 1. Choose local charts (U, κ) on M at m and (V, ρ) on N at $\Phi(m)$ and consider the isomorphisms $F^\kappa_m : \mathrm{T}_m M \to \mathbb{R}^{\dim M}$ and $F^\rho_m : \mathrm{T}_{\Phi(m)} N \to \mathbb{R}^{\dim N}$, defined by (1.4.2). It suffices to show that the composition

$$F^\rho_m \circ \Phi'_m \circ \left(F^\kappa_m\right)^{-1} : \mathbb{R}^{\dim M} \to \mathbb{R}^{\dim N}$$

is linear. To see this, let $\mathbf{X} \in \mathbb{R}^{\dim M}$ and let $\tilde{\gamma}$ denote the curve $t \mapsto \kappa(m) + t\mathbf{X}$. Then,

$$F^\rho_m \circ \Phi'_m \circ \left(F^\kappa_m\right)^{-1}(\mathbf{X}) = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} \left(\rho \circ \Phi \circ \kappa^{-1} \circ \tilde{\gamma}\right)(t) = \left(\rho \circ \Phi \circ \kappa^{-1}\right)'\left(\kappa(m)\right)\mathbf{X},$$

that is, $F^\rho_m \circ \Phi'_m \circ (F^\kappa_m)^{-1}$ is given by the derivative of the local representative of Φ at $\kappa(m)$.

2. This follows immediately from the definition.

3. Let $X_m \in \mathrm{T}_m M$ be represented by $\gamma \in K_m(M)$. Then,

$$(\Psi \circ \Phi)'_m(X_m) = [\Psi \circ \Phi \circ \gamma] = \Psi'_{\Phi(m)}\left([\Phi \circ \gamma]\right) = \Psi'_{\Phi(m)} \circ \Phi'_m(X_m).$$

4. Assertions 2 and 3 imply $(\Phi^{-1})'_{\Phi(m)} \circ \Phi'_m = (\Phi^{-1} \circ \Phi)'_m = (\mathrm{id}_M)'_m = \mathrm{id}_{\mathrm{T}_m M}$ and, analogously, $\Phi'_m \circ (\Phi^{-1})'_{\Phi(m)} = \mathrm{id}_{\mathrm{T}_{\Phi(m)} N}$. □

The tangent mapping can be expressed on the level of derivations, too. Let $X_m \in \mathrm{T}_m M$ be represented by $\gamma \in K_m(M)$. Then $\Phi'_m X_m$ is the derivation at $\Phi(m)$ on N which is given by the directional derivative along the curve $\Phi \circ \gamma$:

$$\left(\Phi'_m X_m\right)(f) = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} (f \circ \Phi \circ \gamma)(t) = X_m(f \circ \Phi), \quad f \in C^k(N). \tag{1.5.1}$$

The assignment of $f \circ \Phi$ to f defines a mapping

$$\Phi^* : C^k(N) \to C^k(M), \qquad \Phi^* f := f \circ \Phi. \tag{1.5.2}$$

In this notation, we have

$$\Phi'_m X_m = X_m \circ \Phi^*. \tag{1.5.3}$$

Remark 1.5.3

1. The mapping Φ^* is called the pull-back (of functions) by Φ. Later on, it will be generalized to differential forms. Φ^* is a homomorphism of algebras and satisfies

$$\mathrm{id}_M^* = \mathrm{id}_{C^k(M)}, \qquad (\Psi \circ \Phi)^* = \Phi^* \circ \Psi^*$$

for all $\Phi \in C^k(M, N)$ and $\Psi \in C^k(N, P)$. In the case of finite k, where $\mathrm{T}_m M$ is a proper subspace of $\mathrm{D}_m M$ according to Remark 1.4.8/2, the right hand side of (1.5.3) may be taken as the extension of the tangent mapping from $\mathrm{T}_m M$ to $\mathrm{D}_m M$.

2. From the proof of point 1 of Proposition 1.5.2 we conclude that for local charts (U, κ) on M at m and (V, ρ) on N at $\Phi(m)$ one has

$$\left(\Phi'_m X_m\right)^{\rho,i}_{\phi(m)} = \left(\rho \circ \Phi \circ \kappa^{-1}\right)'\left(\kappa(m)\right)^i_j X_m^{\kappa,j}. \tag{1.5.4}$$

That is, locally the tangent mapping of Φ is given by the derivative (matrix of partial derivatives) of the local representative $\Phi_{\kappa,\rho} = \rho \circ \Phi \circ \kappa^{-1}$ at $\kappa(m)$.

Example 1.5.4 Let M and N be open subsets of the finite-dimensional real vector spaces V and W, respectively. Let $\Phi \in C^k(M, N)$ and $v \in M$. We determine the tangent mapping Φ'_v using the natural identifications of $\mathrm{T}_v M$ with V and of $\mathrm{T}_{\Phi(v)} N$ with W, cf. Example 1.4.3/1. Since $u \in V$ corresponds to the tangent vector $X_v \in \mathrm{T}_m M$, represented by the curve $t \mapsto \gamma(t) := v + tu$, it is mapped by Φ'_v as follows:

$$u \mapsto \Phi'_v[\gamma] = [\Phi \circ \gamma] = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \Phi(v + tu) = \Phi'(v) \cdot u,$$

where $\Phi'(v)$ denotes the ordinary derivative of mappings between open subsets of finite-dimensional real vector spaces. This shows that the notion of tangent mapping generalizes the notion of derivative of calculus in $\mathbb{R}^n$.

Example 1.5.5 Let M be a C^k-manifold, let $m \in M$ and let $f \in C^k(M)$. We calculate the tangent mapping $f'_m : \mathrm{T}_m M \to \mathrm{T}_{f(m)}\mathbb{R}$ under the natural identification of $\mathrm{T}_{f(m)}\mathbb{R}$ with $\mathbb{R}$, see Example 1.4.3/1. Let $X_m \in \mathrm{T}_m M$ be represented by the curve γ. Then,

$$f'_m(X_m) = [f \circ \gamma] = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 (f \circ \gamma)(t) = X_m(f) = (\mathrm{d}f)_m(X_m),$$

see (1.4.20). Hence, the tangent mapping f'_m is given by the differential $(\mathrm{d}f)_m$.

Example 1.5.6 This example explains the concept of a tangent vector of a curve. Let M be a C^k-manifold, let $I \subset \mathbb{R}$ be an open interval, let $\gamma : I \to M$ be a C^k-curve and let $t \in I$. The tangent vector of I at t represented by the curve $s \mapsto t + s$

corresponds to the derivation $f \mapsto \frac{\mathrm{d}}{\mathrm{d}t}\restriction_t f(t)$. Therefore, it will be denoted by $\frac{\mathrm{d}}{\mathrm{d}t}\restriction_t$. The tangent vector $\dot{\gamma}(t)$ of the curve γ at t is defined by

$$\dot{\gamma}(t) := \gamma'_t\left(\frac{\mathrm{d}}{\mathrm{d}t}\restriction_t\right).$$

It is represented by the curve $s \mapsto \gamma(t+s)$. On the level of derivations, $\dot{\gamma}(t)$ corresponds to

$$\dot{\gamma}(t)(f) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_t (f \circ \gamma).$$

Let us determine the local representative of $\dot{\gamma}(t)$ with respect to the identical chart on I and a local chart (U, κ) on M. Since the local representative of the unit tangent vector at t is given by 1, (1.5.4) yields

$$\big(\dot{\gamma}(t)\big)^{\kappa}_{\gamma(t)} = (\kappa \circ \gamma)'(t) \cdot 1 = \frac{\mathrm{d}}{\mathrm{d}s}\restriction_t (\kappa \circ \gamma)(s). \tag{1.5.5}$$

Thus, the local representative of $\dot{\gamma}(t)$ is given by the tangent vector of the curve $\kappa \circ \gamma$ in $\mathbb{R}^n$.

We now carry over the Inverse Function Theorem, the Implicit Function Theorem and the Constant Rank Theorem of calculus in $\mathbb{R}^n$ to manifolds.

Theorem 1.5.7 (Inverse Mapping Theorem) *Let M and N be C^k-manifolds, let $\Phi \in C^k(M, N)$ and let $m \in M$. If the tangent mapping $\Phi'_m : \mathrm{T}_m M \to \mathrm{T}_{\Phi(m)} N$ is bijective, there exist open neighbourhoods U of m in M and V of $\Phi(m)$ in N such that Φ restricts to a diffeomorphism of class C^k from U onto V.*

Proof Choose local charts (U, κ) on M at m and (V, ρ) on N at $\Phi(m)$. Consider the local representative $\Phi_{\rho,\kappa} = \rho \circ \Phi \circ \kappa^{-1}$. By the chain rule,

$$(\Phi_{\rho,\kappa})'_{\kappa(m)} = \rho'_{\Phi(m)} \circ \Phi'_m \circ \big(\kappa^{-1}\big)'_{\kappa(m)}.$$

By assertion 4 of Proposition 1.5.2, $\rho'_{\Phi(m)}$ and $(\kappa^{-1})'_{\kappa(m)}$ are bijective. Hence, due to the assumption, $(\Phi_{\rho,\kappa})'_{\kappa(m)}$ is bijective. Then, the Inverse Function Theorem of classical calculus implies that U and V can be shrunk so that $\Phi_{\rho,\kappa}$ restricts to a diffeomorphism from $\kappa(U)$ onto $\rho(V)$. Then, $\Phi(U) = V$ and the restricted mapping $\tilde{\Phi} : U \to V$ is a diffeomorphism. □

Corollary 1.5.8 *Let M and N be C^k-manifolds and let $\Phi \in C^k(M, N)$. If Φ is bijective and Φ'_m is bijective for all $m \in M$, then Φ is a diffeomorphism of class C^k.*

Proof Since Φ is bijective, it has an inverse $\Phi^{-1} : N \to M$. Since Φ'_m is bijective for all $m \in M$, Theorem 1.5.7 says, in particular, that every point of N has a neighbourhood on which Φ^{-1} is differentiable of class C^k. This yields the assertion. □

Remark 1.5.9 In Corollary 1.5.8 it suffices to assume that Φ'_m is injective for all $m \in M$. This is due to the fact that bijectivity and differentiability of Φ, together with the property of manifolds to be second countable, imply that M and N have the same dimension, see Exercise 6 of Chap. 1 in [302] for instructions on a proof.

Theorem 1.5.10 (Implicit Mapping Theorem) *Let M_1, M_2 and N be C^k-manifolds and let $\Phi \in C^k(M_1 \times M_2, N)$. Let $(m_{01}, m_{02}) \in M_1 \times M_2$ and $p = \Phi(m_{01}, m_{02})$. If the tangent mapping at m_{02} of the induced mapping*

$$\Phi_{m_{01}} : M_2 \to N, \qquad \Phi_{m_{01}}(m_2) := \Phi(m_{01}, m_2)$$

is bijective, there exist open neighbourhoods U_i of m_{0i} in M_i and a C^k-mapping $\Psi : U_1 \to U_2$ such that for all $(m_1, m_2) \in U_1 \times U_2$ there holds $\Phi(m_1, m_2) = p$ iff $m_2 = \Psi(m_1)$.

Proof Choose local charts at m_{01}, m_{02} and p and apply the Implicit Function Theorem of calculus in $\mathbb{R}^n$ to the corresponding local representative of Φ. The details are left to the reader (Exercise 1.5.2). □

Theorem 1.5.11 (Constant Rank Theorem) *Let M and N be C^k-manifolds, let $\Phi \in C^k(M, N)$ and let $m_0 \in M$. If the linear mapping $\Phi'_m : \mathrm{T}_m M \to \mathrm{T}_{\Phi(m)} N$ has constant rank r for all m in a neighbourhood of $m_0 \in M$, there exist local charts (U, κ) on M at m_0 and (V, ρ) on N at $\Phi(m_0)$ such that the local representative $\Phi_{\rho,\kappa}$ coincides with the restriction to $U \cap \Phi^{-1}(V)$ of the mapping*

$$\mathbb{R}^r \times \mathbb{R}^{\dim M - r} \ni (\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x}, 0) \in \mathbb{R}^r \times \mathbb{R}^{\dim N - r}.$$

Proof Choose local charts (U, κ) on M at m_0 and (V, ρ) on N at $\Phi(m_0)$, apply the ordinary Constant Rank Theorem of calculus in $\mathbb{R}^n$ to the local representative $\Phi_{\rho,\kappa}$ and redefine (U, κ) and (V, ρ) accordingly. The details are left to the reader (Exercise 1.5.2). □

According to the Inverse Mapping Theorem, bijectivity of the tangent mapping has important consequences for the local behaviour of the mapping itself. By weakening the requirement of bijectivity to injectivity or surjectivity, one arrives at the notions of immersion and submersion.

Definition 1.5.12 (Immersion and submersion) Let M and N be C^k-manifolds, let $\Phi \in C^k(M, N)$ and let $m \in M$. Φ is called an immersion at m if Φ'_m is injective. It is called a submersion at m if Φ'_m is surjective. It is called an immersion (submersion) if it is an immersion (submersion) at every $m \in M$.

Equivalent characterizations are

$$\begin{aligned} &\Phi \text{ is an immersion at } m \text{ iff } \operatorname{rank} \Phi'_m = \dim M \text{ iff } \ker \Phi'_m = 0, \\ &\Phi \text{ is a submersion at } m \text{ iff } \operatorname{rank} \Phi'_m = \dim N \text{ iff } \operatorname{im} \Phi'_m = \mathrm{T}_{\Phi(m)} N, \end{aligned}$$

where $\operatorname{rank} \Phi'_m$ denotes the rank of the linear mapping Φ'_m. To be an immersion or a submersion is a local property, whereas to be injective or surjective is a global property. Therefore, an immersion need not be injective and a submersion need not be surjective. Conversely, if Φ is injective it need not be an immersion and if it is surjective it need not be a submersion.

Example 1.5.13 In the examples to follow, the tangent mapping is given by the ordinary derivative of mappings of $\mathbb{R}^n$, see Example 1.5.4.

1. Let $\Phi : \mathbb{R} \to \mathbb{R}^2$, $\Phi(x) := (\cos(x), \sin(x))$. Since $\Phi'_x = (-\sin(x), \cos(x)) \neq 0$ for all $x \in \mathbb{R}$, Φ is an immersion. It is not injective though.
2. Let $\Phi : \mathbb{R} \to \mathbb{R}$, $\Phi(x) := \mathrm{e}^x$. Since $\Phi'_x = \mathrm{e}^x \neq 0$ for all $x \in \mathbb{R}$, Φ is a submersion. It is not surjective though.
3. The mapping $\Phi : \mathbb{R} \to \mathbb{R}$, defined by $\Phi(x) := x^3$, is injective and surjective. The tangent mapping is $\Phi'_x = 3x^2$. Since it is neither injective nor surjective at $x = 0$, Φ is neither an immersion nor a submersion.

Remark 1.5.14 Let M, N and P be C^k-manifolds and let $\Phi \in C^k(M, N)$ and $\Psi \in C^k(N, P)$. By means of the chain rule one can show the following.

1. If Φ and Ψ are immersions (submersions), $\Psi \circ \Phi$ is an immersion (submersion).
2. If $\Psi \circ \Phi$ is an immersion, so is Φ. If it is a submersion, so is Ψ.

These assertions hold also pointwise.

Proposition 1.5.15 *Let M and N be C^k-manifolds and let $\Phi \in C^k(M, N)$. The set of points of M at which Φ is an immersion (submersion) is open in M.*

Proof For a natural number r, define $M_r := \{m \in M : \operatorname{rank} \Phi'_m \geq r\}$. Since for all $m \in M$ there holds $\operatorname{rank} \Phi'_m \leq \dim M$, the set of points at which Φ is an immersion coincides with $M_{\dim M}$. Analogously, since for all $m \in M$ one has $\operatorname{rank} \Phi'_m \leq \dim N$, the set of points at which Φ is a submersion coincides with $M_{\dim N}$. Therefore, it suffices to show that M_r is open in M for all natural numbers r. Let $m_0 \in M_r$. Choose local charts (U, κ) on M at m_0 and (V, ρ) on N at $\Phi(m_0)$ such that $\Phi(U) \subset V$. Then, $\operatorname{rank} \Phi'_m = \operatorname{rank}(\rho \circ \Phi \circ \kappa^{-1})'(\kappa(m))$ for all $m \in U$. Now,

$$\operatorname{rank}\big(\rho \circ \Phi \circ \kappa^{-1}\big)'\big(\kappa(m_0)\big) \geq r$$

is equivalent to the existence of a minor $D_r : L(\mathbb{R}^{\dim M}, \mathbb{R}^{\dim N}) \to \mathbb{R}$ of rank r satisfying $D_r((\rho \circ \Phi \circ \kappa^{-1})'(\kappa(m_0))) \neq 0$. Since the mapping

$$\mathbf{x} \mapsto D_r\big(\big(\rho \circ \Phi \circ \kappa^{-1}\big)'(\mathbf{x})\big)$$

is continuous, there exists a neighbourhood $\tilde{U}$ of $\kappa(m_0)$ such that

$$D_r\big(\big(\rho \circ \Phi \circ \kappa^{-1}\big)'(\mathbf{x})\big) \neq 0$$

for all $\mathbf{x} \in \tilde{U}$. Then, $\kappa^{-1}(\tilde{U})$ is a neighbourhood of m_0 in M and rank $\Phi'_m \geq r$ for all $m \in \kappa^{-1}(\tilde{U})$. □

Remark 1.5.16 (Basic properties of submersions) In the following, proofs are left to the reader (Exercise 1.5.4). According to the Constant Rank Theorem 1.5.11, locally, submersions look like the natural projection to a factor of a direct product. This has the following consequences.

1. Submersions are open mappings.
2. Submersions admit local sections. This means the following. Let $\Phi : M \to N$ be a submersion of class C^k. For every $p \in N$, there exists an open neighbourhood U and a C^k-mapping $s : U \to M$ such that $\Phi \circ s = \mathrm{id}_U$. The mapping s is called a local section of Φ at p.

   The existence of local sections implies, in turn, the following.

3. Let $\Phi : M \to N$ be a surjective submersion of class C^k and let $\Psi : M \to P$ a C^k-mapping. If there exists a mapping $\tilde{\Psi} : N \to P$ such that $\tilde{\Psi} \circ \Phi = \Psi$, then this mapping is unique and of class C^k.
4. Let M be a C^k-manifold, N a set and $\Phi : M \to N$ a surjective mapping. If N admits a C^k-structure such that Φ is a submersion, this structure is unique.

Next, we generalize the notions of regular and critical point and regular and critical value to differentiable mappings between manifolds and state Sard's Theorem.

Definition 1.5.17 Let M and N be C^k-manifolds and let $\Phi \in C^k(M, N)$. A point $m \in M$ is called regular if Φ is a submersion at m. Otherwise, m is called singular or critical for Φ. A point $p \in N$ is called a regular value of Φ if $\Phi^{-1}(p) = \varnothing$ or if all points $m \in \Phi^{-1}(p)$ are regular. Otherwise, p is called a singular or critical value of Φ.

To state Sard's Theorem for manifolds, one needs the notion of Lebesgue measure on a manifold. Such measures[8] are constructed by using an atlas and a subordinate partition of unity, see [73, §16.22.2].

Theorem 1.5.18 (Sard) *Let M and N be C^k-manifolds where $k > \dim M - \dim N$ and let $\Phi \in C^k(M, N)$. The set of critical values of Φ has measure zero with respect to the Lebesgue measures on N. The set of regular values is dense in N.*

Proof See [130, Thm. 3.1.3], or [73, §16.23.1] for the C^∞-case. □

Corollary 1.5.19 *Let M and N be C^k-manifolds and let $\Phi : M \to N$ be of class C^k. If $\dim M < \dim N$, then $\Phi(M)$ is a set of measure zero in N and $N \setminus \Phi(M)$ is dense.*

[8] All such measures are equivalent, that is, they have the same sets of measure zero.

Proof Due to $\dim M < \dim N$, Sard's Theorem can be applied. For the same reason, all elements of M are critical points for Φ and hence all elements of $\Phi(M)$ are critical values of Φ. This yields the assertion. □

Exercises

1.5.1 Let M and N be C^k-manifolds and let $\Phi \in C^k(M, N)$. Show that if M is connected and if $\Phi'_m = 0$ for all $m \in M$, Φ is constant.

1.5.2 Provide the details for the proofs of the Implicit Mapping Theorem 1.5.10 and the Constant Rank Theorem 1.5.11.

1.5.3 Prove the statements of Remark 1.5.14.

1.5.4 Prove the properties of submersions stated in Remark 1.5.16.

1.5.5 Let M, N and P be C^k-manifolds and let $\Phi \in C^k(M, N)$ and $\Psi \in C^k(N, P)$. Show the following.
   (a) If Φ is a submersion and Ψ is an immersion, then $\Psi \circ \Phi : M \to P$ has locally constant rank. ($\Psi \circ \Phi$ is referred to as a subimmersion.)
   (b) If, on the contrary, Φ is an immersion and Ψ is a submersion, then $\Psi \circ \Phi$ need not have locally constant rank.
   Hint. Consider the mappings $\Phi : \mathbb{R} \to \mathbb{R}^3$, $\Phi(t) := (t, t^3, t^3)$ and $\Psi : \mathbb{R}^3 \to \mathbb{R}^2$, $\Psi(x, y, z) := (y, z)$.

1.6 Submanifolds

Let $k \geq 1$ and let N be a C^k-manifold.

Definition 1.6.1 (Submanifold) A C^k-submanifold of N is a pair (M, φ), where M is a C^k-manifold and $\varphi : M \to N$ is an injective immersion of class C^k. Submanifolds (M_1, φ_1) and (M_2, φ_2) are said to be equivalent if there exists a diffeomorphism $\psi : M_1 \to M_2$ such that $\varphi_2 \circ \psi = \varphi_1$.

Remark 1.6.2

1. Let (M, φ) be a C^k-submanifold of N and let $\tilde{\varphi} : M \to \varphi(M)$ denote the induced mapping. Since $\tilde{\varphi}$ is bijective, one can use it to carry over the topological and differentiable structure from M to $\varphi(M)$, thus making $\tilde{\varphi}$ into a diffeomorphism. Therefore, $(\varphi(M), i)$ with the natural inclusion mapping $i : \varphi(M) \to N$ is a submanifold of N equivalent to (M, φ). This shows that, up to equivalence, every submanifold of N may be assumed to be given by a subset and the corresponding natural inclusion mapping. We stress that the topology of this subset (coming with its manifold structure) need not coincide with the relative topology induced from N. Consequences of this fact will be discussed below.
2. The concept of equivalence carries over in an obvious way from submanifolds to immersions. Two immersions $\varphi_1 : M_1 \to M$ and $\varphi_2 : M_2 \to M$ are said to be equivalent if there exists a diffeomorphism $\psi : M_1 \to M_2$ such that $\varphi_2 \circ \psi = \varphi_1$.

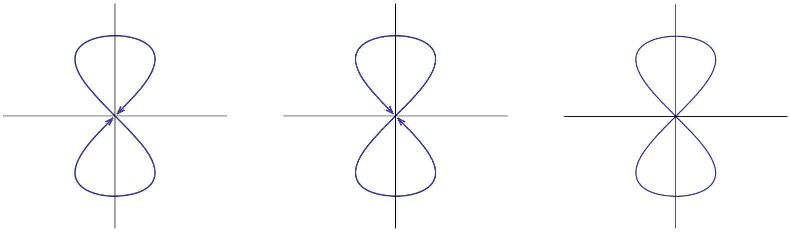


Fig. 1.6 The figure eight submanifolds γ_+ (*left*) and γ_- (*middle*) and the figure eight immersion (*right*) of Example 1.6.6/2. The arrows mean that the curves approach themselves without touching

Let us start with a couple of examples.

Example 1.6.3 (Open subsets) Let N be a C^k-manifold, let M be an open subset of N with the induced smooth structure and let $\varphi : M \to N$ be the natural inclusion mapping. For all $m \in M$, $\mathrm{T}_m M = \mathrm{T}_m N$ and φ'_m is the identical mapping. Hence, φ is an immersion and (M, φ) is a C^k-submanifold.

Example 1.6.4 (Level sets) Let V and W be finite-dimensional real vector spaces and let M be the level set of a regular value of a function $f : V \to W$ of class C^k. Let $\iota : M \to V$ denote the natural inclusion mapping. Then, (M, ι) is a C^k-submanifold of V. Indeed, according to the Level Set Theorem 1.2.1, M is a C^k-manifold. To see that ι is of class C^k, choose a basis in V to identify V with $\mathbb{R}^n$ and recall the construction of local charts on M in the proof of this theorem. The local representative of ι with respect to such a chart is given by the mapping which in this proof is denoted by λ. By the Implicit Function Theorem, this mapping is of class C^k. Finally, for every $v \in M$, under the natural identifications of $\mathrm{T}_v M$ with $\ker f'(v)$ and of $\mathrm{T}_v V$ with V, see Example 1.4.3/1, ι'_v is given by the natural inclusion mapping of the subspace $\ker f'(v)$ of V and is thus injective.

Example 1.6.5 (Graphs) Let $M = \mathbb{R}$, $N = \mathbb{R}^2$ and let $f : \mathbb{R} \to \mathbb{R}$ be a function of class C^k. Define $\varphi(x) := (x, f(x))$, $x \in \mathbb{R}$. The image of φ is the graph of the function f. By construction, φ is of class C^k and injective. The tangent mapping is $\varphi'_x = (1, f'(x))$; it is injective for all $x \in \mathbb{R}$. Hence, $(\mathbb{R}, \varphi)$ is a submanifold of class C^k of $\mathbb{R}^2$.

Example 1.6.6 (Curves) More generally, let $M = I \subset \mathbb{R}$ be an open interval and let $\gamma : I \to N$ be a curve of class C^k. If $\dot\gamma(t) \neq 0$ for all $t \in I$, γ is an immersion. If γ is also injective, that is, if the curve γ does not intersect itself, then (I, γ) is a submanifold of class C^k of N. This holds, in particular, for curves in $N = \mathbb{R}^2$ of the form $\gamma(t) = (t, f(t))$, where $f : \mathbb{R} \to \mathbb{R}$ is a C^k-function (Example 1.6.5). We list three typical examples of curves which are submanifolds.

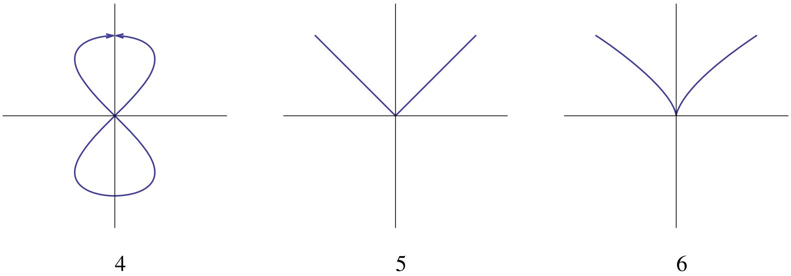


Fig. 1.7 The curves *4–6* of Example 1.6.6. The arrows in *4* mean that the curve approaches the vertical axis without touching

1. Let $I = \mathbb{R}$, $N = \mathbb{R}^2$ and $\gamma(t) = (t, 0)$. The image of γ is the x-axis.
2. Let $I = (0, 1)$, $N = \mathbb{R}^2$ and $\gamma_\pm(t) = (\pm\sin(4\pi t), \sin(2\pi t))$. Both curves have the same image as the corresponding curves with $I = \mathbb{R}$, which are just immersions and not submanifolds. The image is a closed subset of $\mathbb{R}^2$ known as the figure eight, see Fig. 1.6. Correspondingly, submanifolds or immersions with that image will be referred to as a figure eight submanifold or immersion, respectively.
3. Let $I = \mathbb{R}$, $N = \mathrm{T}^2 = \mathrm{S}^1 \times \mathrm{S}^1$, realized as the subset $\{(z, w) \in \mathbb{C}^2 : |z|^2 = |w|^2 = 1\}$. For $(z, w) \in \mathrm{T}^2$ and $\vartheta \in \mathbb{R}$, let

$$\gamma_{(z,w),\vartheta}(t) = \big(z\mathrm{e}^{2\pi\mathrm{i}t}, w\mathrm{e}^{2\pi\mathrm{i}\vartheta t}\big).$$

   This is an immersion. It is injective and hence defines a submanifold iff ϑ is irrational. In this case, it is known that the image $\gamma(\mathbb{R})$ is dense in T^2. Let us remark that the curves $\gamma_{(z,w),\vartheta}$ with ϑ irrational are usually referred to as the orbits of the irrational torus flow. The notion of flow will be introduced in Chap. 3.

For illustration, we also list three curves which are not submanifolds, see Fig. 1.7. Let $N = \mathbb{R}^2$ and $I = (-1, 1)$.

4. The curve $\gamma(t) = (\sin(2\pi(t + 1)), \cos(\pi(t + 1)))$ is an immersion but intersects itself, hence it is not injective.[9]
5. The curve $\gamma(t) = (t, |t|)$ is not differentiable at 0.
6. The curve $\gamma(t) = (t^3, t^2)$ is not an immersion at 0.

Next, we prove that every submanifold admits an atlas whose charts are induced from charts of the ambient manifold.

Proposition 1.6.7 (Charts adapted to a submanifold) *Let (M, φ) be a C^k-submanifold of N. For every $m \in M$ there exists an open neighbourhood U of m in M and a local chart (V, ρ) on N such that*

[9] The image of this curve is not a figure eight, because it is missing the point $(0, 1)$.

1. $\varphi(U) \subset V$,
2. $\rho(\varphi(U))$ *is an open subset of* $(\mathbb{R}^{\dim M} \times \{0\}) \subset \mathbb{R}^{\dim N}$,
3. *the chart* $(U, \rho \circ \varphi_{\restriction U})$ *induced on* M *is compatible with the* C^k*-structure of* M.

Proof Let $r = \dim M$ and $n = \dim N$. The Constant Rank Theorem 1.5.11 yields local charts (U, κ) on M at m and (V, ρ) on N at $\varphi(m)$ such that

$$\rho \circ \varphi \circ \kappa^{-1}(x_1, \dots, x_n) = (x_1, \dots, x_r, 0, \dots, 0), \quad (x_1, \dots, x_n) \in \kappa(U).$$

Since $\rho(\varphi(U))$ is an open subset of the subspace $\mathbb{R}^r$ and $\rho \circ \varphi_{\restriction U}$ is a homeomorphism onto its image, $(U, \rho \circ \varphi_{\restriction U})$ is a local chart on M. Let (W, σ) be a local chart of the C^k-structure on M. Since the transition mapping between $(U, \rho \circ \varphi_{\restriction U})$ and (W, σ) is given by the local representative of φ with respect to the charts (W, σ) and (V, ρ), the charts $(U, \rho \circ \varphi_{\restriction U})$ and (W, σ) are compatible. □

Motivated by the observations made in Remark 1.6.2/1, we continue with introducing two special classes of submanifolds. Let (M, φ) be a submanifold of N. Since φ is continuous, the topology induced on $\varphi(M)$ by means of the induced bijection $\tilde{\varphi} : M \to \varphi(M)$ is at least as fine as the relative topology induced from N. It may be finer, though, see Examples 1.6.12/3 and 1.6.12/5 below. The first class of submanifolds to be introduced is characterized by the property that these two topologies coincide. A necessary and sufficient condition for this is that φ be open onto its image.[10]

Definition 1.6.8 (Embedded submanifold) A C^k-submanifold (M, φ) of N is called embedded if φ is open onto its image.

The second class of submanifolds is characterized by a property related to mappings. Let (M, φ) be a submanifold of N and let P be another manifold. For a C^k-mapping $\chi : N \to P$, we define $\chi_{\restriction M} : M \to P$ by

$$\chi_{\restriction M} := \chi \circ \varphi,$$

and for a C^k-mapping $\psi : P \to N$ with $\psi(P) \subset \varphi(M)$, we define $\psi^{\restriction M} : P \to M$ by

$$\varphi \circ \psi^{\restriction M} := \psi.$$

By a slight abuse of terminology, we refer to $\chi_{\restriction M}$ as the restriction of χ in domain to M and to $\psi^{\restriction M}$ as the restriction of ψ in range to M. Obviously, $\chi_{\restriction M}$ is a C^k-mapping again. For $\psi^{\restriction M}$ this need not hold, because if the topology on $\varphi(M)$ induced from M is finer than the relative topology induced from N, $\psi^{\restriction M}$ need not be continuous; see Example 1.6.12/5.

[10] By definition, this means that the induced mapping $\tilde{\varphi}$ is open with respect to the relative topology. In this case, φ is called an embedding.

Definition 1.6.9 (Initial submanifold) A C^k-submanifold (M, φ) of N is called initial[11] if for any C^k-manifold P and any C^k-mapping $\psi : P \to N$ which satisfies $\psi(P) \subset \varphi(M)$, the restriction in range $\psi^{\restriction M} : P \to M$ is continuous.

Since for a mapping between topological spaces, the property of being continuous is preserved under arbitrary restrictions to subsets with the relative topology, every embedded submanifold is initial.

Proposition 1.6.10 (Restriction in range) *Let N, P be C^k-manifolds, let (M, φ) be a C^k-submanifold of N and let $\psi : P \to N$ be a C^k-mapping with $\psi(P) \subset \varphi(M)$. The restriction in range $\psi^{\restriction M} : P \to M$ is of class C^k iff it is continuous.*

Proof Denote $\dim M = k$ and $\dim N = l$. Continuity of $\psi^{\restriction M}$ is of course necessary for differentiability. To see that it is also sufficient, it is enough to show that it implies that $\psi^{\restriction M}$ is of class C^k in a neighbourhood of an arbitrary point $p \in P$. According to Proposition 1.6.7, there exists an open neighbourhood U of $\psi^{\restriction M}(p)$ in M and a local chart (V, ρ) on N at $\varphi(\psi^{\restriction M}(p)) = \psi(p)$ such that $\rho \circ \varphi_{\restriction U}$ takes values in the subspace $\mathbb{R}^k \times \{0\}$ of $\mathbb{R}^l$ and thus induces a chart $(U, \rho \circ \varphi_{\restriction U})$ on M at $\psi^{\restriction M}(p)$. Since $\psi^{\restriction M}$ is continuous, $(\psi^{\restriction M})^{-1}(U)$ is open in P. Hence, there exists a local chart (W, σ) on P at p such that $\psi^{\restriction M}(W) \subset U$. Due to $\rho \circ \varphi \circ \psi^{\restriction M} \circ \sigma^{-1} = \rho \circ \psi \circ \sigma^{-1}$, the local representative of $\psi^{\restriction M}$ with respect to the charts (W, σ) and $(U, \rho \circ \varphi_{\restriction U})$ coincides, up to the embedding $\mathbb{R}^k \to \mathbb{R}^l$ which is suppressed here, with the local representative of ψ with respect to the charts (W, σ) and (V, ρ). This shows that $\psi^{\restriction M}$ is of class C^k. □

Corollary 1.6.11 *The restriction of a C^k-mapping in range to an initial C^k-submanifold is of class C^k.*

Example 1.6.12

1. Open subsets (Example 1.6.3), level sets of C^k-functions at regular values (Example 1.6.4) and graphs of C^k-functions on $\mathbb{R}$ (Example 1.6.5) are embedded submanifolds. In the first case, φ itself is open. In the second case, M is a subset of $\mathbb{R}^n$ and is equipped with the relative topology, so that the natural inclusion mapping is open onto its image by construction. In the third case, the image of an open interval (a, b) under φ can be written as
$$\varphi\big((a, b)\big) = \big((a, b) \times \mathbb{R}\big) \cap \varphi(M)$$
and is hence open in $\varphi(M)$.
2. The statement about the graph of a C^k-function on $\mathbb{R}$ generalizes to arbitrary C^k-mappings. Let M and N be C^k-manifolds and let $\psi \in C^k(M, N)$. Define $\varphi : M \to M \times N$ by $\varphi(m) := (m, \psi(m))$. Then, (M, φ) is an embedded C^k-submanifold of $M \times N$. The proof is left to the reader (Exercise 1.6.1).

[11] Or, alternatively, weakly embedded.

3. Let $M = \mathbb{R}$, $N = \mathrm{T}^2$ and let $\varphi = \gamma_{(z,w),\vartheta}$ be one of the curves of Example 1.6.6/3, with ϑ irrational. The submanifold (M, φ) is initial but not embedded. That it is not embedded is due to the fact that $\varphi(M)$ is dense in N, because this implies that the preimage under φ of the intersection of $\varphi(M)$ with an arbitrary open subset of N cannot be bounded. To see that these submanifolds are initial, let $\psi : P \to N$ be a C^k-mapping with $\psi(P) \subset \varphi(M)$ and let $p \in P$, $t \in M$ such that $\psi(p) = \varphi(t)$. There exist an open interval I containing t and an open subset $U \subset N$ such that $\varphi(I)$ is an arcwise connected component of the subset $\varphi(M) \cap U$ of N with respect to the relative topology; and φ restricts to a homeomorphism from I onto $\varphi(I)$. Let $W_p \subset P$ denote the arcwise connected component of $\psi^{-1}(U)$ containing p. As a mapping to $\varphi(M)$ with respect to the relative topology, ψ is continuous and thus preserves arcwise connectedness. Hence, $\psi(W_p) \subset \varphi(I)$. Now let $\{p_n\}$ be a sequence in P converging to p. Since W_p is an open neighbourhood of p in P, we may assume $p_n \in W$ for all n. Then, $\psi(p_n) \in \varphi(I)$ for all n and
$$\varphi\big(\psi^{\restriction M}(p_n)\big) = \psi(p_n) \to \psi(p) = \varphi(t),$$
hence $\psi^{\restriction M}(p_n) \to t$. Since p was arbitrary, this shows that $\psi^{\restriction M}$ is continuous.
4. More generally, every orbit of a Lie group action is an initial submanifold. It is embedded if the action is proper, see Chap. 6.
5. Let (M, φ) be the figure eight submanifold of $N = \mathbb{R}^2$ given by the curve γ_+ of Example 1.6.6/2. (M, φ) is not initial, because for $P = \mathbb{R}$ and $\psi = \gamma_-$, the restriction in range $\psi^{\restriction M}$ maps the convergent sequence $\{\frac{1}{n}\}$ in P to a divergent sequence in M.

Remark 1.6.13

1. Let (M, φ) be a C^k-submanifold of N. Proposition 1.6.7 shows that for any $m \in M$ there exists an open neighbourhood U of m in M such that $(U, \varphi_{\restriction U})$ is an embedded C^k-submanifold of N. That is, every submanifold (M, φ) is locally embedded, where locally refers to the topology of M. More generally, one can show that for every compact subset A of M there exists an open neighbourhood U of A in M such that $(U, \varphi_{\restriction U})$ is an embedded C^k-submanifold of N; see Exercise 1.6.2 for instructions on a proof.
2. A compact C^k-submanifold is always embedded. To see this, consider the induced mapping $\tilde{\varphi} : M \to \varphi(M)$, where $\varphi(M)$ carries the relative topology induced from N. Since $\tilde{\varphi}$ is a bijection, it is open iff it is closed. We show the latter. Let $A \subset M$ be closed. Since closed subsets of compact spaces are compact, A is compact. Since compactness is preserved under continuous mappings, $\varphi(A)$ is compact. Since compact subsets of Hausdorff spaces are closed, $\varphi(A)$ is closed in $\varphi(M)$.
3. The following criterion is simple yet useful. Let M be a C^k-manifold and let $\varphi \in C^k(M, N)$. (M, φ) is an embedded C^k-submanifold of N iff for every $m \in M$ there exists an open neighbourhood V of $\varphi(m)$ in N such that

$(\varphi^{-1}(V), \varphi_{\restriction \varphi^{-1}(V)})$ is an embedded C^k-submanifold of V or of N. The statement remains true if embedded is replaced by initial. The proof is left to the reader (Exercise 1.6.3).

4. The property of being initial or embedded is stable with respect to taking preimages of submanifolds under so-called transversal mappings, see Sect. 1.8.
5. Initial C^k-submanifolds (M_1, φ_1) and (M_2, φ_2) of N are equivalent iff $\varphi_1(M_1) = \varphi_2(M_2)$ as subsets of N. In particular, if a subset of N admits a differentiable structure which makes it into an initial submanifold, this structure is unique. To see this, it suffices to show that $\varphi_1(M_1) = \varphi_2(M_2)$ implies equivalence. Let $\psi : M_1 \to M_2$ be the restriction in range of φ_1 to the submanifold (M_2, φ_2). Then, $\varphi_1 = \varphi_2 \circ \psi$. The restriction in range of φ_2 to the submanifold (M_1, φ_1) yields ψ^{-1}. Since both (M_1, φ_1) and (M_2, φ_2) are initial, ψ is a diffeomorphism.
6. In contrast to the case of initial submanifolds, in the general case, one and the same subset may be the image of non-equivalent submanifolds. For example, let $N = \mathbb{R}^2$ and let M be given by the figure eight, see Example 1.6.6/2. By viewing M as the image of $\mathbb{R}$ under the curves γ_+ or γ_-, we obtain two smooth structures on M, denoted by $M_\pm$. Let $j : M \to N$ denote the natural inclusion mapping. Both (M_+, j) and (M_-, j) are smooth submanifolds of $\mathbb{R}^2$. While the manifolds M_+ and M_- are diffeomorphic, because they are both diffeomorphic to $\mathbb{R}$, the submanifolds (M_+, j) and (M_-, j) of $\mathbb{R}^2$ are not equivalent, because the only mapping $\psi : M_+ \to M_-$ satisfying $j = j \circ \psi$ is the identical mapping $M_+ \to M_-$ which however is not continuous.

Proposition 1.6.14 *Let M, N and P be C^k-manifolds.*

1. *Let (M, φ) be a C^k-submanifold of N and let (N, ψ) be a C^k-submanifold of P. Then, $(M, \psi_{\restriction M})$ is a C^k-submanifold of P. If (M, φ) and (N, ψ) are initial (embedded), so is $(M, \psi_{\restriction M})$.*
2. *Let (M, φ) and (P, ψ) be C^k-submanifolds of N such that $\psi(P) \subset \varphi(M)$ and assume that (M, φ) is initial. Then, $(P, \psi^{\restriction M})$ is a C^k-submanifold of M. If (P, ψ) is initial (embedded), so is $(P, \psi^{\restriction M})$.*

Proof 1. A composition of injective mappings is injective. A composition of immersions is an immersion, see Remark 1.5.14. By definition of being initial, $(M, \psi_{\restriction M})$ is initial if so are (M, φ) and (N, ψ). Due to the fact that a composition of mappings which are open onto their images is open onto its image (Exercise 1.6.4), $(M, \psi_{\restriction M})$ is embedded if so are (M, φ) and (N, ψ).

2. The restriction in range $\psi^{\restriction M}$ is differentiable by definition of initial submanifold. It is injective, because so is ψ. It is an immersion, because so is ψ; see Remark 1.5.14. Hence, $(P, \psi^{\restriction M})$ is a C^k-submanifold of M, indeed. Next, assume that (P, ψ) is initial. Let Y be a C^k-manifold and let $\chi \in C^k(Y, M)$ such that $\chi(Y) \subset \psi^{\restriction M}(P)$. For the restriction $\chi^{\restriction P}$ of χ in range to P there holds $\varphi \circ \chi = \psi \circ \chi^{\restriction P}$. Since $\varphi \circ \chi$ is of class C^k and since (P, ψ) is initial, $\chi^{\restriction P}$ is of class C^k. Hence, $(P, \psi^{\restriction M})$ is initial. Finally, assume that (P, ψ) is embedded. Then, ψ is open onto its image and we have to show that the same holds for $\psi^{\restriction M}$.

Let $A \subset P$ be open. Then, $\psi(A)$ is open in $\psi(P)$, that is, $\psi(A) = \psi(P) \cap B$ for some open $B \subset N$. Thus,

$$\varphi^{-1}\big(\psi(A)\big) = \varphi^{-1}\big(\psi(P) \cap B\big) = \varphi^{-1}\big(\psi(P)\big) \cap \varphi^{-1}(B).$$

By definition of $\psi^{\restriction M}$, $\varphi^{-1}(\psi(A)) = \psi^{\restriction M}(A)$ and $\varphi^{-1}(\psi(P)) = \psi^{\restriction M}(P)$. Since $\varphi^{-1}(B)$ is open in M by continuity of φ, the above equality shows that $\psi^{\restriction M}(A)$ is open in $\psi^{\restriction M}(P)$. Hence, $\psi^{\restriction M}$ is open onto its image, as asserted. □

Exercises

1.6.1 Show that the graph of a C^k-mapping $\varphi : M \to N$ is an embedded submanifold of $M \times N$, cf. Example 1.6.12/2.

1.6.2 Prove the statement made in Remark 1.6.13/1 that for every compact subset A of a C^k-submanifold (M, φ) of N there exists an open neighbourhood U of A in M such that $(U, \varphi_{\restriction U})$ is an embedded C^k-submanifold of N.
Hint. Use local charts to show that A possesses an open neighbourhood U of A in M with compact closure. By means of the argument which was used in Remark 1.6.13/2 to prove that a compact submanifold is necessarily embedded, show that the restriction of φ to the closure of U is open onto its image.

1.6.3 Prove the criterion for the embedding property of Remark 1.6.13/3.
Hint. Show that a continuous mapping $f : X \to Y$ of topological spaces X, Y is open onto its image if and only if for every $y \in f(X)$ there exists a neighbourhood V of y in Y such that the induced mapping $f^{-1}(V) \to V$ is open onto its image.

1.6.4 Complete the proof of Proposition 1.6.14 by showing that the composition of mappings which are open onto their images is open onto its image.

1.7 Subsets Admitting a Submanifold Structure

In this section we discuss the question how to characterize those subsets M of N which are images of submanifolds, cf. Remark 1.6.2/1. This will be needed for the Transversal Mapping Theorem in the next section and for the discussion of distributions and foliations in Sect. 3.5. The question under consideration may be rephrased as follows. Let $M \subset N$ and let $\iota : M \to N$ denote the natural inclusion mapping. Under which conditions does there exist a C^k-structure on M such that (M, ι) is a submanifold (initial submanifold, embedded submanifold) of N? Here, by a C^k-structure on M we mean both a topology on M and a maximal atlas whose charts are local homeomorphisms with respect to this topology. Of course, if (M, ι) is embedded, the underlying topology coincides with the relative topology induced from N. Below we will derive existence criteria in terms of local charts on N, starting with the general case and then turning to the initial and embedded cases.

Proposition 1.7.1 *Let N be a C^k-manifold and let $M \subset N$ be a subset. Consider the following condition.*

(S) *There exists $l \in \mathbb{N}$ and a countable covering $\{M_i\}$ of M such that, for every i, one can find a local chart (V_i, ρ_i) on N satisfying*
   (S1) *$M_i \subset V_i$ and $\rho_i(M_i)$ is an open subset of the subspace $\mathbb{R}^l \times \{0\}$ of $\mathbb{R}^{\dim N}$,*
   (S2) *for all $i \neq j$, $\rho_i(M_i \cap M_j)$ is an open subset of $\mathbb{R}^l \times \{0\}$.*

If condition (S) *holds for some $l \in \mathbb{N}$, the family of bijections $\{(M_i, \rho_{i\restriction M_i})\}$ induces a C^k-structure of dimension l on M. With respect to this structure, (M, ι) is a C^k-submanifold of N. Conversely, if there exists a C^k-structure of dimension l on M such that (M, ι) is a C^k-submanifold of N, then* (S) *holds for this l and the family $\{M_i\}$ can be chosen so that the C^k-structure so induced on M coincides with the original one.*

Condition (S2) is necessary, because otherwise one could, for example, cover the figure eight in $\mathbb{R}^2$ by subsets M_1 and M_2 forming a figure S and a reversed S, respectively. Then, there exist local charts on $\mathbb{R}^2$ such that condition (S1) holds. However, since $M_1 \cap M_2$ contains an isolated point (the origin), it cannot be mapped to an open subset of $\mathbb{R}$ by any of these charts. Hence, the figure eight cannot become a topological manifold this way.

Proof Assume that (S) holds. According to Remark 1.1.10, the family of bijections $\{(M_i, \rho_{i\restriction M_i})\}$ defines a topology on M which is Hausdorff and second countable and with respect to which the $(M_i, \rho_{i\restriction M_i})$ are local charts of dimension l on M. Here, by abuse of notation, we view $\rho_{i\restriction M_i}$ as a mapping to $\mathbb{R}^l \times \{0\}$. The transition mappings between two such local charts on M are obtained by restriction of the transition mappings between the original local charts on N to subsets of the subspace $\mathbb{R}^l \times \{0\}$ of $\mathbb{R}^{\dim N}$ which are open by condition (S2). Hence, the transition mappings are of class C^k and thus the charts $(M_i, \rho_{i\restriction M_i})$ define a C^k-structure of dimension l on M. Since for every i, the local representative of ι with respect to the local charts $(M_i, \rho_{i\restriction M_i})$ and (V_i, ρ_i) is given by restriction of the natural embedding $\mathbb{R}^l \times \{0\} \to \mathbb{R}^{\dim N}$ to an open subset of $\mathbb{R}^l \times \{0\}$, ι is a C^k-immersion. Thus, (M, ι) is a C^k-submanifold.

The converse assertion is due to Proposition 1.6.7 and the fact that M contains a countable dense set (because it is second countable). □

For initial submanifolds, following [211] we introduce the following notion. Let N be a C^k-manifold. A piecewise C^k-curve in N is a continuous mapping $[a, b] \to N$ for which there exist $a < t_1 < \cdots < t_r < b$ such that the restriction to (t_i, t_{i+1}) is of class C^k for all $i = 0, \ldots, r$, where $t_0 = a$ and $t_{r+1} = b$. A subset $A \subset N$ is said to be C^k-arcwise connected relative to N if any two points of A can be joined by a piecewise C^k-curve in N which is contained in A. For points of A, the property of being joinable by such a curve defines an equivalence relation in A. The equivalence classes are called the C^k-arcwise connected components of A relative to N. Below we will need the following evident facts.

(a) Any subset of A that is C^k-arcwise connected relative to N is contained in a C^k-arcwise connected component of A relative to N.
(b) If A is open in N, its C^k-arcwise connected components relative to N are open in N.[12]
(c) If $A \subset V$ for some local chart (V, ρ) on N, then A is C^k-arcwise connected relative to N iff $\rho(A)$ is C^k-arcwise connected relative to $\mathbb{R}^{\dim N}$.

Proposition 1.7.2 *Let N be a C^k-manifold and let $M \subset N$ be a subset. Consider the following condition.*

(I) *There exists $l \in \mathbb{N}$ and a countable covering $\{M_i\}$ of M such that, for every i, one can find a local chart (V_i, ρ_i) on N such that*
 (I1) *$M_i \subset V_i$ and $\rho_i(M_i)$ is an open subset of the subspace $\mathbb{R}^l \times \{0\} \subset \mathbb{R}^{\dim N}$,*
 (I2) *M_i is a C^k-arcwise connected component of $V_i \cap M$ relative to N.*

If condition (I) *holds for some $l \in \mathbb{N}$, the family of bijections $\{(M_i, \rho_{i \restriction M_i})\}$ induces a C^k-structure of dimension l on M. With respect to this structure, (M, ι) is an initial C^k-submanifold of N. Conversely, if there exists a C^k-structure of dimension l on M such that (M, ι) is an initial C^k-submanifold of N, then* (I) *holds for this l and the C^k-structure so induced on M coincides with the original one.*

Proof First, assume that (I) holds. The proof that (M, ι) is a submanifold of N is analogous to that of Proposition 1.7.1, except for the fact that, here, we have to give an argument that the domains $\rho_i(M_i \cap M_j)$ of the transition mappings between the charts induced on M are open subsets of $\mathbb{R}^l \times \{0\}$. The case where $M_i \cap M_j$ is empty is trivial. Thus, assume $M_i \cap M_j \neq \varnothing$. Let $m_0 \in M_i \cap M_j$. Let A denote the C^k-arcwise connected component of $M_i \cap V_j$ relative to N which contains m_0. It is easy to see that $A \subset M_i \cap M_j$. Then, $\rho_i(m_0) \in \rho_i(A) \subset \rho_i(M_i \cap M_j)$. Moreover, $\rho_i(A)$ is a C^k-arcwise connected component of the open subset $\rho_i(M_i \cap V_j)$ of $\mathbb{R}^l \times \{0\}$ relative to $\mathbb{R}^{\dim N}$, hence relative to $\mathbb{R}^l \times \{0\}$. Hence, $\rho_i(A)$ itself is open in $\mathbb{R}^l \times \{0\}$. This shows that $\rho_i(M_i \cap M_j)$ is open in $\mathbb{R}^l \times \{0\}$, as asserted.

It remains to show that (M, ι) is initial. Let P be a C^k-manifold and let $\psi \in C^k(P, N)$ be such that $\psi(P) \subset M$. For every $p \in P$, there exists an i such that $\psi(p) \in M_i$. Since $\psi^{-1}(V_i)$ is open, there exists a local chart (W, σ) on P at p such that $\psi(W) \subset V_i$. Then, $\psi(W) \subset V_i \cap M$. W can be chosen to be C^k-arcwise connected relative to P (e.g. by choosing it so that $\sigma(W)$ is convex). Then, $\psi(W)$ is C^k-arcwise connected relative to N, so that $\psi(W) \subset M_i$. It follows that the local representative $(\psi^{\restriction M})_{\sigma, \rho_{i \restriction M_i}}$ is given by the restriction in range of ψ_{σ, ρ_i} to the subspace $\mathbb{R}^l \times \{0\}$, hence it is of class C^k. Since the domain $\sigma(W)$ of $(\psi^{\restriction M})_{\sigma, \rho_{i \restriction M_i}}$ contains $\sigma(p)$ and since p was arbitrary, it follows that $\psi^{\restriction M}$ is of class C^k.

Now, assume that there exists a C^k-structure on M such that (M, ι) is an initial submanifold of N. Since M is second countable, there exists a countable dense

[12] In fact, they coincide with the connected components of A.

family of points $\{m_i\}$ in M. For each i, Proposition 1.6.7 yields an open neighbourhood U_i of m_i in M and a local chart $(\tilde{V}_i, \tilde{\rho}_i)$ on N at m such that $\tilde{\rho}_i(U_i)$ is an open subset of the subspace $\mathbb{R}^l \times \{0\} \subset \mathbb{R}^{\dim N}$. For $\varepsilon > 0$, let $B_{i,\varepsilon}$ denote the open ε-ball in $\mathbb{R}^{\dim N}$ about $\tilde{\rho}_i(m_i)$. Choose ε_i so that $\overline{B_{i,\varepsilon_i}} \subset \tilde{\rho}_i(\tilde{V}_i)$ and $(\overline{B_{i,\varepsilon_i}} \cap (\mathbb{R}^l \times \{0\})) \subset \tilde{\rho}_i(U_i)$. Define

$$V_i := \tilde{\rho}^{-1}(B_{i,\varepsilon_i}), \qquad \rho_i := \tilde{\rho}_{i \restriction V_i}, \qquad M_i := \tilde{\rho}_i^{-1}\big(B_{i,\varepsilon_i} \cap \big(\mathbb{R}^l \times \{0\}\big)\big).$$

Since the subsets M_i are open in M and since the family of points $\{m_i\}$ is dense, the family $\{M_i\}$ covers M. By construction, it satisfies condition (I1). It remains to check condition (I2). Since M_i is C^k-arcwise connected relative to N by construction, it remains to show that M_i contains all $m \in V_i \cap M$ for which there exists a piecewise C^k-curve $\gamma : [0,1] \to N$ such that $\gamma([0,1]) \subset V_i \cap M$, $\gamma(0) = m_i$ and $\gamma(1) = m$. Let such m and γ be given. Let $0 < t_1 < \cdots < t_r < 1$ be such that $\gamma_{\restriction (t_i,t_{i+1})}$ is of class C^k for all $i = 0, \ldots, r$, where $t_0 = 0$ and $t_{r+1} = 1$. Since (M, ι) is initial, $(\gamma_{\restriction (t_i,t_{i+1})})^{\restriction M}$ is continuous for all i, hence $\gamma^{\restriction M}$ is continuous. Thus, if $m \notin M_i$, since M_i is open in M, there must exist $t \in [0,1]$ such that $\gamma(t) \in \overline{M_i} \setminus M_i$, where $\overline{M_i}$ denotes the closure of M_i in M. Since $\tilde{\rho}_i$ maps U_i homeomorphically onto $\tilde{\rho}_i(U_i)$, there holds $\tilde{\rho}_i(\overline{M_i}) = \overline{B_{i,\varepsilon_i}} \cap (\mathbb{R}^l \times \{0\})$. This implies $\tilde{\rho}_i(\gamma(t)) \in (\mathbb{R}^l \times \{0\}) \setminus B_{i,\varepsilon_i}$ and hence $\gamma(t) \notin V_i$ (contradiction). Thus, $m \in M_i$ and condition (I2) holds, indeed. Now, since the family of subsets M_i so constructed satisfies (I1) and (I2), it induces a C^k-structure on M with respect to which (M, ι) is an initial submanifold of M. According to Remark 1.6.13/5, this C^k-structure coincides with the original one. □

In the case of an embedded submanifold, M already carries a topology, namely, the relative topology induced from N.

Proposition 1.7.3 *Let N be a C^k-manifold and let $M \subset N$ be a subset. Assume that M is equipped with the relative topology induced from N. Consider the following condition.*

(E) *There exists $l \in \mathbb{N}$ and a family of local charts $\{(V_i, \rho_i)\}$ on N such that*
  (E1) $M \subset \bigcup_i V_i$,
  (E2) *for every i, $\rho_i(V_i \cap M) = \rho_i(V_i) \cap (\mathbb{R}^l \times \{0\})$.*

If condition (E) *holds for some $l \in \mathbb{N}$, the local charts (V_i, ρ_i) induce local charts $(V_i \cap M, \rho_{i \restriction V_i \cap M})$ of dimension l on M. These charts establish a C^k-atlas and hence a C^k-structure on M. With respect to this structure, (M, ι) is an embedded C^k-submanifold of N. Conversely, if there exists a C^k-structure on M such that (M, ι) is an embedded C^k-submanifold of N of dimension l, then* (E) *holds for this l and the C^k-structure so induced on M coincides with the original one.*

Proof Assume that (E) holds. As a topological subspace of N, M is Hausdorff and second countable.[13] For every i, the subset $V_i \cap M$ of M is open in the relative topology. Being the restriction of a homeomorphism onto its image, $\rho_{i\restriction V_i \cap M}$ is a homeomorphism onto its image. Since by assumption the image is $\rho_i(V_i) \cap (\mathbb{R}^l \times \{0\})$ and is hence open in $\mathbb{R}^l \times \{0\}$, the pair $(V_i \cap M, \rho_{i\restriction V_i \cap M})$ is a local chart on M. By the same arguments as in the proof of Proposition 1.7.1, the transition mappings are of class C^k and (M, ι) is a C^k-submanifold of N. Since, by definition of the relative topology, ι is open onto its image. Thus, (M, ι) is an embedded C^k-submanifold, as asserted.

For the converse assertion, let $\{U_i\}$ be the family of open subsets of M and let $\{(V_i, \rho_i)\}$ be the family of local charts on N provided by Proposition 1.6.7. Since M carries the relative topology, for every i, there exists an open subset W_i in N such that $U_i = M \cap W_i$. By replacing V_i by $V_i \cap W_i$ we obtain the desired family of local charts of N. Since embedded submanifolds are initial, Remark 1.6.13/5 yields that the C^k-structure on M induced by this family coincides with the original one. □

Remark 1.7.4 Conditions (S), (I) and (E) of Propositions 1.7.1, 1.7.2 and 1.7.3 can be reformulated in terms of submanifolds rather than charts in various ways. The following are particularly convenient. Since the proofs are straightforward, they are left to the reader (Exercise 1.7.3).

Condition (S) of Proposition 1.7.1 can be replaced by

(S) *There exists $l \in \mathbb{N}$ and a countable covering $\{M_i\}$ of M such that, for every i, the subset M_i carries a C^k-structure satisfying*
- (S1) *$(M_i, \iota_{\restriction M_i})$ is an embedded C^k-submanifold of N of dimension l,*
- (S2) *for all i, j, the intersection $M_i \cap M_j$ is open in M_i and M_j.*

Condition (I) of Proposition 1.7.2 can be replaced by

(I) *There exists $l \in \mathbb{N}$ and a countable family of open subsets $\{V_i\}$ of N such that*
- (I1) $M \subset \bigcup_i V_i$,
- (I2) *for every i, $V_i \cap M$ admits a C^k-structure of dimension l with respect to which it is an initial C^k-submanifold of N.*

Condition (E) of Proposition 1.7.3 can be replaced by

(E) *There exists $l \in \mathbb{N}$ and a family of open subsets $\{V_i\}$ of N such that*
- (E1) $M \subset \bigcup_i V_i$,
- (E2) *for every i, $V_i \cap M$ admits a C^k-structure of dimension l with respect to which it is an embedded C^k-submanifold of N.*

For later use, from Proposition 1.7.3 we extract

Corollary 1.7.5 *Let N be a C^k-manifold and let $M \subset N$ be an embedded C^k-submanifold. For every $m \in M$ there exists a local chart (V, ρ) on N such that*

[13] Since M already carries a topology, there is no need to require the family $\{(V_i, \rho_i)\}$ to be countable.

$M \cap V$ is the set of solutions of the equations $\rho^i = 0$, $i = \dim M + 1, \ldots, \dim N$, and $(V \cap M, (\rho^1, \ldots, \rho^{\dim M})_{\restriction V \cap M})$ is a local chart on M at m. □

As an application of Proposition 1.7.3, one can prove (Exercise 1.7.4)

Proposition 1.7.6 (Level Set Theorem for mappings of locally constant rank) *Let N and P be C^k-manifolds, let $\varphi \in C^k(N, P)$ and let $p \in P$ such that $M := \varphi^{-1}(p)$ is nonempty. Assume that every $m_0 \in M$ has a neighbourhood in N where* $\operatorname{rank} \varphi'_m$ *is constant.*[14] *Then, every connected component*[15] *M_0 of M is an embedded C^k-submanifold of N of dimension* $\dim N - \operatorname{rank} \varphi'_{m_0}$, *where $m_0 \in M_0$.*

Finally, we note that the terminology concerning submanifolds used here is consistent with e.g. [166] and [302]. It is common as well to take condition (E) of Proposition 1.7.3 as the definition of submanifold (which hence corresponds to our embedded submanifold) and to refer to our submanifolds as immersed or virtual submanifolds, as is done for example in [73], [130], [180] and [232].

Exercises

1.7.1 Use Proposition 1.7.2 to show that the submanifold of T^2 given by the curve $\gamma_{(z,w),\vartheta}$ of Example 1.6.6/3, with ϑ irrational, is initial.
1.7.2 Use Proposition 1.7.2 to show that the figure eight submanifolds of Example 1.6.6/2 are not initial.
1.7.3 Prove that conditions (S), (I) and (E) of Propositions 1.7.1, 1.7.2 and 1.7.3 can be reformulated as given in Remark 1.7.4.
1.7.4 Use Proposition 1.7.3 to prove Proposition 1.7.6.

1.8 Transversality

The Level Set Theorem 1.2.1 states that the preimage of a regular value of a differentiable mapping from an open subset of $\mathbb{R}^n$ to $\mathbb{R}^m$ is an embedded submanifold of $\mathbb{R}^n$. This has a generalization to differentiable mappings between manifolds and to preimages of submanifolds rather than just points. In this context, the condition of regularity is replaced by the condition of transversality.

Definition 1.8.1 (Transversality) Let N, P and Q be C^k-manifolds.

1. C^k-mappings $\varphi : N \to P$ and $\psi : Q \to P$ are transversal if for all $y \in N$ and $q \in Q$ such that $\varphi(y) = \psi(q) \equiv p$ there holds $\varphi'_y(\mathrm{T}_y N) + \psi'_q(\mathrm{T}_q Q) = \mathrm{T}_p P$ (the sum need not be direct).

---

[14] In other words, φ is a subimmersion at every $m_0 \in M$, cf. Exercise 1.5.5/(a).

[15] Since the connected components of M may have different dimensions, M as a whole may not be a manifold. We leave it to the reader to provide examples.

2. A C^k-mapping $\varphi : N \to P$ is transversal to a C^k-submanifold (Q, ψ) of P if φ and ψ are transversal.
3. C^k-submanifolds (N, φ) and (Q, ψ) of P are transversal if φ and ψ are transversal.

If the submanifolds (N, φ) and (Q, ψ) are given by subsets, the condition of transversality is usually written in the form

$$\mathrm{T}_p N + \mathrm{T}_p Q = \mathrm{T}_p P, \quad p \in N \cap Q.$$

Theorem 1.8.2 (Transversal Mapping Theorem) *Let N and P be C^k-manifolds, let (Q, φ) be a C^k-submanifold of P and let $\psi \in C^k(N, P)$ be transversal to (Q, φ). Assume that $M := \psi^{-1}(\varphi(Q))$ is nonempty and let $\iota : M \to N$ denote the inclusion mapping.*

1. *There exists a C^k-structure on M such that (M, ι) is a C^k-submanifold of N of dimension* $\dim N - \dim P + \dim Q$.[16] *For every $m \in M$,*
$$\iota'_m(\mathrm{T}_m M) = \big\{X_m \in \mathrm{T}_m N : \psi'_m(X_m) \in \varphi'_q(\mathrm{T}_q Q)\big\},$$
where $q \in Q$ such that $\varphi(q) = \psi(m)$.
2. *If (Q, φ) is initial, so is (M, ι).*
3. *If (Q, φ) is embedded, so is (M, ι).*

Proof First, we use transversality to reduce the problem locally to the Level Set Theorem 1.2.1. Then, in order to prove the assertions 1–3, we apply Propositions 1.7.1–1.7.3. Denote $n = \dim N$, $r = \dim P$ and $s = \dim Q$. Write $\mathbb{R}^r = \mathbb{R}^s \times \mathbb{R}^{r-s}$ and let $\mathrm{pr}_2 : \mathbb{R}^r \to \mathbb{R}^{r-s}$ denote the projection onto the second factor. Choose a countable subset $\{m_i : i \in I\}$ of M which is dense with respect to the relative topology induced from N. According to Proposition 1.6.7, for every i, there exists an open subset Q_i of Q and a local chart (W_i, σ_i) on P such that $\psi(m_i) \in \varphi(Q_i) \subset W_i$, the image $\sigma_i(\varphi(Q_i))$ is an open subset of the subspace $\mathbb{R}^s \times \{0\} \subset \mathbb{R}^r$ and the local chart $(Q_i, \sigma_i \circ \varphi_{\restriction Q_i})$ is compatible with the C^k-structure on Q. Given (W_i, σ_i), choose a local chart (V_i, ρ_i) on N at m_i such that $\psi(V_i) \subset W_i$ and define

$$M_i := \psi^{-1}\big(\varphi(Q_i)\big) \cap V_i, \qquad \psi_i := \mathrm{pr}_2 \circ \psi_{\sigma_i, \rho_i} : \rho_i(V_i) \to \mathbb{R}^{r-s}.$$

Then, $m_i \in M_i$ and $\rho_i(M_i) = \psi_i^{-1}(0)$. We check that ψ_i is a submersion at $\mathbf{x}$ for all $\mathbf{x} \in \rho_i(M_i)$: denote $m := \rho_i^{-1}(\mathbf{x})$ and let $q \in Q$ be such that $\varphi(q) = \psi(m)$. By transversality,

$$\mathbb{R}^s \times \{0\} + (\sigma_i)'_{\psi(m)} \circ \psi'_m(\mathrm{T}_m N) = (\sigma_i)'_{\psi(m)}\big(\varphi'_q(\mathrm{T}_q Q) + \psi'_m(\mathrm{T}_m N)\big) = \mathbb{R}^r$$

[16] That is, the codimension of M in N equals the codimension of Q in P.

and hence

$$(\psi_i)'_{\mathbf{x}}\big(\mathbb{R}^n\big) = \mathrm{pr}_2\big((\sigma_i)'_{\psi(m)} \circ \psi'_m(\mathrm{T}_m N)\big) = \mathbb{R}^{r-s},$$

as asserted. Thus, the Level Set Theorem 1.2.1 implies that $\rho_i(M_i)$ is an embedded C^k-submanifold of $\rho_i(V_i)$ of dimension $n-r+s$. Then, M_i is an embedded submanifold of N of the same dimension and we can modify (V_i, ρ_i) (and hence M_i) in such a way that

$$\rho_i(M_i) = \rho_i(V_i) \cap \big(\mathbb{R}^{n-r+s} \times \{0\}\big). \tag{1.8.1}$$

1. In order that we can apply Proposition 1.7.1, it remains to check that for all $i_1 \neq i_2$, $\rho_{i_1}(M_{i_1} \cap M_{i_2})$ is open in $\mathbb{R}^{n-r+s} \times \{0\}$. For that purpose, it suffices to show that $M_{i_1} \cap M_{i_2}$ is open in M_{i_1} with respect to the relative topology induced from N. Since φ is injective,

$$M_{i_1} \cap M_{i_2} = \psi^{-1}\big(\varphi(Q_{i_1} \cap Q_{i_2})\big) \cap V_{i_1} \cap V_{i_2}.$$

Since $Q_{i_1} \cap Q_{i_2}$ is open in Q_{i_1} and since $\varphi_{\restriction Q_{i_1}}$ is an embedding, $\varphi(Q_{i_1} \cap Q_{i_2})$ is open in $\varphi(Q_{i_1})$ with respect to the relative topology induced from P. Thus, $\varphi(Q_{i_1} \cap Q_{i_2}) = \varphi(Q_{i_1}) \cap W$ for some open $W \subset P$ and we obtain

$$M_{i_1} \cap M_{i_2} = \psi^{-1}\big(\varphi(Q_{i_1})\big) \cap \psi^{-1}(W) \cap V_{i_1} \cap V_{i_2} = M_{i_1} \cap \psi^{-1}(W) \cap V_{i_2},$$

which is an open subset of M_{i_1}, indeed. Now, Proposition 1.7.1 yields the desired C^k-structure on M. To determine the tangent spaces, let $m \in M_i$ and $\mathbf{x} = \rho_i(m)$. According to the Level Set Theorem 1.2.1, $\mathrm{T}_{\mathbf{x}}(\rho_i(M_i)) = \ker(\psi_i)'_{\mathbf{x}}$. Hence,

$$\mathrm{T}_m M_i = \ker(\mathrm{pr}_2 \circ \sigma_i \circ \psi_{\restriction V_i})'_m.$$

Since $\ker(\mathrm{pr}_2 \circ \sigma_i)'_{\psi(m)} = \varphi'_q \mathrm{T}_q Q$, the assertion follows.

2. If (Q, φ) is initial, according to Proposition 1.7.2, the subsets Q_i and W_i can be chosen so that $\varphi(Q_i)$ is a C^k-arcwise connected component of $W_i \cap \varphi(Q)$ relative to P. We show that, then, M_i is a C^k-arcwise connected component of $V_i \cap M$ relative to N. In view of Proposition 1.7.2, this yields the assertion. Thus, let $m \in V_i \cap M$ be such that there exists a piecewise C^k-curve γ in N which is contained in $V_i \cap M$ and joins m to a point of M_i. We have to show that $m \in M_i$. Now, $\psi \circ \gamma$ is a piecewise C^k-curve in P which is contained in $\psi(V_i \cap M) \subset W_i \cap \varphi(Q)$ and joins $\psi(m)$ to a point in $\psi(M_i) \subset \varphi(Q_i)$. Since $\varphi(Q_i)$ is a C^k-arcwise connected component of $W_i \cap \varphi(Q)$ relative to P, it follows that $\psi \circ \gamma$ lies in $\varphi(Q_i)$ and hence γ lies in $V_i \cap \psi^{-1}(\varphi(Q_i)) = M_i$. Thus, $m \in M_i$, as asserted.

3. If φ is an embedding, the subsets Q_i and W_i can be chosen so that $\varphi(Q_i) = \varphi(Q) \cap W_i$. This implies $M_i = M \cap V_i$. Then, (1.8.1) and Proposition 1.7.3 yield the assertion. □

Since single points are embedded submanifolds of dimension zero and a mapping $N \to P$ is transversal to a point of P if and only if this point is a regular value, Theorem 1.8.2 implies

Corollary 1.8.3 (Level Set Theorem for manifolds) *Let N and P be C^k-manifolds, let $\psi \in C^k(N, P)$ and let $c \in P$ be a regular value of ψ such that $M := \psi^{-1}(c)$ is nonempty. Let $\iota : M \to N$ denote the inclusion mapping. Then (M, ι) is an embedded C^k-submanifold of N of dimension* $\dim N - \dim P$ *and for every $m \in M$ there holds $\iota'_m(\mathrm{T}_m M) = \ker \psi'_m$.* □

Corollary 1.8.3 can also be proved directly by means of the Level Set Theorem 1.2.1 for $\mathbb{R}^n$ and Proposition 1.7.3 in the formulation of Remark 1.7.4 (Exercise 1.8.1).

Remark 1.8.4 Given a C^k-manifold N of dimension n and an embedded C^k-submanifold (M, φ) of N of dimension l, there exists a C^k-mapping $f : N \to \mathbb{R}^{n-l}$ such that $\varphi(M) = f^{-1}(0)$. For the proof, without loss of generality, one may assume that $M \subset N$ and that φ is the natural inclusion mapping. According to Proposition 1.7.3, there exists a family of local charts $\{(V_i, \rho_i) : i \in I\}$ on N such that the V_i cover M and $\rho_i(M \cap V_i) = \rho_i(V_i) \cap \mathbb{R}^r \times \{0\}$. Complement this family to an atlas on N by adding local charts whose domains do not intersect M and whose image does not intersect the subspace $\{0\} \times \mathbb{R}^{n-l}$. According to Proposition 1.3.7, there exists a partition of unity $\{g_j : j \in J\}$ of class C^k subordinate to the open covering $\{V_i : i \in I\}$ of N. Then, for each j, there is an i such that $\mathrm{supp}(g_j) \subset V_i$. By choosing one such i for each j and dropping all the other local charts we obtain a new atlas whose local charts we denote by (V_j, ρ_j). Define

$$f_j : V_j \to \mathbb{R}^{n-l}, \qquad f_j := \left(\left(\rho_j^{l+1}\right)^2, \dots, \left(\rho_j^n\right)^2\right).$$

By extending $f_j g_j$ by 0 to N one obtains a family of C^k-mappings $\tilde{f}_j : N \to \mathbb{R}^{n-l}$. Then, $f := \sum_{j \in J} \tilde{f}_j$ has the desired properties.

Corollary 1.8.5 (Intersection of transversal submanifolds) *Let N be a C^k-manifold and let (M_1, φ_1) and (M_2, φ_2) be transversal C^k-submanifolds of N. Let $M := \varphi_1(M_1) \cap \varphi_2(M_2)$ be nonempty and let $\iota : M \to N$ denote the natural inclusion mapping.*

1. *There exists a C^k-structure on M such that (M, ι) is a C^k-submanifold of N of dimension* $\dim M_1 + \dim M_2 - \dim N$. *For every $m \in M$,*

$$\mathrm{T}_m M = \left((\varphi_1)'_{m_1}(\mathrm{T}_{m_1} M_1)\right) \cap \left((\varphi_2)'_{m_2}(\mathrm{T}_{m_2} M_2)\right),$$

   where $m_i \in M_i$ such that $m = \varphi_i(m_i)$, $i = 1, 2$.
2. *If both (M_1, φ_1) and (M_2, φ_2) are initial, so is (M, ι).*
3. *If both (M_1, φ_1) and (M_2, φ_2) are embedded, so is (M, ι).*

If (M_1, φ_1) and (M_2, φ_2) are given by subsets, one may write

$$M = M_1 \cap M_2 \quad \text{and} \quad \mathrm{T}_m M = \mathrm{T}_m M_1 \cap \mathrm{T}_m M_2, m \in M,$$

where $T_m M_i$, $i = 1, 2$, is viewed as a subspace of $T_m N$, defined by the C^k-structure on M_i.

Proof 1. Application of Theorem 1.8.2/1 to the mapping $\varphi_1 : M_1 \to N$ and the submanifold (M_2, φ_2) of N yields a C^k-structure on $\tilde{M} := \varphi_1^{-1}(\varphi_2(M_2))$ such that $\tilde{M}$, together with the natural inclusion mapping $\tilde{\iota} : \tilde{M} \to M_1$, is a C^k-submanifold of M_1 of dimension $\dim M_1 - \dim N + \dim M_2$. According to Proposition 1.6.14/1, then $(\tilde{M}, \varphi_1 \circ \tilde{\iota})$ is a C^k-submanifold of N. Since $\varphi_1 \circ \tilde{\iota}(\tilde{M}) = M$, $\varphi_1 \circ \tilde{\iota}$ induces a bijection $\varphi : \tilde{M} \to M$. We use this bijection to transport the C^k-structure of $\tilde{M}$ to M. Then, $\iota = \varphi_1 \circ \tilde{\iota} \circ \varphi^{-1}$ and φ^{-1} is a diffeomorphism, hence (M, ι) is a C^k-submanifold of N of the same dimension as $\tilde{M}$. To determine the tangent spaces, let $m \in M$ and $m_i \in M_i$ such that $\varphi_i(m_i) = m$. Theorem 1.8.2/1 yields

$$\tilde{\iota}'_{m_1}(T_{m_1}\tilde{M}) = \left\{X_{m_1} \in T_{m_1}\tilde{M} : (\varphi_1)'_{m_1}(X_{m_1}) \in (\varphi_2)'_{m_2}(T_{m_2}M_2)\right\}.$$

Apply $(\varphi_1)'_m$ to both sides of this equality and use $\varphi'_{m_1}(T_{m_1}\tilde{M}) = T_m M$ to obtain the assertion.

2. and 3. According to Theorem 1.8.2/2, since (M_2, φ_2) is initial, so is $(\tilde{M}, \tilde{\iota})$. Since (M_1, φ_1) is initial, Proposition 1.6.14/1 yields that $(\tilde{M}, \varphi_1 \circ \tilde{\iota})$ is initial. Then, so is (M, ι). A similar argument proves assertion 3. □

Exercises

1.8.1 Prove the Level Set Theorem for manifolds (Theorem 1.8.3) directly by means of the Level Set Theorem for $\mathbb{R}^n$ (Theorem 1.2.1) and Proposition 1.7.3 in the formulation of Remark 1.7.4.

Chapter 2
Vector Bundles

Vector bundles constitute a special class of manifolds, which is of great importance in physics. In particular, all sorts of tensor fields occurring in physical models may be viewed in a coordinate-free manner as sections of certain vector bundles. We start with the observation that the tangent spaces of a manifold combine in a natural way into a bundle, which is called tangent bundle. Next, by taking its typical properties as axioms, we arrive at the general notion of vector bundle. In Sect. 2.2, we discuss elementary aspects of this notion, including the proof that—up to isomorphy—vector bundles are completely determined by families of transition functions. In Sect. 2.3 we discuss sections and frames,[1] and in Sect. 2.4 we present the tool kit for vector bundle operations. We will see that, given some vector bundles over the same base manifold, by applying fibrewise the standard algebraic operations of taking the dual vector space, of building the direct sum and of taking the tensor product, we obtain a universal construction recipe for building new vector bundles. In Sect. 2.5, by applying these operations to the tangent bundle of a manifold, we get the whole variety of tensor bundles over this manifold. The remaining two sections contain further operations, which will be frequently used in this book. In Sect. 2.6, we discuss the notion of induced bundle and Sect. 2.7 is devoted to subbundles and quotient bundles. There is a variety of special cases occurring in applications: regular distributions, kernel and image bundles, annihilators, normal and conormal bundles.

2.1 The Tangent Bundle

Let M be a C^k-manifold, let $I \subset \mathbb{R}$ be an open interval and let $\gamma : I \to M$ be a C^k-curve. According to Example 1.5.6, for every $t \in I$, the tangent vector $\dot{\gamma}(t)$ of γ at t is an element of the tangent space $\mathrm{T}_{\gamma(t)}M$. Hence, while t runs through I, $\dot{\gamma}(t)$ runs through the tangent spaces along γ, see Fig. 2.1.

[1]Here, as well as in Sect. 2.5, in order to keep in touch with the physics literature, the local description is presented in some detail. In particular, we discuss transformation properties. This way, we make contact with classical tensor analysis.

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*,
Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7_2, © Springer Science+Business Media Dordrecht 2013

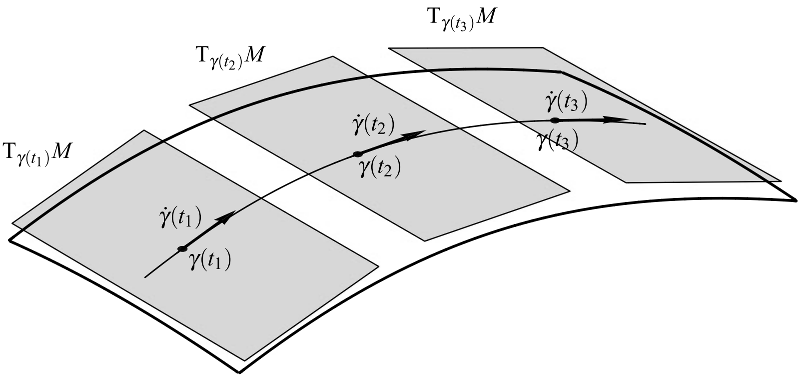


Fig. 2.1 Tangent vectors along a curve γ in M

To follow the tangent vectors along γ it is convenient to consider the totality of all tangent spaces of M. This leads to the notion of tangent bundle of a manifold M, denoted by $\mathrm{T}M$. As a set, $\mathrm{T}M$ is given by the disjoint union of the tangent spaces at all points of M, that is,

$$\mathrm{T}M := \bigsqcup_{m\in M} \mathrm{T}_m M. \tag{2.1.1}$$

Let $\pi : \mathrm{T}M \to M$ be the canonical projection which assigns to an element of $\mathrm{T}_m M$ the point m for every $m \in M$. $\mathrm{T}M$ can be equipped with a manifold structure as follows. Denote $n = \dim M$. Choose a countable atlas $\{(U_\alpha, \kappa_\alpha) : \alpha \in A\}$ on M and define the mappings

$$\kappa_\alpha^{\mathrm{T}} : \pi^{-1}(U_\alpha) \to \mathbb{R}^n \times \mathbb{R}^n, \qquad \kappa_\alpha^{\mathrm{T}}(X_m) := \big(\kappa_\alpha(m), \mathbf{X}_m^{\kappa_\alpha}\big). \tag{2.1.2}$$

The image of $\kappa_\alpha^{\mathrm{T}}$ is given by $\kappa_\alpha(U_\alpha) \times \mathbb{R}^n$ and is hence open in $\mathbb{R}^n \times \mathbb{R}^n$. Using (1.4.9), for the transition mappings we obtain

$$\kappa_\beta^{\mathrm{T}} \circ \big(\kappa_\alpha^{\mathrm{T}}\big)^{-1}(\mathbf{x}, \mathbf{X}) = \big(\kappa_\beta \circ \kappa_\alpha^{-1}(\mathbf{x}), \big(\kappa_\beta \circ \kappa_\alpha^{-1}\big)'(\mathbf{x}) \cdot \mathbf{X}\big), \tag{2.1.3}$$

where $(\mathbf{x}, \mathbf{X}) \in \kappa_\alpha(U_\alpha \cap U_\beta) \times \mathbb{R}^n$. Since $\kappa_\alpha \circ \kappa_\beta^{-1}$ is of class C^k, the transition mappings are of class C^{k-1}. Finally, it is obvious that the subsets $\pi^{-1}(U_\alpha)$ cover $\mathrm{T}M$. Thus, according to Remark 1.1.10, the family of bijections $\{(\pi^{-1}(U_\alpha), \kappa_\alpha^{\mathrm{T}}) : \alpha \in A\}$ defines a differentiable structure of class C^{k-1} and dimension $2n$ on $\mathrm{T}M$, which has the following properties. First, due to (2.1.3), it is independent of the choice of an atlas on M used to construct it. Second, the local representative of the projection $\pi : \mathrm{T}M \to M$ with respect to the charts $\kappa_\alpha^{\mathrm{T}}$ and κ_α is given by the natural projection pr_1 to the first factor in $\kappa_\alpha(U_\alpha) \times \mathbb{R}^n$. Hence, π is a submersion of class C^{k-1}. Third, the charts $\kappa_\alpha^{\mathrm{T}}$ identify the open submanifolds $\pi^{-1}(U_\alpha)$ of $\mathrm{T}M$ with direct products of an open subset of M with a copy of $\mathbb{R}^n$. Under this identification, both the natural projection and the vector space structure on every tangent space

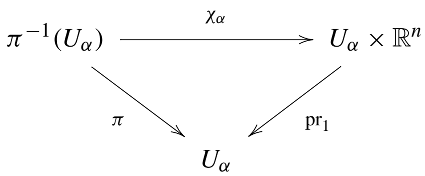
$\mathrm{T}_m M$, $m \in U_\alpha$, is preserved. To formalize this, for every $\alpha \in A$, define a mapping

$$\chi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^n, \qquad \chi_\alpha(X_m) := \big(m, \mathbf{X}_m^{\kappa_\alpha}\big). \tag{2.1.4}$$

Then $\kappa_\alpha^{\mathrm{T}} = (\kappa_\alpha \times \mathrm{id}_{\mathbb{R}^n}) \circ \chi_\alpha$. In particular, the local representative of χ_α with respect to the global charts $\kappa_\alpha^{\mathrm{T}}$ on $\pi^{-1}(U_\alpha)$ and $\kappa_\alpha \times \mathrm{id}_{\mathbb{R}^n}$ on $U_\alpha \times \mathbb{R}^n$ is given by the identical mapping of $\mathbb{R}^n \times \mathbb{R}^n$, restricted to the open subset $\kappa_\alpha(U_\alpha) \times \mathbb{R}^n$. Hence, χ_α is a C^{k-1}-diffeomorphism. Moreover, $\mathrm{pr}_1 \circ \chi_\alpha = \pi_{\restriction \pi^{-1}(U_\alpha)}$ and the restrictions $\chi_{\alpha \restriction \mathrm{T}_m M}$ are vector space isomorphisms for all $m \in U_\alpha$. Let us summarize.

Proposition 2.1.1 *Let M be a C^k-manifold of dimension n and let $\mathrm{T}M$ be defined by* (2.1.1). *There exists a unique C^{k-1}-structure on $\mathrm{T}M$ such that for every local chart (U, κ) on M, the mapping $\kappa^{\mathrm{T}} : \pi^{-1}(U) \to \mathbb{R}^n \times \mathbb{R}^n$, defined by* (2.1.2), *is a local chart on $\mathrm{T}M$. With respect to this structure, $\mathrm{T}M$ has dimension $2n$ and the following holds.*

1. *The natural projection $\pi : \mathrm{T}M \to M$ is a surjective submersion.*
2. *There exists an open covering $\{U_\alpha\}$ of M and an associated family of diffeomorphisms $\chi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^n$ such that*
   (a) *the following diagram commutes,*



   (b) *for every $m \in U_\alpha$, the induced mapping $\mathrm{pr}_2 \circ \chi_{\alpha \restriction \mathrm{T}_m M} : \mathrm{T}_m M \to \mathbb{R}^n$ is a vector space isomorphism.*

Definition 2.1.2 The triple $(\mathrm{T}M, M, \pi)$ is called the tangent bundle of M. $\mathrm{T}M$ is called the total space or the bundle manifold, M the base manifold and π the natural projection. For $m \in M$, $\pi^{-1}(m) \equiv \mathrm{T}_m M$ is called the fibre over m. The vector space $\mathbb{R}^n$ is called the typical fibre and the pairs (U_α, χ_α) are called local trivializations of $\mathrm{T}M$ over U_α.

By an abuse of notation, the tangent bundle will usually be denoted by $\mathrm{T}M$.

Example 2.1.3 Let $M = \mathrm{S}^1$ be realized as the unit circle in $\mathbb{R}^2$. For every $\mathbf{x} \in \mathrm{S}^1$, the tangent space $\mathrm{T}_{\mathbf{x}}\mathrm{S}^1$ can be identified with the subspace of vectors orthogonal to $\mathbf{x}$. This yields a bijection Φ from TS^1 onto the subset

$$T = \big\{(\mathbf{x}, \mathbf{X}) \in \mathrm{S}^1 \times \mathbb{R}^2 : \mathbf{x} \perp \mathbf{X}\big\}$$

of $\mathbb{R}^4$. This is the level set of the smooth mapping

$$F : \mathbb{R}^4 \to \mathbb{R}^2, \qquad F(\mathbf{x}, \mathbf{X}) := \big(\|\mathbf{x}\|^2, \mathbf{x} \cdot \mathbf{X}\big)$$

at the regular value $\mathbf{c} = (1, 0)$. Hence, it carries a smooth structure. One can check that Φ is a diffeomorphism with respect to this structure. (To see this, let $\mathrm{pr}_k : \mathbb{R}^2 \to \mathbb{R}$ denote the natural projection to the k-th component and choose the charts on S^1 and T to be restrictions of pr_k and $\mathrm{pr}_k \times \mathrm{pr}_k$, respectively, $k = 1, 2$.) Thus, TS^1 can naturally be identified with T. The construction carries over to higher-dimensional spheres: as a manifold, the tangent bundle TS^n can be identified with the subset $\{(\mathbf{x}, \mathbf{X}) \in \mathrm{S}^n \times \mathbb{R}^{n+1} : \mathbf{x} \perp \mathbf{X}\}$ of $\mathbb{R}^{2(n+1)}$ which is the level set of a function similar to F at the regular value $\mathbf{c} = (1, 0)$, see also Remark 2.1.4/2 below.

Remark 2.1.4

1. Let V be a finite-dimensional real vector space and let M be an open subset of V. The natural identifications of the tangent spaces $\mathrm{T}_v M$ with V for all $v \in M$, cf. Example 1.4.3/1, combine to a smooth diffeomorphism $\chi : \mathrm{T}M \to M \times V$ which is fibrewise linear. We will refer to χ as the natural identification of $\mathrm{T}M$ with $M \times V$. After choosing a basis in V, this bijection coincides with the (global) trivialization induced via (2.1.4) by the corresponding global chart on M.
2. The construction of Example 2.1.3 generalizes to arbitrary level sets. Let V, W be finite-dimensional real vector spaces and let M be the level set of a C^k-mapping $f : V \to W$ at a regular value $c \in W$. Identifying the tangent space $\mathrm{T}_v M$ with $\ker f'(v)$ for all $v \in M$, see Remark 1.2.2/1, we obtain a bijection Φ from $\mathrm{T}M$ onto the subset

$$T = \big\{(v, X) \in M \times V : f'(v)X = 0\big\}$$

of $V \times V$. This is the level set of the C^{k-1}-mapping

$$F : V \times V \to W \times W, \qquad F(v, X) := \big(f(v), f'(v)X\big)$$

at the value $(c, 0)$, whose regularity follows from that of c with respect to f. It follows that T is an embedded C^{k-1}-submanifold of $V \times V$ and that Φ is a C^{k-1}-diffeomorphism (Exercise 2.1.1). Thus, the tangent bundle of a level set in V can be naturally identified with a level set in $V \times V$.

Just as the tangent spaces of a manifold combine to the tangent bundle, the tangent mappings of a differentiable mapping combine to a mapping of the tangent bundles.

Definition 2.1.5 (Tangent mapping) Let M, N be C^k-manifolds and let $\Phi : M \to N$ be a C^k-mapping. The tangent mapping of Φ is defined by

$$\Phi' : \mathrm{T}M \to \mathrm{T}N, \qquad \Phi'(X_m) := \Phi'_m(X_m).$$

The tangent mapping is of class C^{k-1} (Exercise 2.1.6). The basic properties of the tangent mapping are stated in the next section (Proposition 2.2.9).

Exercises

2.1.1 Prove that the mapping Φ of Remark 2.1.4/2 is a diffeomorphism.
Hint. As local charts on M, use those constructed in the proof of the Level Set Theorem 1.2.1.

2.1.2 Determine the tangent bundle in the form of the level set T of Remark 2.1.4/2 for
(a) the spheres S^n, see Example 1.2.3,
(b) the hyperboloid of Example 1.2.4,
(c) the paraboloid, the ellipsoid and the rotational torus of Exercise 1.2.5,
(d) the classical groups, see Example 1.2.6.
Compare your result for the spheres S^n with Example 2.1.3.

2.1.3 Let M be the level set of a differentiable mapping $f : \mathbb{R}^n \to \mathbb{R}^m$ at a regular value $\mathbf{c} \in \mathbb{R}^m$. Identify $\mathrm{T}M$ with the level set T of Remark 2.1.4/2. The bundle of unit tangent vectors of M is defined to be $\mathrm{E}M := \{(\mathbf{x}, \mathbf{X}) \in \mathrm{T}M : \|\mathbf{X}\| = 1\}$. Show that $\mathrm{E}M$ is an embedded submanifold of $\mathrm{T}M$. What does one get for ES^1 and ES^2?

2.1.4 Let (U, κ) be a local chart on M and let κ^{T} be the local chart induced by κ on the tangent bundle $\mathrm{T}M$ via (2.1.2). Determine the local trivialization (2.1.4) of the tangent bundle $\mathrm{T}(\mathrm{T}M)$ of $\mathrm{T}M$ induced by κ^{T}.

2.1.5 Iterate the construction of Remark 2.1.4/2 by determining the level set T for the tangent bundle $\mathrm{T}(\mathrm{T}M)$ of the tangent bundle $\mathrm{T}M$ of a level set M. Write down the defining equations explicitly for $M = \mathrm{S}^n$.

2.1.6 Let $\Phi : M \to N$ be of class C^k. Show that Φ' is of class C^{k-1}.

2.2 Vector Bundles

The notion of vector bundle arises from the notion of tangent bundle of a manifold by allowing the fibres to be arbitrary finite-dimensional vector spaces, rather than the tangent spaces of that manifold.

Definition 2.2.1 (Vector bundle) Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and let $k \geq 0$. A $\mathbb{K}$-vector bundle of class C^k is a triple (E, M, π), where E and M are C^k-manifolds and $\pi : E \to M$ is a surjective C^k-mapping satisfying the following conditions.

1. For every $m \in M$, $E_m := \pi^{-1}(m)$ carries the structure of a vector space over $\mathbb{K}$.
2. There exists a finite-dimensional vector space F over $\mathbb{K}$, an open covering $\{U_\alpha\}$ of M and an associated family of C^k-diffeomorphisms $\chi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F$ such that, for all α,
   (a) the following diagram commutes,

$$\begin{array}{ccc} \pi^{-1}(U_\alpha) & \xrightarrow{\;\chi_\alpha\;} & U_\alpha \times F \\ & {\scriptstyle \pi}\searrow \quad \swarrow{\scriptstyle \mathrm{pr}_1} & \\ & U_\alpha & \end{array} \tag{2.2.1}$$

(b) for every $m \in U_\alpha$, the induced mapping $\chi_{\alpha,m} := \mathrm{pr}_2 \circ \chi_\alpha{\restriction}_{E_m} : E_m \to F$ is linear.

Like in the case of the tangent bundle, by an abuse of notation, a vector bundle (E, M, π) will usually be denoted by E alone. Like for the tangent bundle, E is called the total space or the bundle manifold, M the base manifold, π the bundle projection and F the typical fibre. For $m \in M$, E_m is called the fibre over m and m is called the base point. The pairs (U_α, χ_α) are called local trivializations. A local trivialization (U, χ) with $U = M$ is called a global trivialization. If a global trivialization exists, the vector bundle is called (globally) trivial.

Remark 2.2.2

1. By condition 2a, since the χ_α are diffeomorphisms, the bundle projection π is a submersion (because so is the projection to a factor of a direct product) and the fibres E_m are embedded submanifolds (because by χ_α they are mapped onto the subsets $\{m\} \times F$ of $U_\alpha \times F$). Being bijective and linear, the mappings $\chi_{\alpha,m}$ are vector space isomorphisms. Hence, all fibres have the same dimension as F; this number is called the dimension or the rank of the vector bundle. Thus, $\dim E = \dim M + \dim F$ for $\mathbb{K} = \mathbb{R}$, and $\dim E = \dim M + 2\dim F$ for $\mathbb{K} = \mathbb{C}$. For a $\mathbb{K}$-vector bundle of dimension n, one can always choose $F = \mathbb{K}^n$.
2. Let A denote the index set of a family of local trivializations $\{(U_\alpha, \chi_\alpha)\}$. The mappings
$$\chi_\beta \circ \chi_\alpha^{-1} : U_\alpha \cap U_\beta \times F \to U_\alpha \cap U_\beta \times F, \quad (\alpha, \beta) \in A \times A, \tag{2.2.2}$$
which are of class C^k, are called the transition mappings of the system of local trivializations $\{(U_\alpha, \chi_\alpha) : \alpha \in A\}$. Since for every $(\alpha, \beta) \in A \times A$, $\chi_\beta \circ \chi_\alpha^{-1}$ maps the subsets $\{m\} \times F$, where $m \in U_\alpha \cap U_\beta$, linearly and bijectively onto themselves, there exists a mapping $\rho_{\beta\alpha} : U_\alpha \cap U_\beta \to \mathrm{GL}(F)$ such that
$$\chi_\beta \circ \chi_\alpha^{-1}(m, u) = \big(m, \rho_{\beta\alpha}(m)u\big) \tag{2.2.3}$$
for all $m \in U_\alpha \cap U_\beta$ and $u \in F$. The mappings $\rho_{\beta\alpha}$ are called the transition functions of the system of local trivializations $\{(U_\alpha, \chi_\alpha) : \alpha \in A\}$. To see that they are of class C^k it suffices to check that for every $(\alpha, \beta) \in A \times A$ and $u \in F$, the mapping $U_\alpha \cap U_\beta \to F$ defined by $m \mapsto (m, \rho_{\beta\alpha}(m)u)$ is of class C^k. This follows at once from the differentiability of the transition mappings $\chi_\beta \circ \chi_\alpha^{-1}$. One can check that the transition functions satisfy
$$\rho_{\gamma\beta}(m)\rho_{\beta\alpha}(m) = \rho_{\gamma\alpha}(m) \tag{2.2.4}$$
for all $\alpha, \beta, \gamma \in A$ and $m \in U_\alpha \cap U_\beta \cap U_\gamma$.
3. A vector bundle is said to be orientable if there exists a family of local trivializations whose transition mappings have positive determinant.

Example 2.2.3

1. Let M be a C^k-manifold, let F be a vector space of dimension r over $\mathbb{K}$ and let $\mathrm{pr}_M : M \times F \to M$ denote the natural projection to the first component. Then, $(M \times F, M, \mathrm{pr}_M)$ is a $\mathbb{K}$-vector bundle of class C^k and dimension r. It is called the product vector bundle of M and F. A product vector bundle is obviously trivial.
2. According to Proposition 2.1.1, the tangent bundle of an n-dimensional C^k-manifold is an n-dimensional real vector bundle of class C^{k-1}.
3. Let (E, M, π) be a vector bundle of class C^k and let $U \subset M$ be open. Define $E_U := \pi^{-1}(U)$. This is an open subset of E and hence a C^k-manifold. By restriction, π induces a surjective C^k-mapping $\pi_U : E_U \to U$, and a system of local trivializations $\{(U_\alpha, \chi_\alpha)\}$ of E induces the system of local trivializations $\{(U_\alpha \cap U, \chi_{\alpha \restriction U_\alpha \cap U})\}$ of E_U. Thus, (E_U, U, π_U) is a $\mathbb{K}$-vector bundle of class C^k. It has the same dimension as E.

Example 2.2.4 (Möbius strip) Let $M = \mathrm{S}^1$ be realized as the unit circle in $\mathbb{C}$ and let E be the Möbius strip of Example 1.1.12 with the open interval $(-1, 1)$ replaced by the whole of $\mathbb{R}$. That is, $E := \mathbb{R}^2/\sim$, where $(s_1, t_1) \sim (s_2, t_2)$ iff there exists $k \in \mathbb{Z}$ such that $s_2 = s_1 + 2\pi k$ and $t_2 = (-1)^k t_1$. Define the projection by

$$\pi : E \to \mathrm{S}^1, \qquad \pi\big(\big[(s,t)\big]\big) := \mathrm{e}^{\mathrm{i}s}.$$

Using the local charts on E constructed in Example 1.1.12 one can easily check that π is smooth. The fibres are $E_{\mathrm{e}^{\mathrm{i}s}} = \pi^{-1}(\mathrm{e}^{\mathrm{i}s}) = \{[(s,t)] : t \in \mathbb{R}\}$. For every $s \in \mathbb{R}$, define

$$\lambda\big[(s,t_1)\big] + \big[(s,t_2)\big] := \big[(s, \lambda t_1 + t_2)\big], \quad \lambda, t_1, t_2 \in \mathbb{R}.$$

This way, the fibres become real vector spaces of dimension one. To construct local trivializations, we choose $U_\pm := \mathrm{S}^1 \setminus \{\pm 1\}$ and define mappings

$$\chi_\pm : \pi^{-1}(U_\pm) \to U_\pm \times \mathbb{R}, \qquad \chi_\pm\big(\big[(s,t)\big]\big) := \big(\mathrm{e}^{\mathrm{i}s}, t\big),$$

where in case of χ_+ and χ_- the representative (s,t) of $[(s,t)]$ used to compute the right hand side is chosen from $]0, 2\pi[\, \times \mathbb{R}$ and from $]-\pi, \pi[\, \times \mathbb{R}$, respectively. We leave it to the reader to check that the $\chi_\pm$ are diffeomorphisms and satisfy conditions 2a and 2b of Definition 2.2.1. Thus, (E, M, π) is a smooth real vector bundle of dimension 1. Figure 2.2 shows E together with the product vector bundle $\mathrm{S}^1 \times \mathbb{R}$. It is quite obvious that E is not trivial. We will be able to give a precise argument for that in the next section.

Remark 2.2.5 Let M be a C^k-manifold, let E be a set and let $\pi : E \to M$ be a surjective mapping such that conditions 1 and 2 of Definition 2.2.1 are satisfied, however, with the following difference. Instead of assuming the χ_α to be C^k-diffeomorphisms, assume that they are bijective and that their transition mappings

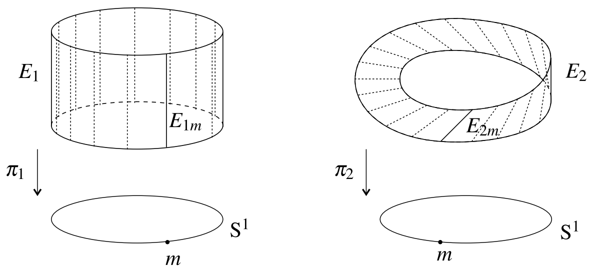


Fig. 2.2 The product vector bundle $S^1 \times \mathbb{R}$ and the Möbius strip as a vector bundle over S^1

(2.2.2) are of class C^k. Since M is second countable, the open covering $\{U_\alpha\}$ contains a countable subcovering. According to Remark 1.1.10, the corresponding subfamily of the family $\{\chi_\alpha\}$ defines a C^k-structure on E. With respect to this structure, (E, M, π) is a $\mathbb{K}$-vector bundle of class C^k and the (U_α, χ_α) are local trivializations. Conversely, if (E, M, π) is a vector bundle of class C^k, then the C^k-structure on E induced in this way coincides with the original one (Exercise 2.2.1).

Next, we consider mappings of vector bundles.

Definition 2.2.6 (Vector bundle morphism) Let (E_1, M_1, π_1) and (E_2, M_2, π_2) be $\mathbb{K}$-vector bundles of class C^k. A C^k-mapping $\Phi : E_1 \to E_2$ is called a morphism if for every $m_1 \in M_1$ there exists $m_2 \in M_2$ such that

1. $\Phi(E_{1,m_1}) \subset E_{2,m_2}$,
2. the induced mapping $\Phi_{m_1} := \Phi_{\restriction E_{1,m_1}} : E_{1,m_1} \to E_{2,m_2}$ is linear.

The rank of Φ is defined to be the integer-valued function which assigns to $m_1 \in M_1$ the rank of the linear mapping Φ_{m_1}. In case $M_1 = M_2 = M$, Φ is called a vertical morphism or a morphism over M if conditions 1 and 2 hold with $m_1 = m_2 = m$.

As usual, together with the notion of morphism there comes the notion of isomorphism (a bijective morphism whose inverse is also a morphism), endomorphism (a morphism of a vector bundle to itself), automorphism (an isomorphism of a vector bundle onto itself). For a vector bundle morphism Φ to be an isomorphism it is obviously sufficient for Φ to be a diffeomorphism. If Φ is vertical, it is sufficient for Φ to be bijective, because then the tangent mapping Φ' is bijective at any point and Theorem 1.5.7 yields that the inverse mapping is of class C^k.

Remark 2.2.7

1. Since Φ is a mapping, condition 1 implies that the point m_2 is uniquely determined by m_1. Thus, every morphism Φ induces a mapping $\varphi : M_1 \to M_2$, defined by

$$\varphi \circ \pi_1 = \pi_2 \circ \Phi.$$

One says that Φ covers φ and calls φ the projection of Φ. If Φ is of class C^k, so is φ. Indeed, if (U_1, χ_1) is a local trivialization of E_1, $\varphi_{\restriction U_1}$ coincides with the composition of the embedding $U_1 \to U_1 \times \{0\} \subset U_1 \times F_1$ with the mapping $\pi_2 \circ \Phi \circ \chi_1^{-1}$. In case $M_1 = M_2 = M$, Φ is a vertical morphism iff $\varphi = \mathrm{id}_M$.

2. Let (U_i, χ_i) be local trivializations of E_i, $i = 1, 2$. The mapping

$$\chi_2 \circ \Phi \circ \chi_1^{-1} : \big(U_1 \cap \varphi^{-1}(U_2)\big) \times F_1 \to U_2 \times F_2 \qquad (2.2.5)$$

is called the local representative of Φ with respect to (U_1, χ_1) and (U_2, χ_2). A fibre-preserving and fibrewise linear mapping $\Phi : E_1 \to E_2$ is a morphism iff all of its local representatives are of class C^k.

3. Let E_1, E_2 be $\mathbb{K}$-vector bundles over M of class C^k. For $\lambda \in \mathbb{K}$ and vertical morphisms $\Phi, \Psi : E_1 \to E_2$ we can define

$$(\lambda\Phi + \Psi)(x) := \lambda\Phi(x) + \Psi(x), \quad x \in E_1,$$

because for all $x \in E_1$, $\Phi(x)$ and $\Psi(x)$ belong to the same fibre of E_2. This provides a $\mathbb{K}$-vector space structure on the set of vertical morphisms from E_1 to E_2.

Example 2.2.8 A local trivialization (U, χ) of a $\mathbb{K}$-vector bundle (E, M, π) with typical fibre F is a vertical isomorphism from the vector bundle E_U, see Example 2.2.3/3, onto the product vector bundle $U \times F$. Accordingly, a global trivialization is a vertical isomorphism from E onto $M \times F$. Thus, a vector bundle is trivial iff it is isomorphic to a product vector bundle.

Probably the most important example of a vector bundle morphism is the tangent mapping. The reader may convince himself that Proposition 1.5.2 implies the following (Exercise 2.2.4).

Proposition 2.2.9 (Properties of the tangent mapping) *Let M and N be C^k-manifolds and let $\varphi \in C^k(M, N)$. The tangent mapping $\varphi' : \mathrm{T}M \to \mathrm{T}N$ has the following properties.*

1. *φ' is a vector bundle morphism of class C^{k-1} with projection φ.*
2. $(\mathrm{id}_M)' = \mathrm{id}_{\mathrm{T}M}$.
3. *If P is another C^k-manifold and $\psi \in C^k(N, P)$, then $(\psi \circ \varphi)' = \psi' \circ \varphi'$.*
4. *If φ is a diffeomorphism, then φ' is an isomorphism and $(\varphi')^{-1} = (\varphi^{-1})'$.*

Remark 2.2.10 (Partial derivatives and product rule) Let M_1, M_2, N be C^k-manifolds and let $\varphi \in C^k(M_1 \times M_2, N)$. We discuss the properties of the tangent mapping φ' which are related to the direct product structure of its domain. Proofs are left to the reader (Exercise 2.2.5). The induced partial mappings

$$\begin{aligned} \varphi_{m_2} &: M_1 \to N, & \varphi_{m_2}(m_1) &:= \varphi(m_1, m_2), & m_2 &\in M_2, \\ \varphi_{m_1} &: M_2 \to N, & \varphi_{m_1}(m_2) &:= \varphi(m_1, m_2), & m_1 &\in M_1, \end{aligned}$$

are of class C^k. Their tangent mappings combine to C^{k-1}-mappings

$$\begin{aligned} \mathrm{T}M_1 \times M_2 &\to \mathrm{T}N, \qquad (X_1, m_2) \mapsto (\varphi_{m_2})'(X_1), \\ M_1 \times \mathrm{T}M_2 &\to \mathrm{T}N, \qquad (m_1, X_2) \mapsto (\varphi_{m_1})'(X_2), \end{aligned}$$

called the partial derivatives of φ. They fulfil the product rule,

$$\varphi'_{(m_1,m_2)}(X_1, X_2) = (\varphi_{m_2})'(X_1) + (\varphi_{m_1})'(X_2), \quad m_i \in M_i, \ X_i \in \mathrm{T}_{m_i} M_i. \tag{2.2.6}$$

If $M_1 = M_2 = M$ and if φ is composed with the diagonal mapping $\Delta : M \to M \times M$, then

$$(\varphi \circ \Delta)'_m(X) = (\varphi_{1,m})'(X) + (\varphi_{2,m})'(X), \quad m \in M, \ X \in \mathrm{T}_m M. \tag{2.2.7}$$

In particular, if $M = I$ is some open interval, then $\varphi \circ \Delta$, φ_{t_1} and φ_{t_2} are C^k-curves in N. For the corresponding tangent vectors at $t \in I$ there holds

$$\frac{\mathrm{d}}{\mathrm{d}s}\restriction_t \varphi(s, s) = \frac{\mathrm{d}}{\mathrm{d}s}\restriction_t \varphi(s, t) + \frac{\mathrm{d}}{\mathrm{d}s}\restriction_t \varphi(t, s), \quad t \in I. \tag{2.2.8}$$

To conclude this section, we show that—up to isomorphy—vector bundles are completely determined by the family of transition functions associated with a system of local trivializations.

Theorem 2.2.11 (Reconstruction theorem) *Let M be a C^k-manifold. Assume that the following data are given*:

1. *a finite-dimensional vector space F over $\mathbb{K}$*,
2. *an open covering $\{U_\alpha : \alpha \in A\}$ of M*,
3. *a family of C^k-mappings $\rho_{\beta\alpha} : U_\alpha \cap U_\beta \to \mathrm{GL}(F)$, $(\alpha, \beta) \in A \times A$, satisfying* (2.2.4).

Then, *there exists a $\mathbb{K}$-vector bundle E over M of class C^k and a family of local trivializations $\{(U_\alpha, \chi_\alpha) : \alpha \in A\}$ of E whose transition functions are given by the functions $\rho_{\beta\alpha}$. E is uniquely determined up to vertical isomorphisms.*

In particular, the last assertion implies that if the $\rho_{\alpha\beta}$ are the transition functions of a vector bundle, then the vector bundle provided by Theorem 2.2.11 is isomorphic over M to the original one.

Proof First, we prove existence. Since M is second countable, the covering $\{U_\alpha : \alpha \in A\}$ contains a countable subcovering. Hence, for the following construction we may assume that A is countable. Moreover, we notice that (2.2.4) implies that $\rho_{\alpha\alpha} = \mathbb{1}$ for all $\alpha \in A$. Take the topological direct sum

$$\mathscr{X} := \bigsqcup_{\alpha \in A} U_\alpha \times F,$$

denote its elements by (α, m, u), where $\alpha \in A$, $m \in U_\alpha$ and $u \in F$, and define a relation on $\mathscr{X}$ by $(\alpha_1, m_1, u_1) \sim (\alpha_2, m_2, u_2)$ iff $m_1 = m_2$ and $u_2 = \rho_{\alpha_2\alpha_1}(m_1)u_1$. Due to $\rho_{\alpha\alpha} = \mathbb{1}$ and (2.2.4), this is an equivalence relation. Let E denote the set of equivalence classes. The mapping $\pi : E \to M$, given by $\pi[(\alpha, m, u)] := m$, is well-defined and surjective. To construct a vector space structure on $\pi^{-1}(m)$ for every $m \in M$, choose α such that $m \in U_\alpha$. Every class in $\pi^{-1}(m)$ has a unique representative of the form (α, m, u) with $u \in F$. Using this, we transport the linear structure from F to $\pi^{-1}(m)$,

$$\lambda\big[(\alpha, m, u_1)\big] + \big[(\alpha, m, u_2)\big] := \big[(\alpha, m, \lambda u_1 + u_2)\big], \quad u_1, u_2 \in F,\ \lambda \in \mathbb{K}.$$

By linearity of the mappings $\rho_{\beta\alpha}(m) : F \to F$, this definition does not depend on the choice of α. The natural injections $U_\alpha \times F \to \mathscr{X}$ induce mappings $U_\alpha \times F \to E$. Due to $\rho_{\alpha\alpha} = \mathbb{1}$, these mappings are injective and hence induce bijective mappings $\chi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F$. A brief computation shows that the transition mappings of the family of bijections $\{\chi_\alpha : \alpha \in A\}$ are given by (2.2.3). Therefore, they are of class C^k and hence define a C^k-structure on E with respect to which (E, M, π) is a $\mathbb{K}$-vector bundle of class C^k and the χ_α are local trivializations, see Remark 2.2.5. To prove uniqueness up to vertical isomorphisms, let $\tilde{E}$ be a $\mathbb{K}$-vector bundle over M of class C^k with projection $\tilde{\pi}$ and let $\tilde{\chi}_\alpha : \tilde{\pi}^{-1}(U_\alpha) \to U_\alpha \times F$ be local trivializations whose transition functions coincide with the $\rho_{\beta\alpha}$. Then, on $\pi^{-1}(U_\alpha \cap U_\beta) \subset E$ we have $\tilde{\chi}_\alpha^{-1} \circ \chi_\alpha = \tilde{\chi}_\beta^{-1} \circ \chi_\beta$ and on $\tilde{\pi}^{-1}(U_\alpha \cap U_\beta) \subset \tilde{E}$ there holds $\chi_\alpha^{-1} \circ \tilde{\chi}_\alpha = \chi_\beta^{-1} \circ \tilde{\chi}_\beta$. Hence, the mappings $\tilde{\chi}_\alpha^{-1} \circ \chi_\alpha$ and $\chi_\alpha^{-1} \circ \tilde{\chi}_\alpha$, $\alpha \in A$, combine to mappings $E \to \tilde{E}$ and $\tilde{E} \to E$, respectively, which are morphisms and inverse to one another. □

Remark 2.2.12

1. Given two finite-dimensional vector spaces F_1, F_2 and two open coverings $\{U_{i,\alpha_i} : \alpha_i \in A_i\}$, $i = 1, 2$, of M with associated systems of C^k-mappings

$$\rho_{i,\beta_i\alpha_i} : U_{i,\alpha_i} \cap U_{i,\beta_i} \to \mathrm{GL}(F), \quad (\alpha_i, \beta_i) \in A_i \times A_i,$$

   there arises the question under which conditions the vector bundles E_1 and E_2, defined by these data according to Theorem 2.2.11, are isomorphic over M. The answer is as follows. First, F_1 and F_2 have to be isomorphic so that they can be replaced by $\mathbb{K}^r$ for some $r \in \mathbb{N}$. Second, there exists a common refinement $\{U_\alpha : \alpha \in A\}$ of the open coverings $\{U_{i,\alpha_i} : \alpha_i \in A_i\}$, $i = 1, 2$. By restriction, the $\rho_{i,\beta_i\alpha_i}$ induce mappings

$$\rho_{i,\beta\alpha} : U_\alpha \cap U_\beta \to \mathrm{GL}(r, \mathbb{K}), \quad (\alpha, \beta) \in A \times A,\ i = 1, 2.$$

   Now, E_1 and E_2 are isomorphic iff there exists a system of C^k-mappings $\rho_\alpha : U_\alpha \to \mathrm{GL}(r, \mathbb{K})$, $\alpha \in A$, such that

$$\rho_{2,\beta\alpha}(m) = \rho_\beta^{-1}(m) \cdot \rho_{1,\beta\alpha}(m) \cdot \rho_\alpha(m), \quad m \in U_\alpha \cap U_\beta. \tag{2.2.9}$$

   The proof is left to the reader (Exercise 2.2.6).

2. An open covering $\{U_\alpha : \alpha \in A\}$ together with an associated family of C^k-mappings $\rho_{\beta\alpha} : U_\alpha \cap U_\beta \to \mathrm{GL}(r, \mathbb{K})$, $(\alpha, \beta) \in A \times A$, with the property (2.2.4) is called a 1-cocycle on M with values in the structure group $\mathrm{GL}(r, \mathbb{K})$. Two 1-cocycles are called cohomologous if there exists a system of C^k-mappings $\rho_\alpha : U_\alpha \to \mathrm{GL}(r, \mathbb{K})$, $\alpha \in A$, such that (2.2.9) holds. To be cohomologous is an equivalence relation in the set of 1-cocycles. Passage to equivalence classes, that is, cohomology classes of 1-cocycles, yields a cohomology theory on M which is called the first Čech cohomology of M with values in the structure group $\mathrm{GL}(r, \mathbb{K})$. According to point 1, the cohomology classes correspond bijectively to the isomorphism classes of $\mathbb{K}$-vector bundles over M of class C^k and dimension r.
3. One can show that the first Čech cohomology of M and, correspondingly, the set of isomorphism classes of vector bundles over M do not depend on the degree of differentiability k, see [130, Ch. 4, Thm. 3.5].

Exercises

2.2.1 Let (E, M, π) be a C^k-vector bundle. Consider the C^k-structure on E induced by a system of local trivializations via the method of Remark 2.2.5. Show that this structure coincides with the original C^k-structure.

2.2.2 Let M be a C^k-manifold. Use the system of bijections (2.1.4) associated with an atlas on M to construct a C^k-structure on $\mathrm{T}M$ via the method of Remark 2.2.5.

2.2.3 Construct a smooth structure on the Möbius strip by means of the method of Remark 2.2.5, using the local trivializations $(U_\pm, \chi_\pm)$ of Example 2.2.4. Show that this structure coincides with the one constructed in Example 1.1.12.

2.2.4 Prove Proposition 2.2.9.

2.2.5 Prove the assertions about partial derivatives stated in Remark 2.2.10.

2.2.6 Prove the criterion for the isomorphy of two vector bundles over M stated in Remark 2.2.12/1.

2.2.7 Let (E_1, π_1, M_1) and (E_2, π_2, M_2) be $\mathbb{K}$-vector bundles of class C^k and of dimensions r_1 and r_2. Define $E := E_1 \times E_2$, $M := M_1 \times M_2$ and $\pi := \pi_1 \times \pi_2 : E_1 \times E_2 \to M_1 \times M_2$. For $(m_1, m_2) \in M_1 \times M_2$, equip $E_{(m_1,m_2)} := \pi^{-1}(m) \equiv E_{1,m_1} \times E_{2,m_2}$ with the linear structure of the direct sum $E_{1,m_1} \oplus E_{2,m_2}$. Show that (E, M, π) is a $\mathbb{K}$-vector bundle of class C^k and dimension $r_1 + r_2$. It is called the direct product of (E_1, π_1, M_1) and (E_2, π_2, M_2).

2.2.8 Let M_1 and M_2 be C^k-manifolds. Let $\mathrm{pr}_i : M_1 \times M_2 \to M_i$ denote the natural projections to the factors. Show that the following mapping is a vertical vector bundle isomorphism:

$$\Phi : \mathrm{T}(M_1 \times M_2) \to \mathrm{T}M_1 \times \mathrm{T}M_2, \qquad \Phi(X) := \big(\mathrm{pr}_1'(X), \mathrm{pr}_2'(X)\big).$$

2.3 Sections and Frames

The notion of section generalizes the concept of a function on a manifold with values in a finite-dimensional vector space.

Definition 2.3.1 (Section) Let (E, M, π) be a $\mathbb{K}$-vector bundle of class C^k. A section (or cross section) of (E, M, π) is a C^k-mapping $s : M \to E$ such that $\pi \circ s = \mathrm{id}_M$.

A local section of (E, M, π) over an open subset $U \subset M$ is a section of the vector bundle (E_U, U, π_U).

Remark 2.3.2

1. Let s be a section and let (U, χ) be a local trivialization of the $\mathbb{K}$-vector bundle (E, M, π) of class C^k. The mapping
$$\mathrm{pr}_F \circ \chi \circ s_{\restriction U} : U \to F \tag{2.3.1}$$
is called the local representative of s with respect to (U, χ). Since local trivializations are diffeomorphisms, a mapping $s : M \to E$ satisfying $\pi \circ s = \mathrm{id}_M$ is of class C^k (and hence a section) iff so are all local representatives of s with respect to a system of local trivializations.
2. Every vector bundle admits a distinguished section $m \mapsto 0_m$, called the zero section.
3. The set of all sections of a $\mathbb{K}$-vector bundle (E, M, π) of class C^k is denoted by $\Gamma(E)$. It carries the structure of a real vector space and of a bimodule over the algebra $C^\infty(M)$, with all operations defined pointwise (Exercise 2.3.1).
4. A local section need not be extendable to a global section, as is shown by the example of $M = \mathbb{R}$, $E = M \times \mathbb{R}$, $U = \mathbb{R}_+$ and $s(x) = (x, \frac{1}{x})$. There holds, however, the following weaker extension property. For every $m \in U$, there exists an open neighbourhood V of m in U and a section $\tilde{s}$ of E such that $s_{\restriction V} = \tilde{s}_{\restriction V}$ (Exercise 2.3.2).

Example 2.3.3

1. (Local) sections of the tangent bundle $\mathrm{T}M$ of a C^k-manifold M are called (local) vector fields on M. They will usually be denoted by $X, Y, \ldots$ and the vector space $\Gamma(\mathrm{T}M)$ will be denoted by $\mathfrak{X}(M)$. To be consistent with the previous notation X_m for a tangent vector at $m \in M$, for the value of the vector field X at the point m we will often write X_m instead of $X(m)$. Note that since $\mathrm{T}M$ is of class C^{k-1}, so are vector fields. If (U, κ) is a local chart on M, for every $i = 1, \ldots, \dim M$, the mapping
$$\partial_i^\kappa : U \to (\mathrm{T}M)_U, \qquad \partial_i^\kappa(m) := \partial_{i,m}^\kappa,$$
is a section of $(\mathrm{T}M)_U = \mathrm{T}U$. Since the representative of this section with respect to the global trivialization of $\mathrm{T}U$ induced by κ is given by the constant mapping

whose value is the i-th standard basis vector in $\mathbb{R}^{\dim M}$, ∂_i^κ is of class C^{k-1}. Hence, ∂_i^κ is a local vector field on M.
2. If the vector bundle E over M is given in terms of a $\mathbb{K}$-vector space F, an open covering $\{U_\alpha\}$ of M and a family of C^k-mappings $\rho_{\beta\alpha} : U_\alpha \cap U_\beta \to \mathrm{GL}(F)$ satisfying (2.2.4), then its sections of class C^k correspond bijectively to families $\{s_\alpha\}$ of C^k-mappings $s_\alpha : U_\alpha \to F$ satisfying $s_\alpha(m) = \rho_{\alpha\beta}(m) s_\beta(m)$ for all $\alpha, \beta \in A$ and $m \in U_\alpha \cap U_\beta$.

Remark 2.3.4

1. Let V be a finite-dimensional real vector space and let $M \subset V$ be an open subset. Via the natural identification of $\mathrm{T}M$ with $M \times V$ of Remark 2.1.4/1, vector fields X on M correspond bijectively to smooth mappings[2] $X : M \to V$. By construction, for all $v \in M$ and $f \in C^\infty(M)$, we have

$$X_v(f) = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} f\big(v + tX(v)\big). \tag{2.3.2}$$

2. Let V and W be finite-dimensional real vector spaces and let $M \subset V$ be the level set of a C^k-mapping $f : V \to W$ at a regular value. Via the natural identification of $\mathrm{T}M$ with the embedded C^{k-1}-submanifold $\{(v, X) \in M \times V : X \in \ker(f'(v))\}$ of $V \times V$, see Remark 2.1.4/2, vector fields X on M correspond bijectively to C^{k-1}-mappings $X : M \to V$ satisfying $X(v) \in \ker f'(v)$.

By means of a local trivialization, sections are identified locally with the graphs of their local representatives. This implies

Proposition 2.3.5 *Let (E, M, π) be a $\mathbb{K}$-vector bundle of class C^k and let $s \in \Gamma(E)$. Then, (M, s) is an embedded C^k-submanifold of E.*

Proof Let $m \in M$. According to Remark 1.6.13/3, we have to show that there exists an open neighbourhood V of $s(m)$ in E such that $(s^{-1}(V), s_{\restriction s^{-1}(V)})$ is an embedded C^k-submanifold of E. Choose a local trivialization (U, χ) of E at m and let $V = \pi^{-1}(U)$. Then, $s^{-1}(V) = U$ and hence we have to show that $(U, s_{\restriction U})$ is an embedded C^k-submanifold of $\pi^{-1}(U)$. Since χ is a diffeomorphism and $(U, \chi \circ s_{\restriction U})$ is the graph of the local representative of s with respect to the local trivialization (U, χ), the latter follows from Example 1.6.12/2. □

Now let (E_1, M_1, π_1) and (E_2, M_2, π_2) be $\mathbb{K}$-vector bundles of class C^k, let $\Phi : E_1 \to E_2$ be a morphism and let $\varphi : M_1 \to M_2$ be the projection of Φ.

[2] Denoted by the same symbol.

Definition 2.3.6 (Φ-relation and transport operator)

1. Sections $s_1 \in \Gamma(E_1)$ and $s_2 \in \Gamma(E_2)$ are said to be Φ-related if they satisfy

$$\Phi \circ s_1 = s_2 \circ \varphi.$$

2. If φ is a C^k-diffeomorphism, the following mapping is called the transport operator of Φ:

$$\Phi_* : \Gamma(E_1) \to \Gamma(E_2), \qquad \Phi_* s := \Phi \circ s \circ \varphi^{-1}. \tag{2.3.3}$$

The following proposition lists the properties of the transport operator (Exercise 2.3.3).

Proposition 2.3.7 *Let (E_1, M_1, π_1) and (E_2, M_2, π_2) be $\mathbb{K}$-vector bundles of class C^k and let $\Phi : E_1 \to E_2$ be a morphism whose projection $\varphi : M_1 \to M_2$ is a diffeomorphism. The transport operator Φ_* has the following properties.*

1. *Φ_* is linear. If Φ is an isomorphism of vector bundles, Φ_* is an isomorphism of vector spaces and there holds $(\Phi^{-1})_* = (\Phi_*)^{-1}$.*
2. *For every $s \in \Gamma(E_1)$, s is Φ-related to $\Phi_* s$.*
3. *For every $s \in \Gamma(E_1)$ and $f \in C^k(M_1)$, there holds $\Phi_*(fs) = ((\varphi^{-1})^* f)\Phi_* s$.*
4. *If $\Psi : E_2 \to E_3$ is another morphism whose projection is a diffeomorphism, then $(\Psi \circ \Phi)_* = \Psi_* \circ \Phi_*$.*

Remark 2.3.8 In the case of vector fields, it is common to speak of φ-relation rather than φ'-relation. Thus, $X_i \in \mathfrak{X}(M_i)$, $i = 1, 2$, are φ-related iff

$$\varphi' \circ X_1 = X_2 \circ \varphi. \tag{2.3.4}$$

Next, we turn to the discussion of (local) frames.

Definition 2.3.9 (Local frame) Let (E, M, π) be a $\mathbb{K}$-vector bundle of class C^k and dimension l, let $U \subset M$ be open and let $\mathscr{B} = \{s_1, \ldots, s_r\}$ be a system of local sections of E over U. $\mathscr{B}$ is said to be pointwise linearly independent if the system $\{s_1(m), \ldots, s_r(m)\}$ is linearly independent in E_m for all $m \in U$. In this case, $\mathscr{B}$ is called a local r-frame (frame if $r = l$) in E over U. If $U = M$, $\mathscr{B}$ is called a global r-frame (global frame if $r = l$).

Local frames provide bases in the fibres over their domain and hence allow for the expansion of local sections.

Proposition 2.3.10 *Let (E, M, π) be a $\mathbb{K}$-vector bundle of class C^k, let $U \subset M$ be open and let $\{s_1, \ldots, s_l\}$ be a local frame in E over U. The assignment of $f^i s_i$ (summation convention) to an l-tuple $(f^1, \ldots, f^l)$ of $\mathbb{K}$-valued C^k-functions on U defines a bijection from $\prod_{i=1}^l C^k(U, \mathbb{K})$ onto $\Gamma(E_U)$.*

Proof Obviously, for every $(f^1, \dots, f^l) \in \prod_{i=1}^l C^k(U, \mathbb{K})$, the sum $f^i s_i$ is a C^k-section of E_U. Conversely, let $s \in \Gamma(E_U)$. By expanding $s(m)$ with respect to the basis $\{s_1(m), \dots, s_l(m)\}$ of E_m for all $m \in U$, we obtain functions $f^i : U \to \mathbb{K}$ satisfying $s_{\restriction U} = f^i s_i$. For every $m \in U$, $(f^1(m), \dots, f^l(m))$ is the unique solution of a system of linear equations whose coefficients depend differentiably of class C^k on m. Hence, $f^i \in C^k(U)$. □

Example 2.3.11

1. Let M be a C^k-manifold of dimension n and let (U, κ) be a local chart on M. Since $\{\partial^\kappa_{1,m}, \dots, \partial^\kappa_{n,m}\}$ is a basis in $\mathrm{T}_m M$ for all $m \in U$, the system $\{\partial^\kappa_1, \dots, \partial^\kappa_n\}$ is a local frame in $\mathrm{T}M$ over U. Thus, over U, vector fields $X \in \mathfrak{X}(M)$ can be represented as $X_{\restriction U} = X^i \partial^\kappa_i$ with $X^i \in C^{k-1}(U)$. According to (1.4.15) and (1.4.16), the coefficient functions X^i are given by $X^i(m) = X_m(\kappa^i)$, where $i = 1, \dots, n$.
2. Let (E, M, π) be a $\mathbb{K}$-vector bundle of class C^k with typical fibre F, let (U, χ) be a local trivialization and let $\{e_1, \dots, e_r\}$ be a linearly independent system in F. Define local sections s_i of E over U by

$$s_i(m) := \chi^{-1}(m, e_i), \quad i = 1, \dots, r. \tag{2.3.5}$$

These sections are of class C^k, because their local representatives with respect to (U, χ) are the constant mappings $m \mapsto e_i$. Hence, the system $\{s_1, \dots, s_r\}$ is a local r-frame in E over U.

As the second example suggests, local frames are closely related to local trivializations.

Proposition 2.3.12 *Let (E, M, π) be a $\mathbb{K}$-vector bundle of class C^k with typical fibre F and let $U \subset M$ be open. By virtue of* (2.3.5), *every basis of F defines a bijection between local trivializations $\chi : \pi^{-1}(U) \to U \times F$ and local frames in E over U. In particular,*

1. *there exists a local trivialization of E over U iff there exists a local frame in E over U.*
2. *E is trivial iff there exists a global frame.*

Proof Let $\{e_1, \dots, e_l\}$ be a basis of F. That a local trivialization over U defines a local frame over U has been shown in Example 2.3.11/2. Conversely, for a given local frame $\{s_1, \dots, s_l\}$ in E over U, expand $x \in E_U$ as $x = x^i s_i(\pi(x))$ and define a mapping $\chi : \pi^{-1}(U) \to U \times F$ by $\chi(x) := (\pi(x), x^i e_i)$, $x \in E_U$. The mapping χ is a bijection and satisfies conditions 2a and 2b of Definition 2.2.1. Thus, to show that χ is a local trivialization, it remains to check that χ and χ^{-1} are of class C^k (Exercise 2.3.4). Finally, assertions 1 and 2 are obvious. □

Example 2.3.13

1. Let M be a C^k-manifold of dimension n, let (U, κ) be a local chart on M and let (U, χ) be the local trivialization of $\mathrm{T}M$ induced by this chart via (2.1.4). The bijection between local frames over U and local trivializations over U, defined by the standard basis of $\mathbb{R}^n$ via (2.3.5), assigns to (U, χ) the local frame $\{\partial_1^\kappa, \dots, \partial_n^\kappa\}$.
2. Consider the smooth real vector bundle E given by the Möbius strip, cf. Example 2.2.4. Since E has dimension 1, a global frame in E is just a nowhere vanishing section. Since the base manifold is S^1, sections of E correspond to closed smooth curves in E winding around exactly once.[3] Since any such curve must cross the zero section, E does not admit a global frame and is hence not globally trivial, cf. Proposition 2.3.12.

Remark 2.3.14 Using the description of vector bundles in terms of coverings and transition functions as explained in Remark 2.2.12, one can show that, up to isomorphy over S^1, the Möbius strip and the product vector bundle $\mathrm{S}^1 \times \mathbb{R}$ are the only real vector bundles of dimension 1 over S^1.

The following proposition collects useful extension results. The proof is left to the reader (Exercise 2.3.5).

Proposition 2.3.15 *Let (E, M, π) be a $\mathbb{K}$-vector bundle of class C^k and dimension l and let $m \in M$.*

1. *Let $\{e_1, \dots, e_l\}$ be a basis of E_m. There exists an open neighbourhood U of m and a local frame $\{s_1, \dots, s_l\}$ over U such that $s_i(m) = e_i$, $i = 1, \dots, l$.*
2. *Let $s_1, \dots, s_r$ be local sections over neighbourhoods $U_1, \dots U_r$ of m such that the system $\{s_1(m), \dots, s_r(m)\}$ is linearly independent in E_m. Then, there exists an open neighbourhood $U \subset U_1 \cap \dots \cap U_r$ of m such that the system $\{s_{1\restriction U}, \dots, s_{r\restriction U}\}$ is a local r-frame in E over U.*
3. *Let $\{s_1, \dots, s_r\}$ be a local r-frame over a neighbourhood U of m. Then, there exist local sections $s_{r+1}, \dots, s_l$ over $V \subset U$ such that the system $\{s_{1\restriction V}, \dots, s_{r\restriction V}, s_{r+1}, \dots, s_l\}$ is a local frame over V.*

As an application, we briefly discuss manifolds whose tangent bundle is trivial.

Definition 2.3.16 A C^k-manifold is called parallelizable if its tangent bundle is trivial.

According to Proposition 2.3.12, a differentiable manifold M of dimension n is parallelizable iff there exist n pointwise linearly independent vector fields on M.

[3] And with tangent vectors being nowhere parallel to the fibres, but this is not relevant for the argument.

Proposition 2.3.17 *The spheres* S^1, S^3 *and* S^7 *are parallelizable.*

Proof Since S^n is a level set of the smooth function $f : \mathbb{R}^{n+1} \to \mathbb{R}$, $f(\mathbf{x}) = \|\mathbf{x}\|^2$, we can use the natural representation of smooth vector fields on S^n by smooth mappings $X : \mathrm{S}^n \to \mathbb{R}^{n+1}$ satisfying $\mathbf{x} \cdot X(\mathbf{x}) = 0$, cf. Example 2.1.3 and Remark 2.3.4/2. In the case of S^1 we identify $\mathbb{R}^2$ with $\mathbb{C}$ via $\mathbf{x} = (x_1, x_2) \mapsto \hat{\mathbf{x}} := x_1 + \mathrm{i}x_2$. Then, $\mathbf{x} \cdot \mathbf{y} = \mathrm{Re}(\bar{\hat{\mathbf{x}}}\hat{\mathbf{y}})$ and vector fields on S^1 are represented by mappings $X : \mathrm{S}^1 \subset \mathbb{C} \to \mathbb{C}$ satisfying $\mathrm{Re}(\overline{z}X(z)) = 0$. This condition holds for example for $X(z) := z\mathrm{i}$. Since this function is nowhere vanishing, the corresponding vector field is nowhere vanishing and hence forms a frame in TS^1. In the case of S^3, we identify $\mathbb{R}^4$ with the quaternions $\mathbb{H}$ via $\mathbf{x} \mapsto \hat{\mathbf{x}} := x_1\mathbf{1} + x_2\mathbf{i} + x_3\mathbf{j} + x_4\mathbf{k}$. Then, $\mathbf{x} \cdot \mathbf{y} = \mathrm{Re}(\bar{\hat{\mathbf{x}}}\hat{\mathbf{y}})$, where $\bar{\hat{\mathbf{x}}}$ now denotes quaternionic conjugation, and vector fields on S^3 are represented by mappings $X : \mathrm{S}^3 \subset \mathbb{H} \to \mathbb{H}$ satisfying $\mathrm{Re}(\overline{\mathbf{q}}X(\mathbf{q})) = 0$. For $l = 1, 2, 3$, define $X_l : \mathbb{H} \to \mathbb{H}$ by

$$X_1(\mathbf{q}) = \mathbf{qi}, \qquad X_2(\mathbf{q}) = \mathbf{qj}, \qquad X_3(\mathbf{q}) = \mathbf{qk}.$$

Then, $\mathrm{Re}(\overline{\mathbf{q}}X_l(\mathbf{q})) = 0$ and $\mathrm{Re}(\overline{X_l(\mathbf{q})}X_j(\mathbf{q})) = \delta_{lj}$. Hence, the X_l restrict to vector fields on S^3 and these vector fields are pointwise linearly independent. In the case of S^7, the proof is analogous, with quaternions replaced by octonions.[4] □

Remark 2.3.18

1. Since TS^1 is isomorphic to the product vector bundle $\mathrm{S}^1 \times \mathbb{R}$, one can rephrase Remark 2.3.14 as follows. Up to isomorphy over S^1, the tangent bundle of S^1 and the Möbius strip are the only real vector bundles of dimension 1 over S^1.
2. The construction of pointwise linearly independent vector fields on the spheres S^1, S^3 and S^7 presented in the proof of Proposition 2.3.17 carries over to the unit spheres of $\mathbb{C}^k$, $\mathbb{H}^k$ and $\mathbb{O}^k$, where $\mathbb{O}$ denotes the octonions. Thus, for $r = 2, 4, 8$ and $k = 1, 2, \ldots$ there exist $r - 1$ pointwise linearly independent vector fields on the sphere S^{rk-1}. In case $k = 1$, these vector fields constitute a global frame, whereas in the other cases they constitute just a global $(r - 1)$-frame. While there may exist more than $r - 1$ pointwise linearly independent vector fields, there does not exist a global frame for any odd-dimensional sphere except for S^1, S^3 and S^7. More precisely, Adams showed that the maximum number of pointwise linearly independent vector fields on an odd-dimensional sphere is given by the corresponding Radon-Hurwitz number [4]. On the other hand, on an even-dimensional sphere, every vector field has a zero. This is known as the Hairy Ball Theorem. For a proof, see for example [6]. As a consequence, S^1, S^3 and S^7 are the only spheres which are parallelizable.

Exercises

2.3.1 Show that $\Gamma(E)$ carries the structure of a real vector space and of a bimodule over the algebra $C^\infty(M)$, cf. Remark 2.3.2/3.

[4]For a guide to octonions, see [29].

2.3.2 Prove the statement of Remark 2.3.2/4.
2.3.3 Prove Proposition 2.3.7.
2.3.4 Complete the proof of Proposition 2.3.12 by showing that the mapping χ defined there as well as its inverse are of class C^k.
2.3.5 Prove Proposition 2.3.15.

2.4 Vector Bundle Operations

Every operation with vector spaces defines an operation with vector bundles by fibrewise application. Below, we will discuss the most important of these operations in the form of examples. The construction uses the method of Remark 2.2.5. It will be explained in some detail for the dual vector bundle and the direct sum of vector bundles. The other operations are then given without further explanations.

Throughout this section, let E, E_1 and E_2 be $\mathbb{K}$-vector bundles over M of class C^k. Let, respectively, π, π_1 and π_2 be their projections and l, l_1 and l_2 their dimensions. Choose, respectively, typical fibres F, F_1 and F_2 and local trivializations (U_α, χ_α), $(U_\alpha, \chi_{1\alpha})$ and $(U_\alpha, \chi_{2\alpha})$ over an appropriate open covering $\{U_\alpha : \alpha \in A\}$ of M.

Example 2.4.1 (Dual vector bundle) Take the dual vector space E_m^* of each fibre E_m of E and define the set E^* as the disjoint union

$$E^* = \bigsqcup_{m \in M} E_m^* .$$

Let $\pi^{E^*} : E^* \to M$ be the natural projection to the index set. Define mappings

$$\chi_\alpha^{E^*} : \left(\pi^{E^*}\right)^{-1}(U_\alpha) \to U_\alpha \times F^*, \qquad \chi_\alpha^{E^*}(\xi) = \left(m, \left(\chi_{\alpha,m}^{\mathrm{T}}\right)^{-1}(\xi)\right), \tag{2.4.1}$$

where $m = \pi^{E^*}(\xi)$ and $\chi_{\alpha,m}^{\mathrm{T}} : F^* \to E_m^*$ denotes the dual linear mapping of $\chi_{\alpha,m} : E_m \to F$. The corresponding transition mappings are given by

$$\chi_\beta^{E^*} \circ \left(\chi_\alpha^{E^*}\right)^{-1}(m, \mu) = \left(m, \left(\left(\chi_\alpha \circ \chi_\beta^{-1}\right)_{\restriction \{m\} \times F}\right)^{\mathrm{T}} \mu\right)$$

with $m \in U_\alpha \cap U_\beta$ and $\mu \in F^*$. They are of class C^k, because so are the transition mappings $\chi_\alpha \circ \chi_\beta^{-1}$ of E. Thus, according to Remark 2.2.5, if we equip E^* with the C^k-structure induced by the family of mappings $\{\chi_\alpha^{E^*}, \alpha \in A\}$, then (E^*, M, π^{E^*}) is a $\mathbb{K}$-vector bundle of class C^k, called the dual vector bundle of E. It has the same dimension as E, typical fibre F^*, and $\{(U_\alpha, \chi_\alpha^{E^*}) : \alpha \in A\}$ is a system of local trivializations.

Let $\{s_1, \ldots, s_l\}$ be a local frame in E over $U \subset M$. For $m \in U$, let $s(m)^{*1}, \ldots, s(m)^{*l}$ denote the elements of the basis of E_m^* which is dual to the basis $\{s_1(m), \ldots, s_l(m)\}$ of E_m. Define local sections s^{*i} in E^* by

$$s^{*i}(m) := s(m)^{*i}, \quad i = 1, \ldots, l. \tag{2.4.2}$$

Using Proposition 2.3.12 it is easy to see that these local sections are of class C^k and form a local frame of E^*, called the dual local frame or coframe.

The pointwise evaluation mappings $E^*_m \times E_m \to \mathbb{K}$ combine to a natural pairing

$$\Gamma\big(E^*_U\big) \times \Gamma(E_U) \to C^k(U, \mathbb{K}), \qquad (\sigma, s) \mapsto \langle \sigma, s\rangle,$$

also denoted by $\sigma(s)$ or $s(\sigma)$. In terms of this pairing, sections of E^* can be expanded over U as

$$\sigma = \sigma(s_i)s^{*i} \equiv \langle \sigma, s_i\rangle s^{*i}. \tag{2.4.3}$$

Let E_a and E_b be $\mathbb{K}$-vector bundles of class C^k over M_a and M_b, respectively, and let $\Phi : E_a \to E_b$ be a morphism projecting to a diffeomorphism $\varphi : M_a \to M_b$. For every $m \in M_b$, the linear mappings $\Phi_{\varphi^{-1}(m)} : E_{a,\varphi^{-1}(m)} \to E_{b,m}$ induce dual linear mappings which combine to a fibre-preserving and fibrewise linear mapping

$$\Phi^{\mathrm{T}} : E^*_b \to E^*_a, \qquad \big\langle \Phi^{\mathrm{T}}(\xi), x\big\rangle := \big\langle \xi, \Phi_{\varphi^{-1}(m)}(x)\big\rangle, \tag{2.4.4}$$

where $m \in M_b$, $\xi \in E^*_{b,m}$ and $x \in E_{a,\varphi^{-1}(m)}$, which is a morphism projecting to the C^k-diffeomorphism φ^{-1} (Exercise 2.4.1). It is called the dual morphism of Φ. Via (2.3.3), the dual morphism induces a transport operator Φ^{T}_* of sections. More generally, by duality, every morphism $\Phi : E_a \to E_b$ induces the following operation on sections, called the pull-back,

$$\Phi^* : \Gamma\big(E^*_b\big) \to \Gamma\big(E^*_a\big), \qquad \big\langle \big(\Phi^*\sigma\big)(m), x\big\rangle := \big\langle \sigma \circ \varphi(m), \Phi(x)\big\rangle, \tag{2.4.5}$$

where $m \in M_a$ and $x \in E_{a,m}$. Indeed, if Φ projects to a diffeomorphism, then the pull-back is given by

$$\Phi^*\sigma = \Phi^{\mathrm{T}} \circ \sigma \circ \varphi = \Phi^{\mathrm{T}}_*\sigma, \tag{2.4.6}$$

that is, it coincides with the transport operator of the dual morphism.[5]

Example 2.4.2 (Direct sum) Take the direct sum $E_{1,m} \oplus E_{2,m}$ of the fibres over each point $m \in M$ and define

$$E_1 \oplus E_2 = \bigsqcup_{m\in M} E_{1,m} \oplus E_{2,m}.$$

Let $\pi^{\oplus} : E_1 \oplus E_2 \to M$ be the natural projection. Define mappings

$$\chi^{\oplus}_\alpha : \big(\pi^{\oplus}\big)^{-1}(U_\alpha) \to U_\alpha \times (F_1 \oplus F_2)$$

by

$$\chi^{\oplus}_\alpha(x_1, x_2) := \big(m, \big(\chi_{1\alpha,m}(x_1), \chi_{2\alpha,m}(x_2)\big)\big),$$

[5]Taking into account that Φ^{T} projects to φ^{-1}.

where $m = \pi^{\oplus}((x_1, x_2))$. By similar arguments as for the dual vector bundle, one can check that $(E_1 \oplus E_2, M, \pi^{\oplus})$ is a $\mathbb{K}$-vector bundle of class C^k and that the mappings $\chi_\alpha^{\oplus}$ provide a system of local trivializations. $E_1 \oplus E_2$ is called the direct sum of E_1 and E_2. It has dimension $l_1 + l_2$ and typical fibre $F_1 \oplus F_2$. Next, note that every local section s_i of E_i over U can be viewed as a local section of $E_1 \oplus E_2$ in an obvious way. Thus, given local frames $\{s_{i,1}, \ldots, s_{i,l_i}\}$ in E_i, the collection

$$\{s_{1,1}, \ldots, s_{1,l_1}, s_{2,1}, \ldots, s_{2,l_2}\}$$

constitutes a local frame in $E_1 \oplus E_2$. Finally, let E_{a1}, E_{a2} and E_{b1}, E_{b2} be $\mathbb{K}$-vector bundles of class C^k over M_a and M_b, respectively, and let $\Phi_i : E_{ai} \to E_{bi}$, $i = 1, 2$, be morphisms projecting to the same mapping $\varphi : M_a \to M_b$. The linear mappings $\Phi_{i,m} : E_{ai,m} \to E_{bi,\varphi(m)}$ induce linear mappings

$$\Phi_{1,m} \oplus \Phi_{2,m} : E_{a1,m} \oplus E_{a2,m} \to E_{b1,\varphi(m)} \oplus E_{b2,\varphi(m)},$$

which combine to a morphism projecting to φ,

$$\Phi_1 \oplus \Phi_2 : E_{a1} \oplus E_{a2} \to E_{b1} \oplus E_{b2}, \qquad (\Phi_1 \oplus \Phi_2)_m := \Phi_{1,m} \oplus \Phi_{2,m}, \quad (2.4.7)$$

where $m \in M_a$. It is called the direct sum of Φ_1 and Φ_2.

Example 2.4.3 (Tensor product) Define

$$E_1 \otimes E_2 = \bigsqcup_{m \in M} E_{1,m} \otimes E_{2,m},$$

denote the canonical projection by $\pi^{\otimes} : E_1 \otimes E_2 \to M$ and take the system of induced local trivializations $\chi_\alpha^{\otimes} : (\pi^{\otimes})^{-1}(U_\alpha) \to U_\alpha \times (F_1 \otimes F_2)$ defined by

$$\chi_\alpha^{\otimes}(x_1 \otimes x_2) = \big(m, \chi_{1\alpha,m}(x_1) \otimes \chi_{2\alpha,m}(x_2)\big),$$

where $m = \pi^{\otimes}(x_1 \otimes x_2)$. Then, $(E_1 \otimes E_2, M, \pi^{\otimes})$ is a C^k-vector bundle, called the tensor product of E_1 and E_2. Its typical fibre is $F_1 \otimes F_2$ and its dimension is $l_1 l_2$. Every pair of local sections s_i of E_i over U, $i = 1, 2$, defines a local section $s_1 \otimes s_2$ of $E_1 \otimes E_2$ by

$$(s_1 \otimes s_2)(m) := s_1(m) \otimes s_2(m), \quad m \in U, \quad (2.4.8)$$

which is called the tensor product of s_1 and s_2. If $\{s_{i,1}, \ldots, s_{i,l_i}\}$ are local frames in E_i, $i = 1, 2$, then

$$\{s_{1,i} \otimes s_{2,j} : i = 1, \ldots, l_1, j = 1, \ldots, l_2\}$$

is a local frame in $E_1 \otimes E_2$. For $\mathbb{K}$-vector bundle morphisms $\Phi_j : E_{aj} \to E_{bj}$, $j = 1, 2$, projecting to the same mapping $\varphi : M_a \to M_b$, the tensor product is the morphism $\Phi_1 \otimes \Phi_2 : E_{a1} \otimes E_{a2} \to E_{b1} \otimes E_{b2}$ defined by

$$(\Phi_1 \otimes \Phi_2)_m(x_1 \otimes x_2) := \Phi_{1,m}(x_1) \otimes \Phi_{2,m}(x_2). \quad (2.4.9)$$

It projects to φ as well.

Example 2.4.4 (Tensor bundles) The tensor bundle of E of type (p, q) is defined to be

$$\mathbb{T}^q_p E := E^* \otimes \overset{p}{\cdots} \otimes E^* \otimes E \otimes \overset{q}{\cdots} \otimes E.$$

Its fibres are the p-fold covariant and q-fold contravariant tensor products $\mathbb{T}^q_p E_m$. Hence, the dimension is l^{p+q} and the elements of $\mathbb{T}^q_p E_m$ are linear combinations of $\xi_1 \otimes \cdots \otimes \xi_p \otimes x_1 \otimes \cdots \otimes x_q$, where $x_i \in E_m$ and $\xi_i \in E^*_m$. The projection is denoted by $\pi^{\otimes} : \mathbb{T}^q_p E \to M$ and the typical fibre is $\mathbb{T}^q_p F$. We will view elements of $\mathbb{T}^q_p E_m$ as $(p+q)$-linear mappings

$$u : E_m \times \overset{p}{\cdots} \times E_m \times E^*_m \times \overset{q}{\cdots} \times E^*_m \to \mathbb{R},$$

thus using the natural isomorphism which assigns to $\xi_1 \otimes \cdots \otimes \xi_p \otimes x_1 \otimes \cdots \otimes x_q$ the mapping

$$u(y_1, \ldots, y_p, \eta_1, \ldots, \eta_q) = \xi_1(y_1) \cdots \xi_p(y_p) \eta_1(x_1) \cdots \eta_q(x_q).$$

Then, the tensor product of $u_i \in \mathbb{T}^{q_i}_{p_i} E_m$, $i = 1, 2$, is given by

$$\begin{aligned} &u_1 \otimes u_2(x_1, \ldots, x_{p_1+p_2}, \xi_1, \ldots, \xi_{q_1+q_2}) \\ &\quad := u_1(x_1, \ldots, x_{p_1}, \xi_1, \ldots, \xi_{q_1}) u_2(x_{p_1+1}, \ldots, x_{p_1+p_2}, \xi_{q_1+1}, \ldots, \xi_{q_1+q_2}) \end{aligned} \tag{2.4.10}$$

for all $x_j \in E_m$ and $\xi_j \in E^*_m$. Accordingly, local sections τ of $\mathbb{T}^q_p E$ over U can be viewed as mappings

$$\tau : \Gamma(E_U) \times \overset{p}{\cdots} \times \Gamma(E_U) \times \Gamma\big(E^*_U\big) \times \overset{q}{\cdots} \times \Gamma\big(E^*_U\big) \to C^k(U) \tag{2.4.11}$$

which are $C^k(U)$-linear in every argument. Every pair of local sections τ_i of $\mathbb{T}^{q_i}_{p_i} E$, $i = 1, 2$, defines a local section $\tau_1 \otimes \tau_2$ in $\mathbb{T}^{q_1+q_2}_{p_1+p_2} E$ by

$$(\tau_1 \otimes \tau_2)(m) := \tau_1(m) \otimes \tau_2(m).$$

On the level of the mappings (2.4.11), $\tau_1 \otimes \tau_2$ is given by (2.4.10), with u_i replaced by τ_i and x_j and ξ_j replaced by local sections in E and E^*, respectively. In particular, if $\{s_1, \ldots, s_l\}$ is a local frame in E over U, then

$$\big\{s^{*i_1} \otimes \cdots \otimes s^{*i_p} \otimes s_{j_1} \otimes \cdots \otimes s_{j_q} : i_1, \ldots, i_p, j_1, \ldots, j_q = 1, \ldots, l\big\}$$

is a local frame in $\mathbb{T}^q_p E$. Every $\tau \in \Gamma(\mathbb{T}^q_p E)$ can be decomposed over U as

$$\tau_{\restriction U} = \tau^{j_1 \ldots j_q}_{i_1 \ldots i_p} s^{*i_1} \otimes \cdots \otimes s^{*i_p} \otimes s_{j_1} \otimes \cdots \otimes s_{j_q}$$

with

$$\tau^{j_1\dots j_q}_{i_1\dots i_p}(m) = \tau(m)\big(s_{i_1}(m),\dots,s_{i_p}(m),s^{*j_1}(m),\dots,s^{*j_q}(m)\big) \tag{2.4.12}$$

(Exercise 2.4.2). Finally, according to (2.4.9), every isomorphism $\Phi : E_a \to E_b$ of $\mathbb{K}$-vector bundles of class C^k induces isomorphisms

$$\Phi^{\otimes} : \mathbb{T}^q_p E_a \to \mathbb{T}^q_p E_b, \qquad \Phi^{\otimes} := \big(\Phi^{-1}\big)^{\mathrm{T}} \otimes \overset{p}{\cdots} \otimes \big(\Phi^{-1}\big)^{\mathrm{T}} \otimes \Phi \otimes \overset{q}{\cdots} \otimes \Phi, \tag{2.4.13}$$

with the same projection. On the level of $(p+q)$-linear mappings, $\Phi^{\otimes}$ takes the form

$$\begin{aligned}&\big(\Phi^{\otimes}u\big)(x_1,\dots,x_p,\xi_1,\dots,\xi_q)\\ &\quad = u\big(\Phi^{-1}(x_1),\dots,\Phi^{-1}(x_p),\Phi^{\mathrm{T}}(\xi_1),\dots,\Phi^{\mathrm{T}}(\xi_q)\big)\end{aligned} \tag{2.4.14}$$

for all $u \in \mathbb{T}^q_p E_{a,m}$, $x_i \in E_{b,\varphi(m)}$ and $\xi_i \in E^*_{b,\varphi(m)}$. The corresponding transport operators $\Phi^{\otimes}_*$ satisfy

$$\Phi^{\otimes}_*(\tau_1 \otimes \tau_2) = \big(\Phi^{\otimes}_*\tau_1\big) \otimes \big(\Phi^{\otimes}_*\tau_2\big) \tag{2.4.15}$$

for all $\tau_i \in \Gamma(\mathbb{T}^{q_i}_{p_i} E_a)$, and

$$\begin{aligned}&\big(\Phi^{\otimes}_*\tau\big)(s_1,\dots,s_p,\sigma_1,\dots,\sigma_q)\\ &\quad = \tau\big(\Phi^{-1}_* s_1,\dots,\Phi^{-1}_* s_p,\Phi^*\sigma_1,\dots,\Phi^*\sigma_q\big) \circ \varphi^{-1}\end{aligned} \tag{2.4.16}$$

for all $\tau \in \Gamma(\mathbb{T}^q_p E_a)$, $s_i \in \Gamma(E_b)$ and $\sigma_i \in \Gamma(E^*_b)$. Here, $\varphi : M_a \to M_b$ is the projection of Φ.

Example 2.4.5 (Exterior powers) The r-fold exterior power $\bigwedge^r E^*$ has the vector spaces $\bigwedge^r E^*_m$ of antisymmetric r-linear forms on E_m as its fibres. Hence, the dimension is $\binom{l}{r}$. In particular, $\bigwedge^0 E^* = M \times \mathbb{K}$, $\bigwedge^1 E^* = E^*$ and $\bigwedge^r E^* = M \times \{0\}$ (the zero-dimensional vector bundle over M) for $r > l$. The projection is denoted by $\pi^{\wedge} : \bigwedge^r E^* \to M$ and the typical fibre is $\bigwedge^r F^*$. The exterior product of $\eta_i \in \bigwedge^{r_i} E^*_m$ is defined to be the (r_1+r_2)-linear form on E_m given by[6]

$$\begin{aligned}&(\eta_1 \wedge \eta_2)(x_1,\dots,x_{r_1+r_2})\\ &\quad := \frac{1}{r_1!r_2!}\sum_{\pi \in \mathrm{S}_{r_1+r_2}} \operatorname{sign}(\pi)\eta_1(x_{\pi(1)},\dots,x_{\pi(r_1)})\eta_2(x_{\pi(r_1+1)},\dots,x_{\pi(r_1+r_2)}),\end{aligned} \tag{2.4.17}$$

[6] Beware that there exist different conventions concerning the choice of the factor in Formula (2.4.17).

for all $x_i \in E_m$. A local section σ in $\bigwedge^r E^*$ over U can be viewed as an antisymmetric mapping

$$\sigma : \Gamma(E_U)\times \overset{r}{\cdots} \times\Gamma(E_U) \to C^k(U) \tag{2.4.18}$$

which is $C^k(U)$-linear in every argument. Every pair of local sections σ_i of $\bigwedge^{r_i} E^*$, $i = 1, 2$, defines a local section $\sigma_1 \wedge \sigma_2$ of $\bigwedge^{r_1+r_2} E^*$ by

$$(\sigma_1 \wedge \sigma_2)(m) := \sigma_1(m) \wedge \sigma_2(m), \quad m \in U. \tag{2.4.19}$$

If we view $\sigma_1 \wedge \sigma_2$ as a mapping (2.4.18), it is given by (2.4.17) with ξ_i replaced by σ_i and x_i replaced by local sections in E. If $\{s_1, \dots, s_l\}$ is a local frame in E, then

$$\left\{s^{*i_1} \wedge \cdots \wedge s^{*i_r} : 1 \le i_1 < \cdots < i_r \le l\right\} \tag{2.4.20}$$

is a local frame in $\bigwedge^r E^*$. Every $\sigma \in \Gamma(\bigwedge^r E^*)$ can be decomposed over U as

$$\sigma_{\restriction U} = \sum_{i_1<\cdots<i_r} \sigma_{i_1\dots i_r} s^{*i_1} \wedge \cdots \wedge s^{*i_r} \tag{2.4.21}$$

with $\sigma_{i_1\dots i_r}(m) = \sigma(m)(s_{i_1}(m), \dots, s_{i_r}(m))$ (Exercise 2.4.2). Next, every $\mathbb{K}$-vector bundle morphism $\Phi : E_a \to E_b$ projecting to a diffeomorphism $\varphi : M_a \to M_b$ induces a morphism $\Phi^{\mathrm{T}\wedge} : \bigwedge^r E_b^* \to \bigwedge^r E_a^*$ projecting to φ^{-1}, defined by

$$\big(\Phi_m^{\mathrm{T}\wedge}(\eta)\big)(x_1, \dots, x_r) := \eta\big(\Phi_{\varphi^{-1}(m)}(x_1), \dots, \Phi_{\varphi^{-1}(m)}(x_r)\big). \tag{2.4.22}$$

This generalizes Formula (2.4.4). Via (2.3.3), $\Phi^{\mathrm{T}\wedge}$ induces a transport operator $(\Phi^{\mathrm{T}\wedge})_*$ of sections. Moreover, the pull-back operation (2.4.5) generalizes in an obvious way to a mapping $\Phi^* : \Gamma(\bigwedge^r E_b^*) \to \Gamma(\bigwedge^r E_a^*)$, given by

$$\big(\big(\Phi^*\sigma\big)(m)\big)(x_1, \dots, x_r) := \big(\sigma \circ \varphi(m)\big)\big(\Phi(x_1), \dots, \Phi(x_r)\big). \tag{2.4.23}$$

Again, if Φ projects to a diffeomorphism, then $\Phi^* = (\Phi^{\mathrm{T}\wedge})_*$.

Example 2.4.6 (Exterior algebra bundle) By composing the operations of exterior power and direct sum one obtains the exterior algebra bundle $\bigwedge E^* = \bigoplus_{i=0}^l \bigwedge^i E^*$, which has dimension 2^l. We retain the notations $\pi^\wedge : \bigwedge E \to M$ for the projection and $\Phi^{\mathrm{T}\wedge} : \bigwedge E_b^* \to \bigwedge E_a^*$ for the morphism induced by a morphism $\Phi : E_a \to E_b$. The local frame in $\bigwedge E^*$ associated with a local frame $\{s_1, \dots, s_l\}$ in E consists of the constant mapping $U \to \mathbb{K}$ given by $m \mapsto 1$ and the local sections (2.4.20) with $r = 1, \dots, l$. In addition to being a vector bundle, $\bigwedge E^*$ is an associative $\mathbb{K}$-algebra bundle[7] of class C^k over M. The exterior product of local sections (2.4.19) induces

[7] In the definition of vector bundle, replace "$\mathbb{K}$-vector space" by "$\mathbb{K}$-algebra" and "linear mapping" by "algebra homomorphism".

a bilinear mapping

$$\Gamma\Big(\bigwedge^{r_1} E^*\Big) \times \Gamma\Big(\bigwedge^{r_2} E^*\Big) \to \Gamma\Big(\bigwedge^{r_1+r_2} E^*\Big)$$

and hence defines on $\Gamma(\bigwedge E^*)$ the structure of an associative $\mathbb{K}$-algebra. By (2.4.17) and (2.4.23), the pull-back is a homomorphism with respect to this algebra structure,

$$\Phi^*(\sigma_1 \wedge \sigma_2) = \big(\Phi^*\sigma_1\big) \wedge \big(\Phi^*\sigma_2\big), \quad \sigma_1, \sigma_2 \in \Gamma\Big(\bigwedge E_b^*\Big). \tag{2.4.24}$$

Remark 2.4.7 (Homomorphism and endomorphism bundles) Analogously, one can construct the homomorphism bundle $\mathrm{Hom}(E_1, E_2)$ of E_1 and E_2, which has the fibres $\mathrm{Hom}(E_{1,m}, E_{2,m})$, and the endomorphism bundle $\mathrm{End}(E)$ of E, which has the fibres $\mathrm{End}(E_m)$. For every m, the vector space $\mathrm{Hom}(E_{1,m}, E_{2,m})$ is naturally isomorphic to the vector space $E^*_{1,m} \otimes E_{2,m}$ and all these isomorphisms combine to a natural isomorphism of $\mathrm{Hom}(E_1, E_2)$ with the tensor product $E_1^* \otimes E_2$ (Exercise 2.4.4). Therefore, we may always identify $\mathrm{Hom}(E_1, E_2)$ with $E_1^* \otimes E_2$. Accordingly, we may identify $\mathrm{End}(E)$ with the tensor bundle $E^* \otimes E \equiv \mathbb{T}^1_1 E$ of E. Then, since vertical C^k-morphisms $E_1 \to E_2$ correspond to C^k-sections of $\mathrm{Hom}(E_1, E_2)$, the vector space of these morphisms is naturally isomorphic to $\Gamma(E_1^* \otimes E_2)$. Accordingly, since vertical endomorphisms of E correspond to sections of $\mathrm{End}(E)$, the vector space of these endomorphisms is naturally isomorphic to $\Gamma(\mathbb{T}^1_1 E)$. The proof is left to the reader (Exercise 2.4.5).

Exercises

2.4.1 Show that the mapping Φ^{T} defined by (2.4.4) is a morphism of vector bundles.

2.4.2 Verify Formulae (2.4.12) and (2.4.21).

2.4.3 Let (E, M, π) be a smooth $\mathbb{K}$-vector bundle. Consider the tangent mapping $\pi' : \mathrm{T}E \to \mathrm{T}M$.
- (a) Show that $(\mathrm{T}E, \mathrm{T}M, \pi')$ is a $\mathbb{K}$-vector bundle by determining the linear structure of the fibres and constructing a system of local trivializations.
- (b) Show that in the cases $E = \mathrm{T}M$ and $E = \mathrm{T}^*M$, local charts on M induce local trivializations of $(\mathrm{T}E, \mathrm{T}M, \pi')$.
- (c) If $E = \mathrm{T}M$, then $(\mathrm{T}E, \mathrm{T}M, \pi')$ has the same base manifold as the tangent bundle of E. Are these two vector bundles isomorphic?

2.4.4 Let E, E_1 and E_2 be $\mathbb{K}$-vector bundles over M of class C^k. Construct the homomorphism bundle $\mathrm{Hom}(E_1, E_2)$ and the endomorphism bundle $\mathrm{End}(E)$ as explained in Remark 2.4.7. Show that $\mathrm{Hom}(E_1, E_2)$ and $\mathrm{End}(E)$ are naturally isomorphic to $E_1^* \otimes E_2$ and $\mathbb{T}^1_1 E$, respectively.

2.4.5 Show that the natural isomorphisms of Exercise 2.4.4 induce natural isomorphisms between the vector space of vertical C^k-morphisms $E_1 \to E_2$ and $\Gamma(E_1^* \otimes E_2)$, as well as between the vector space of vertical C^k-endomorphisms of E and $\Gamma(\mathbb{T}^1_1 E)$, cf. Remark 2.4.7.

2.4.6 The image of the identical mapping id_E under the isomorphism from the vector space of C^k-endomorphisms of E to $\Gamma(\mathbb{T}^1_1 E)$ of Exercise 2.4.5 is called the Kronecker tensor field of E and is denoted by δ. Determine the coefficient functions δ^i_j of δ with respect to the local frame in $\mathbb{T}^1_1 E$ induced by a local frame in E.

2.4.7 Show that if E is one-dimensional, the tensor bundles $\mathbb{T}^2_0 E$, $\mathbb{T}^0_2 E$ and $\mathbb{T}^1_1 E$ are trivial.

2.5 Tensor Bundles and Tensor Fields

Let M be a C^k-manifold of dimension n. By tensor bundles over M one means the various vector bundles arising from the tangent bundle $\mathrm{T}M$ by applying the vector bundle operations of Sect. 2.4. These are

(a) the cotangent bundle $\mathrm{T}^*M := (\mathrm{T}M)^*$. Its fibres are the cotangent spaces[8] $\mathrm{T}^*_m M$ introduced in Sect. 1.4. (Local) sections of T^*M are called (local) covector fields or (local) differential 1-forms.

(b) The bundle of alternating r-vectors $\bigwedge^r \mathrm{T}M$, the bundle of alternating r-forms $\bigwedge^r \mathrm{T}^*M$ and the bundles of exterior algebras

$$\bigwedge \mathrm{T}M = \bigoplus_{r=0}^{n} \bigwedge^{r} \mathrm{T}M, \qquad \bigwedge \mathrm{T}^*M = \bigoplus_{r=0}^{n} \bigwedge^{r} \mathrm{T}^*M.$$

Their (local) sections are called (local) multivector fields and (local) differential forms, respectively. The number r is called the degree. We denote

$$\mathfrak{X}^r(M) := \Gamma\Big(\bigwedge^r \mathrm{T}M\Big), \ \Omega^r(M) := \Gamma\Big(\bigwedge^r \mathrm{T}^*M\Big), \ \Omega^*(M) := \Gamma\Big(\bigwedge \mathrm{T}^*M\Big)$$

and, as before, $\mathfrak{X}(M) \equiv \mathfrak{X}^1(M)$. One has $\mathfrak{X}^0(M) = \Omega^0(M) = C^k(M)$.

(c) The tensor bundles $\mathbb{T}^q_p M := \mathbb{T}^q_p(\mathrm{T}M)$, $p, q = 0, 1, 2, \ldots$. Their (local) sections are called (local) tensor fields of type (p, q). The algebraic operations of symmetrization, antisymmetrization and contraction of tensors over a vector space carry over to tensor fields in an obvious way.

Since $\mathrm{T}M$ is of class C^{k-1}, so are all the tensor bundles over M. Recall from Example 2.4.1 that pointwise evaluation $\mathrm{T}^*_m M \times \mathrm{T}_m M \to \mathbb{R}$ defines a natural pairing

$$\Omega^1(M) \times \mathfrak{X}(M) \to C^{k-1}(M), \qquad (\alpha, X) \mapsto \langle \alpha, X \rangle, \tag{2.5.1}$$

which depending on the context can also be written as $\alpha(X)$ or $X(\alpha)$.

[8] Like for the tangent bundle we will stick to this notation (instead of writing $(\mathrm{T}^*M)_m$).

Example 2.5.1 Let $f : U \to \mathbb{R}$, with $U \subset M$ open, be a real-valued local C^k-function. Then, the differentials $(\mathrm{d}f)_m$ of f at $m \in U$, defined by (1.4.20), combine to a local C^{k-1}-covector field $\mathrm{d}f$ on U, called the differential of f.

Now, let (U, κ) be a local chart on M. The differentials of the coordinate functions κ^i form a local frame $\{\mathrm{d}\kappa^1, \dots, \mathrm{d}\kappa^n\}$ in T^*M which is dual to $\{\partial_1, \dots, \partial_n\}$, cf. Examples 2.3.11/1 and 2.4.1 and Formula (1.4.21). The induced local frame in the tensor bundle $\mathbb{T}^q_p M$ consists of the local sections

$$\mathrm{d}\kappa^{i_1} \otimes \cdots \otimes \mathrm{d}\kappa^{i_p} \otimes \partial_{j_1} \otimes \cdots \otimes \partial_{j_q} : \quad i_1, \dots, i_p, j_1, \dots, j_q = 1, \dots, n,$$

see Example 2.4.4. Using these local frames, a tensor field T of type (p, q) can be represented locally as follows:

$$T_{\restriction U} = \left(T^\kappa\right)^{j_1 \dots j_q}_{i_1 \dots i_p} \mathrm{d}\kappa^{i_1} \otimes \cdots \otimes \mathrm{d}\kappa^{i_p} \otimes \partial_{j_1} \otimes \cdots \otimes \partial_{j_q}, \tag{2.5.2}$$

where, according to (2.4.12), pointwise we have

$$\left(T^\kappa\right)^{j_1 \dots j_q}_{i_1 \dots i_p} = T\left(\partial_{i_1}, \dots, \partial_{i_p}, \mathrm{d}\kappa^{j_1}, \dots, \mathrm{d}\kappa^{j_q}\right). \tag{2.5.3}$$

Remark 2.5.2 We determine the transformation laws for the local frames and for the corresponding coefficient functions of tensor fields under a change of local chart. Thus, let (V, ρ) be another local chart on M. The following formulae hold over $U \cap V$. From (1.4.17) and (1.4.23) we read off

$$\partial^\rho_i = \tilde{A}^j_i \partial^\kappa_j, \qquad \mathrm{d}\rho^i = A^i_j \mathrm{d}\kappa^j$$

where

$$A^i_j := \left[\left(\rho \circ \kappa^{-1}\right)' \circ \kappa\right]^i_j, \qquad \tilde{A}^i_j := \left[\left(\kappa \circ \rho^{-1}\right)' \circ \rho\right]^i_j,$$

and an according formula for the induced local frames in $\mathbb{T}^q_p M$. Then, (2.5.3) implies

$$\left(T^\rho\right)^{j_1 \dots j_q}_{i_1 \dots i_p} = \tilde{A}^{k_1}_{i_1} \cdots \tilde{A}^{k_p}_{i_p} A^{j_1}_{l_1} \cdots A^{j_q}_{l_q} \left(T^\kappa\right)^{l_1 \dots l_q}_{k_1 \dots k_p}. \tag{2.5.4}$$

To pass to coefficient functions which depend on the coordinates, denote the elements of $\kappa(U \cap V)$ by $\mathbf{x}$ and the elements of $\rho(U \cap V)$ by $\mathbf{y}$ and write

$$y^i(\mathbf{x}) := \rho^i \circ \kappa^{-1}(\mathbf{x}), \qquad x^i(\mathbf{y}) := \kappa \circ \rho^{-1}(\mathbf{y}).$$

Then, from (2.5.4) we read off

$$\left(T^\rho\right)^{j_1 \dots j_q}_{i_1 \dots i_p} \circ \rho^{-1} = \frac{\partial x^{k_1}}{\partial y^{i_1}} \cdots \frac{\partial x^{k_p}}{\partial y^{i_p}} \frac{\partial y^{j_1}}{\partial x^{l_1}} \cdots \frac{\partial y^{j_q}}{\partial x^{l_q}} \left(T^\kappa\right)^{l_1 \dots l_q}_{k_1 \dots k_p} \circ \kappa^{-1}. \tag{2.5.5}$$

This formula is well-known from classical tensor analysis. The argument in (2.5.5) can be either $\mathbf{x}$, in which case $\frac{\partial x^{i_a}}{\partial y^{k_a}}$ and $(T^\rho)^{j_1\dots j_q}_{i_1\dots i_p} \circ \rho^{-1}$ have to be evaluated at $\mathbf{y}(\mathbf{x})$, or $\mathbf{y}$, in which case $\frac{\partial y^{l_a}}{\partial x^{j_a}}$ and $(T^\kappa)^{l_1\dots l_q}_{k_1\dots k_p} \circ \kappa^{-1}$ have to be evaluated at $\mathbf{x}(\mathbf{y})$.

Next, let M and N be C^k-manifolds and let $\varphi : M \to N$ be a C^k-mapping. According to Proposition 2.2.9, $\varphi' : \mathrm{T}M \to \mathrm{T}N$ is a vector bundle morphism of class C^{k-1} projecting to φ. The corresponding pull-back operation (2.4.23) applies to differential r-forms of class C^{k-1}. It will be denoted by $\varphi^* : \Omega^r(N) \to \Omega^r(M)$. According to Examples 2.4.1–2.4.5, if φ is a diffeomorphism, φ' induces isomorphisms of tensor bundles. The corresponding transport operator (2.3.3) will be denoted by φ_* in case the induced isomorphism projects to φ and by φ^* in case it projects to φ^{-1}. Then, for $T \in \Gamma(\mathbb{T}^q_p M)$, we have

$$\varphi_* T \equiv \big((\varphi')^{\otimes}\big)_* T = (\varphi')^{\otimes} \circ T \circ \varphi^{-1}, \tag{2.5.6}$$

with $(\varphi')^{\otimes}$ given by (2.4.13), and Formula (2.4.16) takes the form

$$\begin{aligned}(\varphi_* T)&(X_1, \dots, X_p, \alpha_1, \dots, \alpha_q) \\ &= T\big(\varphi_*^{-1} X_1, \dots, \varphi_*^{-1} X_p, \varphi^* \alpha_1, \dots, \varphi^* \alpha_q\big) \circ \varphi^{-1}\end{aligned} \tag{2.5.7}$$

with $X_i \in \mathfrak{X}(M)$ and $\alpha_i \in \Omega^1(M)$. Moreover, Eq. (2.4.15) reads

$$\varphi_*(T_1 \otimes T_2) = (\varphi_* T_1) \otimes (\varphi_* T_2). \tag{2.5.8}$$

Recall from Example 2.4.5 that for differential forms, the transport operation φ^* coincides with the pull-back under φ.

Remark 2.5.3

1. Let $T \in \Gamma(\mathbb{T}^q_p M)$ and let $\varphi : M \to N$ be a diffeomorphism. Given local charts (U, κ) and (V, ρ) on M and N, respectively, the local formula for the transport (2.5.6) of T is given by (2.5.4), with T replaced by $\varphi_* T$ on the left hand side and by $T \circ \varphi^{-1}$ on the right hand side, and with A and $\tilde{A}$ given by

$$A^i_j = \big[\big(\rho \circ \varphi \circ \kappa^{-1}\big)' \circ \kappa \circ \varphi^{-1}\big]^i_k, \qquad \tilde{A}^i_j = \big[\big(\kappa \circ \varphi^{-1} \circ \rho^{-1}\big)' \circ \rho\big]^i_j.$$

   The proof of this fact is left to the reader (Exercise 2.5.3).
2. Let (U, κ) be a local chart on M. We compare the corresponding local representative $\kappa_*(T_{\restriction U})$ of a tensor field $T \in \Gamma(\mathbb{T}^q_p M)$ with the local representative of the mapping $T : M \to \mathbb{T}^q_p M$ with respect to the induced chart $((\pi^{\otimes})^{-1}(U), \kappa^{\otimes})$ on $\mathbb{T}^q_p M$, given by

$$\kappa^{\otimes} \circ T_{\restriction U} \circ \kappa^{-1} : \kappa(U) \to \kappa(U) \times \mathbb{T}^q_p \mathbb{R}^n.$$

Proofs are left to the reader (Exercise 2.5.4). Since

$$\kappa_* \partial_i = \frac{\partial}{\partial x^i}, \qquad \kappa^* \mathrm{d}x^i = \mathrm{d}\kappa^i, \tag{2.5.9}$$

we have

$$\kappa_*(T_{\restriction U}) = \left(\left(T^\kappa\right)^{j_1 \ldots j_q}_{i_1 \ldots i_p} \circ \kappa^{-1}\right) \mathrm{d}x^{i_1} \otimes \cdots \otimes \frac{\partial}{\partial x^{j_q}}. \tag{2.5.10}$$

On the other hand,

$$\left(\kappa^\otimes \circ T_{\restriction U} \circ \kappa^{-1}\right)(\mathbf{x}) = \left(\mathbf{x}, \left(\left(T^\kappa\right)^{j_1 \ldots j_q}_{i_1 \ldots i_p} \circ \kappa^{-1}(\mathbf{x})\right) \mathbf{e}^{*i_1} \otimes \cdots \otimes \mathbf{e}_{j_q}\right),$$

where, as before, $\mathbf{e}_i$ denote the elements of the standard basis of $\mathbb{R}^n$ and $\mathbf{e}^{*i}$ the elements of the dual basis. The relation to $\kappa_*(T_{\restriction U})$ is as follows. The natural identifications of the tangent spaces $\mathrm{T}_{\mathbf{x}}(\kappa(U))$ with $\mathbb{R}^n$ and of the cotangent spaces $\mathrm{T}^*_{\mathbf{x}}(\kappa(U))$ with $\mathbb{R}^{n*}$ induce a natural identification of tensor fields on $\kappa(U)$ with C^{k-1}-mappings $\kappa(U) \to \mathbb{T}^q_p \mathbb{R}^n$. Since the latter identifies the elements of the global frames $\{\frac{\partial}{\partial x^i}\}$ in $\mathrm{T}(\kappa(U))$ and $\{\mathrm{d}x^i\}$ in $\mathrm{T}^*(\kappa(U))$ with the constant mappings $\mathbf{x} \mapsto \mathbf{e}_i$ and $\mathbf{x} \mapsto \mathbf{e}^{*i}$, respectively, it identifies $\kappa_*(T_{\restriction U})$ with $\kappa^\otimes \circ T_{\restriction U} \circ \kappa^{-1}$. Note that for $M = \mathbb{R}^n$ and $\kappa = \mathrm{id}$, (2.5.9) yields $\partial_i = \frac{\partial}{\partial x^i}$.

Exercises

2.5.1 Let $M_1 = \mathbb{R}_+ \times \mathrm{S}^1$, with S^1 realized as the unit sphere in $\mathbb{R}^2$, and $M_2 = \mathbb{R}^2 \setminus \{0\}$. Consider the mapping $\varphi : M_1 \to M_2$, $\varphi(r, (a, b)) := (ra, rb)$. Let r denote the standard coordinate on $\mathbb{R}_+$ and let ϕ denote the angle coordinate of S^1. Determine the coefficient functions of $\varphi_* \frac{\partial}{\partial r}$ and $\varphi_* \frac{\partial}{\partial \phi}$ with respect to the global frame $\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\}$ in $\mathrm{T}(\mathbb{R}^2 \setminus \{0\})$.

2.5.2 Let $M_1 = \mathbb{R}_+ \times \mathrm{S}^2$, with S^2 realized as the unit sphere in $\mathbb{R}^3$, and $M_2 = \mathbb{R}^3 \setminus \{0\}$. Consider the mapping $\varphi : M_1 \to M_2$, $\varphi(r, (a, b, c)) = (ra, rb, rc)$. Let r denote the natural coordinate on $\mathbb{R}_+$ and let the angle coordinates ϑ, ϕ on S^2 be defined by $a = \cos\phi \sin\vartheta$, $b = \sin\phi \sin\vartheta$, $c = \cos\vartheta$. Determine the coefficient functions of $\varphi_* \frac{\partial}{\partial r}$, $\varphi_* \frac{\partial}{\partial \phi}$ and $\varphi_* \frac{\partial}{\partial \theta}$ with respect to the global frame $\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\}$ in $\mathrm{T}(\mathbb{R}^3 \setminus \{0\})$.

2.5.3 Prove the transformation formula for the transport of tensor fields under diffeomorphisms given in Remark 2.5.3/1.

2.5.4 Prove the assertions of Remark 2.5.3/2.

2.6 Induced Bundles

Let (E, N, π) be a $\mathbb{K}$-vector bundle of class C^k, let M be a C^k-manifold and let $\varphi \in C^k(M, N)$. Using φ, one can construct from E a vector bundle $\varphi^* E$ over M by attaching to $m \in M$ the fibre $E_{\varphi(m)}$ as follows. Define

$$\varphi^* E := \{(m, x) \in M \times E : \varphi(m) = \pi(x)\}$$

and consider the surjective mapping $\pi^{\varphi^*} : \varphi^* E \to M$ defined by $\pi^{\varphi^*}(m, x) := m$. The fibres are

$$(\varphi^* E)_m \equiv (\pi^{\varphi^*})^{-1}(m) = \{m\} \times E_{\varphi(m)}.$$

They inherit a natural $\mathbb{K}$-vector space structure from E.

Proposition 2.6.1 *Under the above assumptions, $\varphi^* E$ admits a C^k-structure such that it is an embedded submanifold of $M \times E$. Then,*

1. *$(\varphi^* E, M, \pi^{\varphi^*})$ is a $\mathbb{K}$-vector bundle of class C^k,*
2. *the natural projection $M \times E \to E$ restricts to a C^k-morphism $\varphi^* E \to E$ covering φ,*
3. *every local section s of E induces a local section of $\varphi^* E$ defined by*

$$(\varphi^* s)(m) := (m, s \circ \varphi(m)).$$

Proof We apply Proposition 1.7.3 in the formulation of Remark 1.7.4. Choose a typical fibre F and a system of local trivializations $\{(U_\alpha, \chi_\alpha) : \alpha \in A\}$ for E. For every $\alpha \in A$, consider the open subset $V_\alpha := \varphi^{-1}(U_\alpha)$ of M and the mapping

$$\psi_\alpha : V_\alpha \times F \to M \times E, \qquad \psi_\alpha(m, u) := (m, \chi_\alpha^{-1}(\varphi(m), u)).$$

Since ψ_α is obtained by composing the diffeomorphism χ_α^{-1} with the natural inclusion mapping of the graph of $\varphi_{\restriction V_\alpha} : V_\alpha \to U_\alpha$, by Example 1.6.12/2, it is a C^k-embedding. Hence, the image $\psi_\alpha(V_\alpha \times F)$ inherits a C^k-structure from $V_\alpha \times F$ and with respect to this structure it is an embedded C^k-submanifold of $M \times E$. Since the image is $\varphi^* E \cap (V_\alpha \times \pi^{-1}(U_\alpha))$ and since the $V_\alpha \times \pi^{-1}(U_\alpha)$ are open subsets of $M \times E$ covering $\varphi^* E$, we conclude that $\varphi^* E$ is an embedded submanifold. It remains to prove assertion 1; assertions 2 and 3 are then obvious. Since π^{φ^*} is the restriction of the natural projection $M \times E \to M$ to the C^k-submanifold $\varphi^* E$, it is of class C^k. Since, by construction, the ψ_α restrict to C^k-diffeomorphisms from $V_\alpha \times F$ to $\varphi^* E \cap (V_\alpha \times \pi^{-1}(U_\alpha)) = (\pi^{\varphi^*})^{-1}(V_\alpha)$, by inverting them we obtain C^k-diffeomorphisms

$$\chi_\alpha^{\varphi^*} : (\pi^{\varphi^*})^{-1}(V_\alpha) \to V_\alpha \times F, \qquad \chi_\alpha^{\varphi^*}(m, x) = (m, \chi_{\alpha,\varphi(m)}(x)). \tag{2.6.1}$$

The latter satisfy conditions 2a and 2b of Definition 2.2.1. Thus, $(\varphi^* E, M, \pi^{\varphi^*})$ is a $\mathbb{K}$-vector bundle of class C^k. □

Definition 2.6.2 (Induced vector bundle) The $\mathbb{K}$-vector bundle $(\varphi^* E, M, \pi^{\varphi^*})$ is called the vector bundle induced from E by φ or the pull-back of E by φ. For a local section s of E, the local section $\varphi^* s$ of $\varphi^* E$ is said to be induced from s by φ or to be the pull-back of s by φ.

Another common notation for the induced vector bundle is $\varphi^* E \equiv M \times_N E$.

Remark 2.6.3

1. From the proof of Proposition 2.6.1 we note that via (2.6.1), every local trivialization (U, χ) of E induces a local trivialization $(\varphi^{-1}(U), \chi^{\varphi^*})$ of $\varphi^* E$. In particular, the pull-back of a trivial vector bundle is trivial. Moreover, if $\rho_{\alpha\beta} : U_\alpha \cap U_\beta \to \mathrm{GL}(F)$ are the transition functions of a system of local trivializations of E, then $\varphi^* \rho_{\alpha\beta} : \varphi^{-1}(U_\alpha) \cap \varphi^{-1}(U_\beta) \to \mathrm{GL}(F)$ are the transition functions of the induced system of local trivializations of $\varphi^* E$.
2. Let (E_i, M_i, π_i), $i = 1, 2$, be $\mathbb{K}$-vector bundles of class C^k and let $\Phi : E_1 \to E_2$ be a morphism with projection $\varphi : M_1 \to M_2$. Φ naturally decomposes as

$$E_1 \xrightarrow{\Phi_{\mathrm{ver}}} \varphi^* E_2 \xrightarrow{\Phi_{\mathrm{hor}}} E_2, \tag{2.6.2}$$

where Φ_{ver} is given by $\Phi_{\mathrm{ver}}(x) = (\pi_1(x), \Phi(x))$, $x \in E_1$, and Φ_{hor} denotes the induced vector bundle morphism of Proposition 2.6.1/2. One can check that Φ_{ver} is a vertical morphism, with differentiability of class C^k following from Proposition 1.6.10 and the fact that $\varphi^* E$ is an embedded submanifold of $M \times E$. Using this decomposition, one can derive the following characterization of isomorphisms in terms of their projections and fibre mappings (Exercise 2.6.1): a morphism is an isomorphism iff its projection is a diffeomorphism and its fibre mappings are bijective.

Example 2.6.4

1. If $\varphi : M \to N$ is constant with $\varphi(m) = p$, then $\varphi^* E$ coincides with the product vector bundle $M \times E_p$.
2. If $M \subset N$ is an open subset and $j : M \to N$ is the natural inclusion mapping, $j^* E$ can be identified with the restriction $E_{\restriction M}$, see Example 2.2.3/3.
3. Let $M = N = \mathrm{S}^1$ and let E be the Möbius strip of Example 2.2.4. Realize S^1 as the unit circle in $\mathbb{C}$ and consider the n-fold covering $\varphi_n : \mathrm{S}^1 \to \mathrm{S}^1$, $\varphi_n(z) = z^n$. Since $\varphi_n^* E$ is a differentiable real vector bundle over S^1 of dimension 1, according to Remark 2.3.14, it must be isomorphic to either E or the product vector bundle $\mathrm{S}^1 \times \mathbb{R}$. Indeed, one finds (Exercise 2.6.2)

$$\varphi_n^* E \cong \begin{cases} E \mid n \text{ odd}, \\ \mathrm{S}^1 \times \mathbb{R} \mid n \text{ even}. \end{cases}$$

4. Let E_1 and E_2 be $\mathbb{K}$-vector bundles of class C^k over M, let $E_1 \times E_2$ denote the product vector bundle over $M \times M$, see Exercise 2.2.7, and let $\Delta : M \to M \times M$ denote the diagonal mapping, $\Delta(m) = (m, m)$. The pull-back $\Delta^*(E_1 \times E_2)$ is naturally isomorphic to the direct sum $E_1 \oplus E_2$ (Exercise 2.6.3).
5. If E is a $\mathbb{K}$-vector bundle of class C^k over N and (M, φ) is a C^k-submanifold of N, the induced vector bundle $\varphi^* E$ is referred to as the restriction of E to M and is usually denoted by $E_{\restriction M}$. This applies in particular to $E = \mathrm{T}N$, where $\varphi^* \mathrm{T}N$ is a real vector bundle over M of class C^{k-1} and dimension $\dim N$.

Exercises

2.6.1 Use the natural decomposition (2.6.2) of vector bundle morphisms to show that a morphism is an isomorphism iff its projection is a diffeomorphism and the fibre mappings are bijective, cf. Remark 2.6.3/2.

2.6.2 Prove the statement of Example 2.6.4/3 about the pull-back of the Möbius strip by means of a covering of S^1.

2.6.3 Show that $\Delta^*(E_1 \times E_2) \cong E_1 \oplus E_2$, see Example 2.6.4/4.

2.7 Subbundles and Quotient Bundles

Definition 2.7.1 (Vector subbundle) Let (E_i, M_i, π_i), $i = 1, 2$, be $\mathbb{K}$-vector bundles of class C^k and let $\Phi : E_1 \to E_2$ be a morphism. The pair (E_1, Φ) is called a subbundle, an initial subbundle or an embedded subbundle of E_2 if it is, respectively, a submanifold, an initial submanifold or an embedded submanifold. If $M_1 = M_2 = M$ and Φ is vertical, (E_1, Φ) is called a vertical subbundle or a subbundle over M.

At the very beginning, we observe that Propositions 1.6.10 and 1.6.14 remain true if the term submanifold is replaced by subbundle and C^k-mapping by morphism. The following two specific types of subbundles are the building blocks for arbitrary subbundles.

Example 2.7.2 (Vertical subbundle) If E_1 and E_2 are $\mathbb{K}$-vector bundles of class C^k over M and $\Phi : E_1 \to E_2$ is an injective vertical morphism, then (E_1, Φ) is a vertical subbundle of E_2. Vertical subbundles are embedded. To see this, it suffices to show that (E_1, Φ) is an embedded submanifold of E_2. Let l_i denote the dimensions of E_i. Necessarily, $l_1 \leq l_2$. Let $x \in E_1$ and $m := \pi_1(x)$. Choose a local frame in E_1 at m. By injectivity, the image under Φ is a local l_1-frame in E_2. According to Proposition 2.3.15/3, the latter can be complemented, over a possibly smaller domain U, to a local frame in E_2 at m. The local representative of Φ with respect to the local trivializations associated with these local frames in E_1 and E_2 is given by

$$U \times \mathbb{K}^{l_1} \to U \times \mathbb{K}^{l_2}, \qquad (m, \mathbf{x}) \mapsto \big(m, (\mathbf{x}, 0)\big).$$

Hence, it is an embedding. Since $\pi_1^{-1}(U) = \Phi^{-1}(\pi_2^{-1}(U))$, this implies that the restriction $\Phi_{\restriction \Phi^{-1}(\pi_2^{-1}(U))}$ is an embedding. Since $\pi_2^{-1}(U)$ is an open neighbourhood of $\Phi(x)$ and x was arbitrary, Remark 1.6.13/3 yields the assertion.

Example 2.7.3 (Restriction of the base manifold) Let (E, N, π) be a $\mathbb{K}$-vector bundle of class C^k and let (M, φ) be a C^k-submanifold of N. Let $\Phi : \varphi^* E \to E$ denote the induced vector bundle morphism of Proposition 2.6.1/2. Recall from Example 2.6.4/5 that $\varphi^* E$ is referred to as the restriction of E to M and is alternatively denoted by $E_{\restriction M}$. We show that $(\varphi^* E, \Phi)$ is a C^k-subbundle of E. If (M, φ) is initial or embedded, so is $(\varphi^* E, \Phi)$. Indeed, the local representatives of Φ with respect to a system of local trivializations $\{(U_\alpha, \chi_\alpha) : \alpha \in A\}$ of E and the induced system of local trivializations of $\varphi^* E$ are given by

$$\chi_\alpha \circ \Phi_{\restriction (\pi^{\varphi^*})^{-1}(\varphi^{-1}(U_\alpha))} \circ \left(\chi_\alpha^{\varphi^*}\right)^{-1} : \varphi^{-1}(U_\alpha) \times F \to U_\alpha \times F,$$
$$(m, u) \mapsto \big(\varphi(m), u\big).$$

First, this implies that Φ is an immersion. Second, since χ_α and $\chi_\alpha^{\varphi^*}$ are diffeomorphisms and since

$$\left(\pi^{\varphi^*}\right)^{-1}\left(\varphi^{-1}(U_\alpha)\right) = \Phi^{-1}\left(\pi^{-1}(U_\alpha)\right),$$

this implies that the submanifolds $(\Phi^{-1}(\pi^{-1}(U_\alpha)), \Phi_{\restriction \Phi^{-1}(\pi^{-1}(U_\alpha))})$ inherit the property of being initial or embedded from (M, φ). Then, Remark 1.6.13/3 yields the assertion.

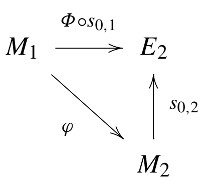
The following proposition states criteria for a morphism to define a subbundle.

Proposition 2.7.4 *Let (E_i, M_i, π_i), $i = 1, 2$, be $\mathbb{K}$-vector bundles of class C^k, let $\Phi : E_1 \to E_2$ be a morphism and let $\varphi : M_1 \to M_2$ be the projection. The following statements are equivalent.*

1. *(E_1, Φ) is, respectively, a subbundle, initial subbundle or embedded subbundle of E_2.*
2. *(M_1, φ) is, respectively, a submanifold, initial submanifold or embedded submanifold of M_2 and the fibre mappings $\Phi_m : E_{1,m} \to E_{2,\varphi(m)}$ are injective for all $m \in M_1$.*
3. *In the decomposition (2.6.2), $(E_1, \Phi_{\mathrm{ver}})$ is a vertical subbundle of $\varphi^* E_2$ and $(\varphi^* E_2, \Phi_{\mathrm{hor}})$ is, respectively, a subbundle, initial subbundle or embedded subbundle of E_2.*

Item 3 gives a precise meaning to the statement made above that vertical subbundles (Example 2.7.2) and restrictions of the base manifold (Example 2.7.3) provide the building blocks for arbitrary subbundles.

Proof $1 \Rightarrow 2$: The fibre mappings Φ_m are obviously injective. Since they are linear, one has the commutative diagram



where $s_{0,i}$ denotes the zero sections of E_i. According to Proposition 2.3.5, $s_{0,1}$ and $s_{0,2}$ are embeddings. Hence, the assertion follows by applying Proposition 1.6.14. (We encourage the reader to work out the argument for each case.)

2 $\Rightarrow$ 3: Since the mappings Φ_m are injective, Φ_{ver} is injective, hence the assertion on (E_1, Φ_{ver}) holds due to Example 2.7.2. The assertion on $(\varphi^* E_2, \Phi_{\text{hor}})$ was proved in Example 2.7.3.

3 $\Rightarrow$ 1: Since vertical subbundles are embedded, this follows from Proposition 1.6.14/1. □

In the following proposition we give criteria for a family of fibre subspaces of a vector bundle to define a vertical subbundle. The proof is left to the reader (Exercise 2.7.1).

Proposition 2.7.5 (Families of fibre subspaces) *Let (E_2, M, π) be a $\mathbb{K}$-vector bundle of class C^k. For every $m \in M$, let $E_{1,m} \subset E_{2,m}$ be a linear subspace. Define $E_1 := \bigcup_{m \in M} E_{1,m}$. The following statements are equivalent.*

1. *E_1 admits a C^k-structure such that it is a vertical subbundle of E_2 of dimension r.*
2. *There exists a covering of M by local r-frames in E_2 which span E_1.*
3. *There exists a covering of M by local frames in E_2 whose first r elements span E_1.*
4. *There exists a system of local trivializations $\{(U_\alpha, \chi_\alpha) : \alpha \in A\}$ of E_2 and a subspace F_1 of dimension r of the typical fibre F_2 such that the restrictions of the χ_α to E_1 take values in $U_\alpha \times F_1$.*

Example 2.7.6 (Regular distribution) Let M be a C^k-manifold. A vertical subbundle (D, Φ) of $\mathrm{T}M$ is called a regular distribution (in the geometrical sense) on M. According to Proposition 2.7.5, a family of r-dimensional subspaces $D_m \subset \mathrm{T}_m M$, $m \in M$, defines a distribution iff for every $m_0 \in M$ there exists an open neighbourhood U and pointwise linearly independent local vector fields $X_1, \dots, X_r$ on U such that D_m is spanned by $X_{1,m}, \dots, X_{r,m}$ for all $m \in U$. There is a more general notion of distribution on M which will be defined and studied in Sect. 3.5.

Example 2.7.7 (Kernel and image) Let E_i be $\mathbb{K}$-vector bundles over M of class C^k and dimension l_i, $i = 1, 2$, and let $\Phi : E_1 \to E_2$ be a vertical morphism of constant rank r. Define the image and the kernel of Φ to be

$$\operatorname{im} \Phi := \bigcup_{m \in M} \operatorname{im} \Phi_m, \qquad \ker \Phi := \bigcup_{m \in M} \ker \Phi_m,$$

respectively. We show that $\operatorname{im} \Phi$ is a vertical subbundle of E_2 of dimension r and that $\ker \Phi$ is a vertical subbundle of E_1 of dimension $l_1 - r$.

Let $m_0 \in M$. Choose a basis $\{e_1, \dots, e_{l_1}\}$ of E_{1,m_0} such that $e_{r+1}, \dots, e_{l_1}$ span $\ker \Phi_{m_0}$. Extend this basis to a local frame $\{s_1, \dots, s_{l_1}\}$ in E_1, cf. Proposition 2.3.15. By construction, the vectors $\Phi_{m_0}(e_1), \dots, \Phi_{m_0}(e_r)$ form a basis of the subspace

$\operatorname{im}\Phi_{m_0} \subset E_{2,m_0}$. In particular, the local sections $\Phi \circ s_1, \ldots, \Phi \circ s_r$ of E_2 are linearly independent at m_0 so that, by possibly shrinking the domain of definition of the s_i, we may assume that they form a local r-frame in E_2. Since Φ has rank r, this local r-frame spans $\operatorname{im}\Phi_m$ for all m belonging to the domain of definition. First, in view of Proposition 2.7.5, this yields the assertion for $\operatorname{im}\Phi$. Second, this implies that there exist local C^k-functions a_{ij}, $i = r+1, \ldots, l_1$, $j = 1, \ldots, r$, on M such that

$$\Phi \circ s_i = \sum_{j=1}^{r} a_{ij} \Phi \circ s_j, \quad r+1 \leq i \leq l_1.$$

Then the local sections $\tilde{s}_{r+1}, \ldots, \tilde{s}_{l_1}$ given by

$$\tilde{s}_i := s_i - \sum_{j=1}^{r} a_{ij} s_j, \quad r+1 \leq i \leq l_1,$$

form a local $(l_1 - r)$-frame in E_1 spanning $\ker\Phi_m$. Applying Proposition 2.7.5 once again, we obtain the assertion for $\ker\Phi$.

Example 2.7.8 (Annihilator) Let V be a vector space. The annihilator of a subspace $W \subset V$ is the subspace

$$W^0 := \left\{\nu \in V^* : \nu_{\restriction W} = 0\right\}$$

of the dual vector space V^*. Let E_2 be a $\mathbb{K}$-vector bundle over M of class C^k and dimension l_2 and let (E_1, Φ) be a vertical subbundle of dimension l_1. Then,

$$E_1^0 := \bigcup_{m \in M} \left(\Phi(E_{1,m})\right)^0$$

is a vertical subbundle of dimension $l_2 - l_1$ of the dual vector bundle E_2^*, called the annihilator of E_1 in E_2. In view of Proposition 2.7.5/3, this follows from the obvious fact that for every local frame in E_2 whose first l_1 elements span (E_1, Φ), the last $l_2 - l_1$ elements of the corresponding dual local frame in E_2^* span E_1^0. The annihilator of a general vector subbundle (E_1, Φ) is defined to be $(E_1^0, (\Phi_{\mathrm{hor}})_{\restriction E_1^0})$, where E_1^0 is the annihilator of the vertical subbundle $(E_1, \Phi_{\mathrm{ver}})$ of $\varphi^* E_2$ and φ is the projection of Φ. It has the same base manifold as E_1.

Remark 2.7.9

1. For every vertical subbundle (E_1, Φ) of E_2 there exists a complement in E_2, that is, a vertical subbundle $(\tilde{E}_1, \tilde{\Phi})$ of E_2 such that $E_2 = E_1 \oplus \tilde{E}_1$. The proof is in two steps.
   (a) Show that for every $\mathbb{K}$-vector bundle (E, M, π) of class C^k there exists a C^k-function $\mathsf{h} : E \otimes E \to \mathbb{K}$ such that $\mathsf{h}_m := \mathsf{h}_{\restriction E_m \otimes E_m}$ is a scalar product on E_m for all $m \in M$ (Exercise 2.7.2).[9]

[9] (E, h) is called a Euclidean vector bundle if $\mathbb{K} = \mathbb{R}$ and a Hermitian vector bundle if $\mathbb{K} = \mathbb{C}$.

(b) Show that the family of h_m-orthogonal complements of the subspaces $E_{1,m} \subset E_{2,m}$ defines a vertical subbundle of E_2 (Exercise 2.7.3).

2. Let M be a compact smooth manifold. The statement of 1 provides part of the proof that for every smooth vector bundle E over M there exists a smooth vector bundle $\tilde{E}$ over M such that $E \oplus \tilde{E}$ is trivial. For the remaining part, see for example [125, Prop. 1.4].[10] This is known as the cancellation property and is an important ingredient in what is called the K-theory of M. Let us have a glimpse at the reduced version of the latter. Two smooth $\mathbb{K}$-vector bundles E and $\tilde{E}$ over M are said to be stably equivalent if $E \oplus (M \times \mathbb{K}^r)$ is isomorphic to $\tilde{E} \oplus (M \times \mathbb{K}^s)$ for some r, s. The set of stable equivalence classes is an Abelian semigroup with respect to the operation of direct sum, where the unit element is given by the class of trivial bundles. Now, the cancellation property yields that every element of this semigroup has an inverse, hence the semigroup is in fact a group, called the reduced real (for $\mathbb{K} = \mathbb{R}$) or complex (for $\mathbb{K} = \mathbb{C}$) K-group of M. Together with the operation of tensor product, it is an Abelian ring.

Next, we discuss quotient vector bundles. Let (E_2, M, π_2) be a $\mathbb{K}$-vector bundle of class C^k and let (E_1, Φ) be a vertical subbundle of E_2 of rank r with projection π_1. Since vertical subbundles are embedded, we may assume that $E_1 \subset E_2$ and Φ is the natural inclusion mapping. $E_{1,m}$ is a vector subspace of $E_{2,m}$ for all $m \in M$, and we can form the quotient spaces $E_{2,m}/E_{1,m}$. Let

$$E_2/E_1 := \bigsqcup_{m \in M} E_{2,m}/E_{1,m}$$

and let $\pi : E_2/E_1 \to M$ denote the natural projection to the index set. By construction, the fibres $\pi^{-1}(m)$ are vector spaces. According to Proposition 2.7.5, there exists a family of local trivializations $\{(U_\alpha, \chi_{2\alpha}) : \alpha \in A\}$ of E_2 and an r-dimensional subspace F_1 of the typical fibre F_2 of E_2 such that the restrictions of $\chi_{2\alpha}$ to E_1 take values in $U_\alpha \times F_1$. For any such $\chi_{2\alpha}$, we define a mapping

$$\chi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F_2/F_1, \qquad \chi_\alpha\big([x]\big) := \big(m, \big[\chi_{2\alpha,m}(x)\big]\big),$$

where $m = \pi_2(x)$. To check differentiability of the corresponding transition mappings, choose a complement $\tilde{F}_1$ of F_1 in F_2 and let $\lambda : F_2/F_1 \to F_2$ denote the linear mapping which assigns to each class its unique representative in $\tilde{F}_1$. Moreover, let $\mathrm{pr} : F_2 \to F_2/F_1$ be the natural projection. Since $\chi_\beta \circ \chi_\alpha^{-1}$ decomposes as

$$\chi_\beta \circ \chi_\alpha^{-1} = (\mathrm{id} \times \mathrm{pr}) \circ \big(\chi_{2\beta} \circ \chi_{2\alpha}^{-1}\big) \circ (\mathrm{id} \times \lambda),$$

it is of class C^k. Then, Remark 2.2.5 yields that the family $\{(U_\alpha, \chi_\alpha) : \alpha \in A\}$ defines a C^k-structure on E_2/E_1 such that $(E_2/E_1, M, \pi)$ is a $\mathbb{K}$-vector bundle of class C^k over M. This C^k-structure obviously does not depend on the choice of the subspace F_1.

[10] Compactness of M is necessary here, see Example 3.6 in [125].

Definition 2.7.10 (Quotient vector bundle) The vector bundle $(E_2/E_1, M, \pi)$ constructed above is called the quotient vector bundle of E_2 by E_1.

Remark 2.7.11

1. The fibrewise natural projections $E_{2,m} \to E_{2,m}/E_{1,m}$ to classes combine to a natural projection $E_2 \to E_2/E_1$. The latter is a vertical morphism, because its local representative with respect to a local trivialization $(U_\alpha, \chi_{2\alpha})$ of E_2 whose restriction to E_1 takes values in $U_\alpha \times F_1$, and the induced local trivialization of E_2/E_1 is given by the natural projection $F_2 \to F_2/F_1$. By composing a local section s of E_2 with the natural projection $E_2 \to E_2/E_1$ one obtains a local section of E_2/E_1, denoted by $[s]$.
2. Let l_i denote the dimension of E_i, $i = 1, 2$. For any local frame $\{s_1, \dots, s_{l_2}\}$ in E_2 with the property[11] that $s_1, \dots, s_{l_1}$ span E_1, $\{[s_{l_1+1}], \dots, [s_{l_2}]\}$ is a local frame in E_2/E_1.
3. According to Remark 2.7.9/1, E_1 admits a complement $\tilde{E}_1$ in E_2. For any such complement, the natural projection $E_2 \to E_2/E_1$ restricts to a vertical isomorphism $\tilde{E}_1 \to E_2/E_1$. This follows at once by observing that the induced mapping is a bijective vertical morphism. Thus, every complement defines a vector bundle isomorphism
$$E_2 \cong E_1 \oplus (E_2/E_1).$$
4. By a coorientation, or transversal orientation, of E_1 in E_2 one means an orientation of the quotient vector bundle E_2/E_1. Accordingly, E_1 is said to be coorientable, or transversally orientable, in E_2 if E_2/E_1 is orientable.

Example 2.7.12 (Homomorphism theorem) Let E_1 and E_2 be $\mathbb{K}$-vector bundles of class C^k over M and let $\Phi\colon E_1 \to E_2$ be a vertical morphism of constant rank. Then, the induced mapping

$$\tilde{\Phi} : E_1/\ker\Phi \to \operatorname{im}\Phi$$

is an isomorphism. Indeed, $\tilde{\Phi}$ is obviously bijective and fibrewise linear. To see that it is of class C^k, one may choose a complement E_0 of $\ker\Phi$ in E_1 and write $\tilde{\Phi}$ as the composition of the isomorphism $E_1/\ker\Phi \to E_0$ and the restriction of $\tilde{\Phi}$ in domain to E_0. Thus, $\tilde{\Phi}$ is a bijective vertical morphism and hence an isomorphism.

Example 2.7.13 (Dual quotient vector bundle) Let E_2 be a $\mathbb{K}$-vector bundle of class C^k over M and let E_1 be a vertical subbundle. The dual vector bundle $(E_2/E_1)^*$ is called the dual quotient vector bundle. It is naturally isomorphic over M to the annihilator E_1^0. Indeed, the mapping

$$\Phi : E_1^0 \to (E_2/E_1)^*, \qquad \big(\Phi_m(\xi)\big)\big([x]\big) := \xi(x), \quad \xi \in E_{2,m}^*,\ x \in E_{2,m},$$

[11] Such local frames exist by Proposition 2.7.5.

is well-defined, bijective and fibrewise linear. Hence, it remains to show that Φ is of class C^k. To see this, choose a local frame $\{s_1, \ldots, s_{l_2}\}$, whose first l_1 elements span E_1 over U. Then, the elements s^{*i}, $i = l_1 + 1, \ldots, l_2$, of the dual local frame span E_1^0 over U and, according to Remark 2.7.11/2, the dual local frame $\{[s_{l_1+1}]^*, \ldots, [s_{l_2}]^*\}$ spans $(E_2/E_1)^*$ over U. By construction, the local representative of Φ with respect to the local trivializations defined by these local frames does not depend on m and is hence of class C^k, as asserted.

To conclude this section, we discuss vector bundle structures induced by submanifolds. Thus, let N be a C^k-manifold and let (M, φ) be a C^k-submanifold of N.

Proposition 2.7.14 $(\mathrm{T}M, \varphi')$ *is a vector subbundle of* $\mathrm{T}N$. *It is initial or embedded iff so is* (M, φ).

Proof Recall from Example 2.6.4/5 that the restriction of $\mathrm{T}N$ to the submanifold (M, φ) is defined to be the induced vector bundle $(\mathrm{T}N)_{\restriction M} := \varphi^*\mathrm{T}N$. Since φ is an immersion, the vertical morphism $(\varphi')_{\mathrm{ver}} : \mathrm{T}M \to (\mathrm{T}N)_{\restriction M}$ in the natural decomposition (2.6.2) of φ' is injective. Hence, $(\mathrm{T}M, (\varphi')_{\mathrm{ver}})$ is a vertical subbundle of $(\mathrm{T}N)_{\restriction M}$. Then, Proposition 2.7.4/3 yields that $(\mathrm{T}M, \varphi')$ is a subbundle of $\mathrm{T}N$ and that it is initial or embedded if so is (M, φ). The converse direction follows from Proposition 1.6.14 and the fact that the zero sections of the tangent bundles of M and N are embeddings. The details are left to the reader (Exercise 2.7.4). □

Remark 2.7.15 Let V be a finite-dimensional real vector space and let $M \subset V$ be an embedded C^k-submanifold. For every $v \in M$, the natural identification of $\mathrm{T}_v V$ with V of Example 1.4.3/2 identifies $\mathrm{T}_v M$ with a subspace of V, which we denote by the same symbol. In particular, in case M is open in V, one has $\mathrm{T}_v M = V$; and in case M is a level set of a C^k-mapping f, one has $\mathrm{T}_v M = \ker f'(v)$. In the general case, $\mathrm{T}_v M$ is just the tangent plane of M at v, shifted by $-v$ to the origin. Thus, together with the induced natural identification of $\mathrm{T}V$ with $V \times V$, Proposition 2.7.14 yields a natural identification of $\mathrm{T}M$ with the embedded C^{k-1}-submanifold

$$\big\{(v, X) \in M \times V : X \in \mathrm{T}_v M\big\}$$

of $M \times V$ and a natural representation of vector fields on M by C^{k-1}-mappings $X : M \to V$ satisfying $X(v) \in \mathrm{T}_v M$ for all $v \in M$. This generalizes Remarks 2.1.4/2 and 2.3.4/2.

A further consequence of the observation that $(\mathrm{T}M, (\varphi')_{\mathrm{ver}})$ is a vertical subbundle of $(\mathrm{T}N)_{\restriction M}$ is the following. A vector field X on N is said to be tangent to the submanifold (M, φ) if $X_{\varphi(m)} \in \varphi'(\mathrm{T}_m M)$ for all $m \in M$.

Proposition 2.7.16 *Let N be a manifold and let (M, φ) be a submanifold of N. For every vector field X on N which is tangent to (M, φ), there exists a unique vector field $\tilde{X}$ on M such that $\varphi' \circ \tilde{X} = X \circ \varphi$, that is, $\tilde{X}$ and X are φ-related.*

We will say that $\tilde{X}$ is induced from X and call it the restriction of X to (M, φ).

Proof Due to the assumption, the equation $\varphi' \circ \tilde{X} = X \circ \varphi$ defines a mapping $\tilde{X} : M \to \mathrm{T}M$. $\tilde{X}$ is the restriction in range to $\mathrm{T}M$ of the section of $(\mathrm{T}N)_{\restriction M} = \varphi^*\mathrm{T}N$ induced from X by φ. Since vertical subbundles are embedded, Proposition 1.6.10 yields that $\tilde{X}$ is differentiable,[12] that is, of class C^{k-1}. □

Finally, we introduce

Definition 2.7.17 (Normal and conormal bundle) Let N be a manifold and let (M, φ) be a submanifold of N.

1. The quotient vector bundle $\mathrm{N}M := (\mathrm{T}N)_{\restriction M}/\mathrm{T}M$ is called the normal bundle of (M, φ). Its fibres are called the normal spaces of M at $m \in M$. They are denoted by $\mathrm{N}_m M$.
2. The dual vector bundle $\mathrm{N}^*M := (\mathrm{N}M)^*$ is called the conormal bundle of (M, φ). Its fibres are called the conormal spaces of M at $m \in M$. They are denoted by $\mathrm{N}^*_m M$.

Remark 2.7.18

1. The normal and the conormal bundle of (M, φ) are real vector bundles over M of class C^{k-1} and dimension $\dim N - \dim M$. According to Remark 2.7.11/3, $\mathrm{N}M$ is isomorphic to an arbitrary complement of $\mathrm{T}M$ in $(\mathrm{T}N)_{\restriction M}$, and it is often realized in this way. For an example, see Exercise 2.7.6. According to Example 2.7.13, N^*M is naturally isomorphic to the annihilator $(\mathrm{T}M)^0$ of $\mathrm{T}M$ in $(\mathrm{T}N)_{\restriction M}$.
2. By a coorientation, or a transversal orientation, of (M, φ) one means an orientation of $\mathrm{N}M$. Accordingly, (M, φ) is said to be coorientable, or transversally orientable, if the normal bundle $\mathrm{N}M$ of (M, φ) is orientable. This is consistent with the terminology for vector subbundles introduced in Remark 2.7.11/4: a coorientation of (M, φ) is the same as a coorientation of $\mathrm{T}M$ in $(\mathrm{T}N)_{\restriction M}$.
3. We discuss local frames in $\mathrm{N}M$ and N^*M induced by local charts on N adapted to M. Denote $r := \dim M$ and $s := \dim N$. For simplicity, we consider the case of M being a subset of N. We leave it to the reader to write down the respective local frames for the general situation. According to Proposition 1.6.7, for every $m \in M$, there exists an open neighbourhood U of m in M and a local chart (V, ρ) on N at m such that $U \subset V$ and $(U, \rho_{\restriction U})$ is a local chart on M, taking values in the subspace $\mathbb{R}^r \times \{0\} \subset \mathbb{R}^s$. Then, $\{\partial_{i \restriction U} : i = 1, \dots, s\}$ is a local frame in $(\mathrm{T}N)_{\restriction M}$ whose first r elements span $\mathrm{T}M$ over U. According to Remark 2.7.11/2, then $\{[\partial_{i \restriction U}] : i = r + 1, \dots, s\}$ is a local frame in $\mathrm{N}M$. This, in turn, induces a dual local frame $\{[\partial_{i \restriction U}]^* : i = r + 1, \dots, s\}$ in N^*M, see Example 2.4.1. According

[12] By construction, $\tilde{X}$ is also the restriction of X in domain to the submanifold (M, φ) and in range to the subbundle $(\mathrm{T}M, \varphi')$. This does not help for the argument though, because the latter need not be embedded, so that Proposition 1.6.10 does not apply here.

to Example 2.7.13, the natural isomorphism $\mathrm{N}^*M \to (\mathrm{T}M)^0$ maps the latter to the local frame in $(\mathrm{T}M)^0$ consisting of the last $s-r$ elements of the local frame $\{(\mathrm{d}\rho^i)_{\restriction U} : i = 1, \ldots, s\}$ in $((\mathrm{T}N)_{\restriction M})^* \equiv (\mathrm{T}^*N)_{\restriction M}$.

4. Assume that (M, φ) is embedded. The subset

$$C^k_M(N) = \left\{f \in C^k(N) : \varphi^* f = 0\right\} \tag{2.7.1}$$

is an ideal of the associative algebra $C^k(N)$, called the vanishing ideal of M. By means of this ideal, for $m \in M$, the subspaces $\mathrm{T}_m M$ of $\mathrm{T}_{\varphi(m)}N$ and $\mathrm{N}^*_m M \cong (\mathrm{T}_m M)^0$ of $\mathrm{T}^*_{\varphi(m)}(N)$ can be characterized as follows:

$$\varphi'(\mathrm{T}_m M) = \left\{X \in \mathrm{T}_{\varphi(m)}N : X(f) = 0 \text{ for all } f \in C^k_M(N)\right\}, \tag{2.7.2}$$

$$\mathrm{N}^*_m M = \left\{\xi \in \mathrm{T}^*_{\varphi(m)}N : \xi = \mathrm{d}f\big(\varphi(m)\big) \text{ for some } f \in C^k_M(N)\right\}. \tag{2.7.3}$$

The proof is left to the reader (Exercise 2.7.5). Beware that (2.7.2) or (2.7.3) need not hold if M is not embedded. A counterexample is provided by the figure eight submanifold $(\mathbb{R}, \gamma_\pm)$ of Example 1.6.6/2. At the crossing point, the derivative of any element of $C^\infty_M(N)$ vanishes. Hence, for the right hand side of (2.7.2) one obtains $\mathrm{T}_{\varphi(m)}N$.

Exercises

2.7.1 Prove Proposition 2.7.5 by means of Proposition 2.3.15.

2.7.2 Let (E, M, π) be a $\mathbb{K}$-vector bundle of class C^k. Use a system of local trivializations and a subordinate partition of unity of M to construct a C^k-function $\mathsf{h} : E \otimes E \to \mathbb{K}$ such that $\mathsf{h}_m := \mathsf{h}_{\restriction E_m \otimes E_m}$ is a scalar product on E_m for all $m \in M$.

2.7.3 Show that every vertical subbundle admits a complement.

2.7.4 Complete the proof of Proposition 2.7.14.

2.7.5 Prove Eqs. (2.7.2) and (2.7.3) of Remark 2.7.18, characterizing the tangent and the conormal spaces of an embedded submanifold.

2.7.6 Using the Euclidean metric, construct the normal bundle of the submanifold S^n of $\mathbb{R}^{n+1}$ as a complement of TS^n in $(\mathrm{T}\mathbb{R}^{n+1})_{\restriction \mathrm{S}^n}$. Is this bundle trivial?

Chapter 3
Vector Fields

In the first four sections, we discuss elementary aspects of the theory of vector fields. In Sect. 3.1 we show that it is fruitful to view vector fields as derivations of the algebra of functions on the manifold. Next, in Sect. 3.2, we discuss in detail the notions of integral curve and flow and, in Sect. 3.3 we introduce the Lie derivative of a tensor field with respect to a given vector field. Finally, in Sect. 3.4, we extend the notion of an ordinary vector field to that of a time-dependent vector field. Next, we pass to more advanced topics. In Sect. 3.5, we give an introduction to the theory of (geometric) distributions, a notion which generalizes that of a vector field: a distribution is a subset of the tangent bundle consisting of linear subspaces of the tangent spaces. Following the theory developed by Stefan and Sussmann, we discuss the concept of integrability in some detail. The special case of a distribution of constant rank is built in here and the classical Frobenius Theorem occurs as a special case of a general theorem yielding integrability criteria. In the remaining four sections, we give an introduction to the study of the qualitative behaviour of the flows of vector fields.[1] In Sect. 3.6 we collect the basic notions related to critical integral curves.[2] In Sect. 3.7, we introduce the concept of a Poincaré mapping and in Sect. 3.8 we pass to the study of elementary aspects of stability. This notion comprises a variety of concepts characterizing, in effect, two aspects of the long-time behaviour of a flow, namely, returning properties and attraction properties of integral curves or, more generally, of invariant subsets. Here, we limit our attention to the concept of so-called orbital stability, to which for simplicity we refer to merely as stability. At the end of this section we briefly discuss the relation to Lyapunov stability, a notion physicists are probably more familiar with. Finally, in Sect. 3.9, we present the concept of invariant manifolds which plays a basic role in the analysis of the qualitative behaviour of flows near critical integral curves. In all of the four final sections, the reader will find a number of illustrative examples.

---

[1]That is, to the theory of dynamical systems.

[2]Equilibrium points or periodic integral curves.

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7_3, © Springer Science+Business Media Dordrecht 2013

From now on, we restrict attention to smooth manifolds and smooth mappings. Let M be a smooth manifold of dimension n. Recall from Chap. 2 that a vector field on M is a section of the tangent bundle $\mathrm{T}M$ and that the space of vector fields is denoted by $\mathfrak{X}(M)$. Recall, furthermore, that for every local chart (U,κ), the local sections ∂_i form a local frame in $\mathrm{T}M$.

3.1 Vector Fields as Derivations

In this section, we will relate vector fields to derivations of the associative algebra $C^\infty(M)$ and use this to define their commutator. Let $X \in \mathfrak{X}(M)$. Recall from Sect. 1.4 that the tangent vectors X_m, $m \in M$, define directional derivatives of functions and hence mappings $X_m : C^\infty(M) \to \mathbb{R}$. Thus, for $f \in C^\infty(M)$, we can define a function $X(f)$ on M by

$$X(f)(m) := X_m(f), \quad m \in M.$$

This function is smooth, because its local representative with respect to a chart (U,κ) is

$$X(f) \circ \kappa^{-1} = \big(X^i \circ \kappa^{-1}\big) \frac{\partial (f \circ \kappa^{-1})}{\partial x^i},$$

where $X^i \in C^\infty(U)$ are the coefficient functions of X with respect to the local frame $\{\partial_i\}$. By assigning to f the function $X(f)$ we obtain a linear mapping

$$X : C^\infty(M) \to C^\infty(M),$$

denoted by the same symbol. According to (1.4.11), it satisfies

$$X(fg) = X(f)g + fX(g), \quad f, g \in C^\infty(M),$$

hence it is a derivation of the algebra $C^\infty(M)$.[3]

Proposition 3.1.1 *Vector fields on M correspond bijectively to derivations of the algebra $C^\infty(M)$.*

Proof It remains to show that every derivation of $C^\infty(M)$ is generated by a vector field. Thus, for a given derivation δ define a mapping $X_m : C^\infty(M) \to \mathbb{R}$ by

$$X_m(f) := \big(\delta(f)\big)(m)$$

for every $m \in M$. X_m is linear and satisfies

$$X_m(fg) = \big(\delta(fg)\big)(m) = \big(\delta(f)g + f\delta(g)\big)(m) = X_m(f)g(m) + f(m)X_m(g),$$

hence it is a derivation of $C^\infty(M)$ at m. Then, according to Proposition 1.4.7, $X_m \in \mathrm{T}_m M$, and so the assignment $m \mapsto X_m$ defines a mapping $X : M \to \mathrm{T}M$

[3]The derivations of $C^\infty(M)$ are frequently referred to as the first order differential operators on M.

with $\pi \circ X = \mathrm{id}_M$. By construction, $X(f) = \delta(f) \in C^\infty(M)$ for any $f \in C^\infty(M)$. We show that this implies that X is smooth. In view of Proposition 2.3.10, it suffices to show that for every $m \in M$ there exists a local chart (U, κ) at m such that the coefficient functions X^i of X with respect to the induced local frame $\{\partial_i\}$ are smooth. According to Remark 2.3.2/4, by possibly shrinking U we may assume that $\kappa^i = \tilde{\kappa}^i_{\restriction U}$ for some $\tilde{\kappa}^i \in C^\infty(M)$, $i = 1, \ldots, n$. Then, Example 2.3.11/1 and Proposition 1.4.5/2 imply

$$X^i = X_{\restriction U}(\kappa^i) = X_{\restriction U}(\tilde{\kappa}^i_{\restriction U}) = \big(X(\tilde{\kappa}^i)\big)_{\restriction U} = \big(\delta(\tilde{\kappa}^i)\big)_{\restriction U},$$

thus showing that X^i is smooth, indeed. □

For what follows, let us recall the notion of Lie algebra.

Definition 3.1.2 A Lie algebra over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ is a vector space $\mathfrak{g}$ over $\mathbb{K}$ together with a bilinear mapping $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, called multiplication, satisfying the following axioms.

1. Anticommutativity: $[A, A] = 0$ for all $A \in \mathfrak{g}$.
2. Jacobi identity: $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$ for all $A, B, C \in \mathfrak{g}$.

A homomorphism of Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ is a mapping $\varphi : \mathfrak{g} \to \mathfrak{h}$ which is linear and satisfies $\varphi([A, B]) = [\varphi(A), \varphi(B)]$ for all $A, B \in \mathfrak{g}$.

One can check that the commutator $[\delta_1, \delta_2] = \delta_1 \circ \delta_2 - \delta_2 \circ \delta_1$ of derivations δ_1, δ_2 of an associative algebra $\mathfrak{A}$ is a derivation of $\mathfrak{A}$ and that the commutator defines a Lie algebra structure on the space of all derivations of $\mathfrak{A}$ (Exercise 3.1.1). Therefore, Proposition 3.1.1 allows for

Definition 3.1.3 (Commutator of vector fields) The commutator of the vector fields X, Y on M is the vector field $[X, Y]$ on M corresponding to the derivation $[X, Y] = X \circ Y - Y \circ X$.

Thus, $\mathfrak{X}(M)$ with the commutator as a product is a real Lie algebra.

Remark 3.1.4

1. Recall that $\mathfrak{X}(M)$, as a space of sections in a vector bundle over M, is naturally a $C^\infty(M)$-module. The commutator and the module structure are related by
$$[fX, gY] = fg[X, Y] + fX(g)Y - gY(f)X, \tag{3.1.1}$$
with $X, Y \in \mathfrak{X}(M)$ and $f, g \in C^\infty(M)$. The proof is left to the reader (Exercise 3.1.2).
2. Let (U, κ) be a local chart on M and let $X, Y \in \mathfrak{X}(M)$. The coefficient functions of $[X, Y]$ with respect to the local frame $\{\partial_i\}$ are given by (Exercise 3.1.4)
$$[X, Y]^i = X^j \partial_j Y^i - Y^j \partial_j X^i. \tag{3.1.2}$$

3. Let V be a finite-dimensional real vector space, let $M \subset V$ be an open subset and let $X, Y \in \mathfrak{X}(M)$, viewed as mappings $M \to V$, cf. Remark 2.3.4/1. Using (2.3.2), we find that the mapping representing $[X, Y]$ is given by

$$[X, Y](v) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 Y\big(v + tX(v)\big) - \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 X\big(v + tY(v)\big), \quad v \in M. \tag{3.1.3}$$

By writing down this equation in terms of the global chart on M associated with a basis in V one recovers (3.1.2) in this particular situation.
4. For the more general situation of $M \subset V$ being an embedded submanifold, Formula (3.1.3) generalizes to

$$[X, Y](v) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \tilde{Y}\big(v + tX(v)\big) - \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \tilde{X}\big(v + tY(v)\big), \quad v \in M, \tag{3.1.4}$$

where $\tilde{X}$ and $\tilde{Y}$ are arbitrary smooth extensions of X and Y to an open neighbourhood of M in V.

There are two further ways to view a vector field X as a mapping, namely, as a smooth function,

$$X : \mathrm{T}^*M \to \mathbb{R}, \qquad X(\eta) := \eta(X_m), \tag{3.1.5}$$

where $\eta \in \mathrm{T}^*_m M$, and as the morphism of $C^\infty(M)$-modules induced by the natural pairing (2.5.1),

$$X : \Omega^1(M) \to C^\infty(M), \qquad X(\alpha) := \langle \alpha, X \rangle. \tag{3.1.6}$$

In calculations involving vector fields one often switches forth and back between these viewpoints, as for example in the proof of the next proposition.

Proposition 3.1.5 *Let M and N be manifolds and let $\varphi \in C^\infty(M, N)$. Let X, X_1, X_2 be vector fields on M and let Y, Y_1, Y_2 be vector fields on N.*

1. *X is φ-related to Y iff the corresponding derivations satisfy $X \circ \varphi^* = \varphi^* \circ Y$.*
2. *If X_i is φ-related to Y_i, $i = 1, 2$, then $[X_1, X_2]$ is φ-related to $[Y_1, Y_2]$.*

If φ is a diffeomorphism, then, in addition, for the transport operator φ_* the following holds.

3. *As a derivation of $C^\infty(N)$, $\varphi_* X$ is given by $\varphi_* X = (\varphi^{-1})^* \circ X \circ \varphi^*$.*
4. *$[\varphi_* X_1, \varphi_* X_2] = \varphi_*[X_1, X_2]$.*

Recall that if φ is a diffeomorphism, $\varphi_* X$ is the unique vector field on N which is φ-related to X. Hence, assertions 1 and 2 hold in particular for X and $Y = \varphi_* X$.

Proof 1. X is φ-related to Y iff $\varphi' \circ X = Y \circ \varphi$, hence iff $(\varphi' X_m)(f) = Y_{\varphi(m)}(f)$ for all $m \in M$ and $f \in C^\infty(N)$. According to (1.5.3), the left hand side can be rewritten as $(X \circ \varphi^*(f))(m)$, whereas the right hand side yields $(\varphi^* \circ Y(f))(m)$.

2. On the level of derivations, one has

$$\begin{aligned}[X_1, X_2] \circ \varphi^* &= X_1 \circ X_2 \circ \varphi^* - X_2 \circ X_1 \circ \varphi^* \\ &= \varphi^* \circ Y_1 \circ Y_2 - \varphi^* \circ Y_2 \circ Y_1 \\ &= \varphi^* \circ [Y_1, Y_2].\end{aligned}$$

Points 3 and 4 follow from points 1 and 2 by putting $Y = \varphi_* X$ and $Y_i = \varphi_* X_i$, respectively. □

Corollary 3.1.6 *Let (N, φ) be a submanifold of M. Let $X, Y \in \mathfrak{X}(M)$ be tangent to N and let $\tilde{X}, \tilde{Y} \in \mathfrak{X}(N)$ be the restrictions to N, see Proposition* 2.7.16. *Then, $[X, Y]$ is tangent to N and $[\tilde{X}, \tilde{Y}]$ is its restriction.*

Proof By Proposition 2.7.16, $\tilde{X}$ and $\tilde{Y}$ are φ-related to X and Y, respectively. By Proposition 3.1.5/2, the vector field $[\tilde{X}, \tilde{Y}]$ on N is φ-related to $[X, Y]$. This implies that $[X, Y]$ is tangent to N and $[\tilde{X}, \tilde{Y}]$ is its restriction. □

Corollary 3.1.6 reproduces the following direct consequence of Proposition 1.4.5:

$$[X, Y]_{\restriction U} = [X_{\restriction U}, Y_{\restriction U}] \quad \text{for all } X, Y \in \mathfrak{X}(M),\ U \subset M \text{ open.} \tag{3.1.7}$$

Exercises

3.1.1 Let $\mathfrak{A}$ be an associative $\mathbb{K}$-algebra and let $\mathrm{Der}(\mathfrak{A})$ denote the $\mathbb{K}$-vector space of derivations of $\mathfrak{A}$. Show that
  (a) for $D_1, D_2 \in \mathrm{Der}(\mathfrak{A})$, the commutator $[D_1, D_2] := D_1 \circ D_2 - D_2 \circ D_1$ is in $\mathrm{Der}(\mathfrak{A})$,
  (b) $\mathrm{Der}(\mathfrak{A})$ with the commutator as a product is a Lie algebra.

3.1.2 Prove Formula (3.1.1).

3.1.3 Compute the commutator of $X_1 = (x+y)\frac{\partial}{\partial x} - \frac{\partial}{\partial y}$ and $X_2 = (y^2+1)\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$ on $\mathbb{R}^2$.

3.1.4 Verify Formula (3.1.2).

3.1.5 Use (2.3.2) to verify (3.1.3).

3.2 Integral Curves and Flows

Let M be a manifold and let X be a vector field on M. Consider a smooth curve $\gamma : I \to M$, where $I \subset \mathbb{R}$ is some open interval. According to Example 1.5.6, the tangent vector $\dot{\gamma}(t)$ of γ at $t \in I$ is defined to be the tangent vector of M at $\gamma(t)$ given by the curve $s \mapsto \gamma(t+s)$.

Definition 3.2.1 (Integral curve) Let $X \in \mathfrak{X}(M)$. A smooth curve γ on M is an integral curve of X if

$$\dot{\gamma}(t) = X_{\gamma(t)} \quad \text{for all } t \in I. \tag{3.2.1}$$

Less formally, an integral curve of a vector field is a smooth curve whose tangent vectors are given by the values of this vector field along the curve. In terms of derivations, the defining condition is equivalent to

$$X_{\gamma(t)}(f) = \frac{\mathrm{d}}{\mathrm{d}t}{\restriction_t} (f \circ \gamma) \quad \text{for all } f \in C^\infty(M).$$

By an extension of an integral curve γ of X we mean an integral curve $\tilde{\gamma}$ of X which extends γ in the sense of mappings.

Definition 3.2.2 (Maximality and completeness) An integral curve of a vector field on M is called maximal if there does not exist a proper extension. It is called complete if it has domain $I = \mathbb{R}$. A vector field is called complete if its maximal integral curves are complete.

We analyse Eq. (3.2.1) in a local chart (U, κ) on M. According to (1.5.5), for $t \in I$ such that $\gamma(t) \in U$, (3.2.1) is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}s}{\restriction_t} (\kappa \circ \gamma)(s) = \mathbf{X}^\kappa_{\gamma(t)}. \tag{3.2.2}$$

Using the notation $\mathbf{x}(t) := \kappa \circ \gamma(t)$ and $X^i(\mathbf{x}) := X^{\kappa,i}_{\kappa^{-1}(\mathbf{x})}$, one may rewrite (3.2.2) as

$$\dot{x}^i(t) = X^i\big(\mathbf{x}(t)\big), \quad i = 1, \ldots, n. \tag{3.2.3}$$

This is a system of ordinary first order differential equations with smooth coefficients on the open subset $\kappa(U)$ of $\mathbb{R}^n$. The fundamental existence and uniqueness theorems for the solutions of such systems state the following.

(a) For every $\mathbf{x}_0 \in \kappa(U)$ there exists a unique solution $I_{\mathbf{x}_0} \ni t \mapsto \mathbf{x}(t; \mathbf{x}_0) \in \kappa(U)$ which is maximal among all solutions defined on open intervals containing 0 and satisfying the initial condition $\mathbf{x}(0; \mathbf{x}_0) = \mathbf{x}_0$.
(b) The set $\mathscr{D} = \bigcup_{\mathbf{x}_0 \in \kappa(U)} I_{\mathbf{x}_0} \times \{\mathbf{x}_0\}$ is open in $\mathbb{R} \times \kappa(U)$.
(c) The mapping $\mathscr{D} \to \kappa(U)$, $(t, \mathbf{x}_0) \mapsto \mathbf{x}(t; \mathbf{x}_0)$, is smooth.

Theorem 3.2.3 (Existence and uniqueness of maximal integral curves) *Let M be a manifold and let $X \in \mathfrak{X}(M)$. For every $m \in M$ there exists a unique maximal integral curve $\gamma_m : I_m \to M$ of X with $\gamma_m(0) = m$.*

Proof Let $m \in M$ be given. According to the points (a)–(c) above, the fundamental existence and uniqueness theorems for solutions of (3.2.3) imply that

1. integral curves through m exist,
2. if two integral curves through m coincide at t, then they coincide on an open interval containing t.

Uniqueness follows from 2. To prove existence, let I_m be the union of the domains of all integral curves through m. Due to 1 and since the domain of any integral curve through m is an open interval containing the common point 0, I_m is a nonempty open interval. Thus, in order that all the integral curves through m combine to a single maximal integral curve γ_m with domain I_m, it remains to show that any two of them, say $\gamma_1 : I_1 \to M$ and $\gamma_2 : I_2 \to M$, coincide on their common domain $I_1 \cap I_2$. Let $J \subset I_1 \cap I_2$ be the set of t such that $\gamma_1(t) = \gamma_2(t)$. Since J contains 0, it is nonempty. Due to 2, J is open in $I_1 \cap I_2$. J is also closed: if a sequence $\{t_n\}$ in J converges to a point t in $I_1 \cap I_2$ then, by continuity of γ_i,

$$\gamma_1(t) = \lim_{n\to\infty} \gamma_1(t_n) = \lim_{n\to\infty} \gamma_2(t_n) = \gamma_2(t).$$

Thus, J is a nonempty open and closed subset of the interval $I_1 \cap I_2$. This implies $J = I_1 \cap I_2$ and hence the assertion. □

Corollary 3.2.4 *For* $t \in I_m$, $I_{\gamma_m(t)} = \{s \in \mathbb{R} : s + t \in I_m\}$. *For* $s \in I_{\gamma_m(t)}$, *one has*

$$\gamma_{\gamma_m(t)}(s) = \gamma_m(t + s).$$

Proof Denote $I_m - t := \{s \in \mathbb{R} : s + t \in I_m\}$. The smooth curve $\gamma : I_m - t \to M$, defined by $\gamma(s) := \gamma_m(t + s)$, is an integral curve of X through $\gamma(0) = \gamma_m(t)$. Hence, $I_m - t \subset I_{\gamma_m(t)}$. Interchanging the roles of m and $\gamma_m(t)$ and applying the same argument with t replaced by $-t$, one finds $I_{\gamma_m(t)} + t \subset I_m$. It follows that $I_m - t = I_{\gamma_m(t)}$ and that $\gamma_{\gamma_m(t)}(s) = \gamma(s) = \gamma_m(t + s)$ for all $s \in I_m - t$. □

The maximal integral curves combine to what is called a flow.[4]

Definition 3.2.5 (Flow) Let $\mathscr{D}$ be an open neighbourhood of $\{0\} \times M$ in $\mathbb{R} \times M$ and let $\Phi\colon \mathscr{D} \to M$ be a smooth mapping. For $m \in M$ and $t \in \mathbb{R}$, denote

$$\mathscr{D}_m := \big\{t \in \mathbb{R} : (t, m) \in \mathscr{D}\big\}, \qquad \mathscr{D}_t := \big\{m \in M : (t, m) \in \mathscr{D}\big\}$$

and let $\Phi_m : D_m \to M$ and $\Phi_t : D_t \to M$ denote the induced mappings given by

$$\Phi_m(t) = \Phi_t(m) = \Phi(t, m).$$

The mapping Φ is called a (smooth) flow on M if the following holds:

1. $\Phi_{t=0} = \mathrm{id}_M$,
2. if $(s, m) \in \mathscr{D}$ and $(t, \Phi_s(m)) \in \mathscr{D}$, then $(s + t, m) \in \mathscr{D}$ and $\Phi_t(\Phi_s(m)) = \Phi_{s+t}(m)$,
3. $\mathscr{D}_m$ is connected and hence an open interval for all $m \in M$,
4. $\Phi_t(\mathscr{D}_t) \subset \mathscr{D}_{-t}$ for all $t \in \mathbb{R}$.

A flow is called maximal if it does not admit an extension, in the sense of mappings, which itself is a flow. It is called complete if $\mathscr{D} = \mathbb{R} \times M$.

[4]This name alludes to the intuitive picture of the points of M moving synchronously along their individual maximal integral curves.

Now, for a given vector field X, by analogy with the local system (3.2.3), we form the subset

$$\mathscr{D} := \bigcup_{m \in M} I_m \times \{m\} \tag{3.2.4}$$

of $\mathbb{R} \times M$ and define the mapping

$$\Phi : \mathscr{D} \to M, \qquad \Phi(t, m) := \gamma_m(t), \tag{3.2.5}$$

where $\gamma_m : I_m \to M$ is the maximal integral curve of X through m provided by Theorem 3.2.3.

Proposition 3.2.6 (Vector fields and flows) *The assignment of Φ to X given by* (3.2.4) *and* (3.2.5) *defines a bijection between vector fields and maximal flows on M. Complete vector fields thereby correspond to complete flows.*

We will refer to Φ given by (3.2.4) and (3.2.5) as the flow of the vector field X. If we have to deal with the flows of several vector fields we will occasionally write Φ^X and $\mathscr{D}^X$. The notion of flow extends in an obvious way to local vector fields: by definition, the flow of the local vector field X over $U \subset M$ is the flow of the vector field X on the manifold U.

Proof First, let $X \in \mathfrak{X}(M)$. We must show that Φ is a flow. By Theorem 3.2.3, we have $\{0\} \times M \subset \mathscr{D}$. Recall the consequences (a)–(c) of the fundamental existence and uniqueness theorems for solutions of the local system (3.2.3), listed before Theorem 3.2.3. Due to (b), $\mathscr{D}$ is open in $\mathbb{R} \times M$. Due to (c), Φ is smooth. Properties 1, 3 and 4 of Definition 3.2.5 hold by construction and property 2 follows from Corollary 3.2.4. To check the maximality property, let $\tilde{\mathscr{D}}$ be an open subset of $\mathbb{R} \times M$ containing $\mathscr{D}$ and let $\tilde{\Phi} : \mathscr{D} \to M$ be a smooth mapping possessing the properties 1–4 and satisfying $\tilde{\Phi}_{\restriction \mathscr{D}} = \Phi$. Let $m \in M$. Due to property 3, the mapping $\tilde{\Phi}_m : \tilde{\mathscr{D}}_m \to M$ is a curve. Using property 2, for $t \in \tilde{\mathscr{D}}_m$ one obtains

$$\frac{\mathrm{d}}{\mathrm{d}s}_{\restriction_t} \tilde{\Phi}_m(s) = \frac{\mathrm{d}}{\mathrm{d}s}_{\restriction_0} \tilde{\Phi}_s\big(\tilde{\Phi}_t(m)\big) = \frac{\mathrm{d}}{\mathrm{d}s}_{\restriction_0} \Phi_s\big(\tilde{\Phi}_t(m)\big) = X_{\tilde{\Phi}_t(m)}.$$

This shows that $\tilde{\Phi}_m$ is an integral curve of X through m. It follows that $\tilde{\mathscr{D}}_m \subset \mathscr{D}_m$ for all $m \in M$. This implies $\tilde{\mathscr{D}} = \mathscr{D}$ and $\tilde{\Phi} = \Phi$.

Conversely, let $\tilde{\Phi} : \tilde{\mathscr{D}} \to M$ be a maximal flow on M. For $m \in M$, $\tilde{\Phi}_m : \tilde{\mathscr{D}}_m \to M$ is a smooth curve through m and hence defines a tangent vector

$$X_m := \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction_0} \tilde{\Phi}_m(t).$$

The assignment of X_m to m defines a section $X : M \to \mathrm{T}M$. Its local representative with respect to a chart (U, κ) on M and the induced chart on $\mathrm{T}M$ is given by

$$\mathbf{x} \mapsto \left(\mathbf{x}, \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction_0} \kappa \circ \tilde{\Phi}\big(\kappa^{-1}(\mathbf{x}), t\big)\right).$$

Hence, it is smooth. Therefore, X is a smooth vector field on M.

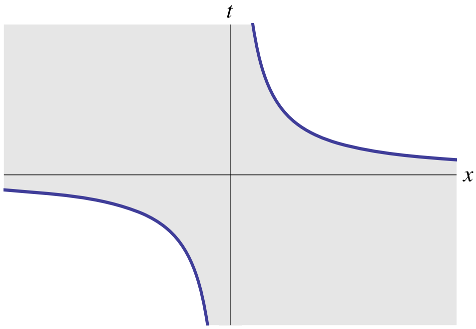


Fig. 3.1 Domain $\mathscr{D}$ of the flow of the vector field $X(x) = x^2\partial_x$ on $M = \mathbb{R}$

Finally, we have to check that the above assignments between vector fields and flows are inverse to one another. Given a vector field X, the vector field assigned to the flow of X obviously coincides with X. Given a flow $\tilde{\Phi} : \tilde{\mathscr{D}} \to M$, let $\Phi : \mathscr{D} \to M$ be the flow of the vector field defined by $\tilde{\Phi}$. Then, $\tilde{\Phi}$ and Φ coincide on $\tilde{\mathscr{D}} \cap \mathscr{D}$. Since both $\tilde{\Phi}$ and Φ are maximal, it follows that $\tilde{\mathscr{D}} = \mathscr{D}$ and $\tilde{\Phi} = \Phi$. This proves the proposition. □

Example 3.2.7

1. Let $M = U \subset \mathbb{R}^n$ be an open subset and let $X = \partial_i$. The system (3.2.3) reads $\dot{\mathbf{x}} = \mathbf{e}_i$ and the unique maximal solution with initial condition $\mathbf{x}(0) = \mathbf{x}$ is given by $\mathbf{x}(t) = \mathbf{x} + t\mathbf{e}_i$, where $t \in \mathbb{R}$ is such that $\mathbf{x} + s\mathbf{e}_i \in U$ for all $s \in [0, t]$. Hence,
$$\mathscr{D} = \big\{(t, \mathbf{x}) \in \mathbb{R} \times U : \mathbf{x} + s\mathbf{e}_i \in U \text{ for all } s \in [0, t]\big\}, \qquad \Phi_t(\mathbf{x}) = \mathbf{x} + t\mathbf{e}_i.$$
2. Let $M = \mathbb{R}$ and $X(x) = x^2\partial_x$. The system (3.2.3) consists of the single equation $\dot{x} = x^2$. Integration with initial condition $x(0) = x$ yields
$$x(t) = \frac{x}{1 - tx}.$$
This is defined for all $t < 1/x$. Hence, the flow is given by
$$\mathscr{D} = \left\{(t, x) \in \mathbb{R} \times M : t < \frac{1}{x}\right\}, \qquad \Phi_t(x) = \frac{x}{1 - tx}.$$
The domain $\mathscr{D}$ is shown in Fig. 3.1. In particular, X is not complete.

We use Example 3.2.7/2 to show that neither the sum nor the commutator of complete vector fields need be complete. Let $M = \mathbb{R}^2$ with coordinates x and y. For $f_1, f_2 \in C^\infty(\mathbb{R})$, the vector fields
$$X_1(x, y) = f_1(y)\partial_x, \qquad X_2(x, y) = f_2(x)\partial_y$$
are complete and have the flows
$$\Phi_t^{X_1}(x, y) = \big(x + tf_1(y), y\big), \qquad \Phi_t^{X_2}(x, y) = \big(x, y + tf_2(x)\big).$$
For $f_1(s) = f_2(s) = \frac{1}{2}s^2$, the sum $X_1 + X_2$ is tangent to the diagonal $\Delta \subset M$ defined by $x = y$. Hence, by Proposition 2.7.16, it induces a vector field X on Δ. Using the

global coordinate x on Δ, one finds $X = x^2\partial_x$, which was shown above to be not complete. Hence, $X_1 + X_2$ is not complete. For $f_1(s) = s$ and $f_2(s) = s^2$, we have

$$[X_1, X_2](x) = -2xy\partial_x + y^2\partial_y.$$

This vector field is tangent to the y-axis. Using y as a global coordinate on this submanifold, for the induced vector field one finds again $X = y^2\partial_y$. Hence, $[X_1, X_2]$ is not complete.

Example 3.2.8 (Linear vector fields) Let V be a finite-dimensional real vector space. In terms of the natural representation of vector fields on V by smooth mappings $X : V \to V$, cf. Remark 2.3.4/1, the system (3.2.1) reads

$$\dot{v} = X\big(v(t)\big). \tag{3.2.6}$$

A vector field on V is called linear if the corresponding mapping is linear, that is, $X(v) = Av$ for some $A \in \mathrm{End}(V)$. In this case, X is said to be generated by A. The solution of (3.2.6) with initial condition $v(0) = v_0$ is

$$v(t) = \mathrm{e}^{tA}v_0, \quad \mathrm{e}^{tA} = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!},$$

where $(tA)^n$ means n-fold composition of the endomorphism tA with itself.[5] Thus, X is complete and its flow Φ is given by the one-parameter group of automorphisms of V

$$\Phi_t = \mathrm{e}^{tA}, \quad t \in \mathbb{R}. \tag{3.2.7}$$

Conversely, if Φ is a one-parameter group of automorphisms of V, the generating vector field is given by $X_v = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \Phi_t(v) = (\frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \Phi_t)(v)$, that is, it is linear and corresponds to the endomorphism $\frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \Phi_t$ of V.

For later use, we determine the flow $\Phi_t = \mathrm{e}^{tA}$ explicitly. The necessary calculations are left to the reader (Exercise 3.2.2). Choose a basis in V such that the matrix associated with A has Jordan normal form with Jordan blocks $B_1, \dots, B_r$. Let λ_i denote the eigenvalue of A with nonnegative imaginary part corresponding to B_i and let m_i denote its multiplicity.[6] Let N_k denote the $(k \times k)$-matrix with entries 1 on the upper secondary diagonal and 0 elsewhere. For $\lambda = \alpha + \mathrm{i}\beta$ with α, β real, define

$$\mathfrak{R}(\lambda) = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}.$$

---

[5]The exponential series $\sum_{n=0}^{\infty} \frac{A^n}{n!}$ is convergent for all $A \in \mathrm{End}(V)$. This follows from the fact that the operator norm on $\mathrm{End}(V)$ induced by an arbitrary norm on V satisfies $\|A^n\| \le \|A\|^n$.

[6]I.e., the dimension of the corresponding Jordan block of the extension of A to the complexification of V.

For a complex $(k \times k)$-matrix C, let $\mathfrak{R}(C)$ be the real $(2k \times 2k)$-matrix obtained by replacing each entry C_{ij} by the (2×2)-block $\mathfrak{R}(C_{ij})$. The Jordan normal form of A is given by

$$\mathrm{diag}(B_1, \dots, B_r), \qquad B_i = \begin{cases} \lambda_i \mathbb{1}_{m_i} + N_{m_i} & \lambda_i \in \mathbb{R}, \\ \mathfrak{R}(\lambda_i \mathbb{1}_{m_i} + N_{m_i}) & \lambda_i \notin \mathbb{R}. \end{cases} \tag{3.2.8}$$

Then, e^{tA} is represented by the matrix $\mathrm{diag}(\mathrm{e}^{tB_1}, \dots, \mathrm{e}^{tB_r})$, where

$$\mathrm{e}^{tB_i} = \mathrm{e}^{\lambda_i t} \begin{bmatrix} 1 & t & \cdots & \frac{t^{m_i-1}}{(m_i-1)!} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t \\ 0 & \cdots & 0 & 1 \end{bmatrix} \tag{3.2.9}$$

for $\lambda_i \in \mathbb{R}$ and

$$\mathrm{e}^{tB_i} = \mathrm{e}^{\alpha_i t} \begin{bmatrix} D_{\beta_i t} & t\, D_{\beta_i t} & \cdots & \frac{t^{m_i-1}}{(m_i-1)!} D_{\beta_i t} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t\, D_{\beta_i t} \\ 0 & \cdots & 0 & D_{\beta_i t} \end{bmatrix}, \qquad D_\phi = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}, \tag{3.2.10}$$

for $\lambda_i \notin \mathbb{R}$. For a detailed discussion of the linear vector fields on $\mathbb{R}^2$, see Example 3.6.13.

Remark 3.2.9

1. Successive application of the vector fields $X_1, \dots, X_r$ to a function $f \in C^\infty(M)$ yields

$$\big(X_r \circ \cdots \circ X_1(f)\big)(m) = \frac{\mathrm{d}}{\mathrm{d}t_1}_{\restriction 0} \cdots \frac{\mathrm{d}}{\mathrm{d}t_r}_{\restriction 0} f\big(\Phi^{X_1}_{t_1} \circ \cdots \circ \Phi^{X_r}_{t_r}(m)\big). \tag{3.2.11}$$

In particular, for $X, Y \in \mathfrak{X}(M)$,

$$[X, Y]_m(f) = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} \frac{\mathrm{d}}{\mathrm{d}s}_{\restriction 0} f\big(\Phi^Y_s\big(\Phi^X_t(m)\big)\big) - \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} \frac{\mathrm{d}}{\mathrm{d}s}_{\restriction 0} f\big(\Phi^X_s\big(\Phi^Y_t(m)\big)\big). \tag{3.2.12}$$

Furthermore, for $X \in \mathfrak{X}(M)$, $f \in C^\infty(M)$ and $m \in M$, Taylor expansion of the smooth function $t \mapsto f(\Phi^X_t(m))$ at $t = 0$ and computation of the derivatives using (3.2.11) yields the following Taylor formula for manifolds:

$$f\big(\Phi^X_t(m)\big) = \sum_{k=1}^{n} \frac{t^k}{k!}\big(X^k(f)\big)(m) + O\big(t^{n+1}\big), \quad t \in \mathscr{D}^X_m, \tag{3.2.13}$$

where $O(t^{n+1})$ is a smooth function on $\mathscr{D}^X_m$ such that $O(t^{n+1})/t^{n+1}$ is bounded. Repeated application of this formula yields the iterated Taylor formula for manifolds

$$f\big(\Phi^X_t\big(\Phi^Y_s(m)\big)\big) = \sum_{k,l=1}^{n} \frac{t^k s^l}{k!l!} Y^l\big(X^k(f)\big)(m) + O\big(t^{n+1}, s^{n+1}\big) \tag{3.2.14}$$

which can be easily generalized to an arbitrary number of vector fields.

2. Let (N, φ) be a submanifold of M, let X be a vector field on M which is tangent to N and let $\tilde{X}$ be the restriction to N, see Proposition 2.7.16. The flow $\tilde{\Phi} : \tilde{\mathscr{D}} \to N$ of $\tilde{X}$ is related to the flow $\Phi : \mathscr{D} \to M$ of X as follows. The domain $\tilde{\mathscr{D}}$ consists of the pairs (t, p) in $\mathbb{R} \times N$ satisfying $(t, \varphi(p)) \in \mathscr{D}$ and $\Phi_s(\varphi(p)) \in \varphi(N)$ for all s between 0 and t. The mapping $\tilde{\Phi}$ fulfils $\varphi(\tilde{\Phi}_t(p)) = \Phi_t(\varphi(p))$.

In the remainder of this section, we derive the basic properties of flows.

Proposition 3.2.10 *Let X be a vector field on M and let $\Phi : \mathscr{D} \to M$ be its flow.*

1. *For every $t \in \mathbb{R}$ such that $\mathscr{D}_t$ is nonempty, Φ_t is a diffeomorphism from $\mathscr{D}_t$ onto $\mathscr{D}_{-t}$ with inverse Φ_{-t}.*
2. *For every $t \in \mathbb{R}$ and every s between 0 and t, one has $\mathscr{D}_t \subset \mathscr{D}_s$ and $\Phi_s(\mathscr{D}_t) \subset \mathscr{D}_{t-s}$.*
3. *If Φ is complete, then $\Phi_t \circ \Phi_s = \Phi_{t+s}$ for all $t, s \in \mathbb{R}$.*
4. *For $s \in \mathbb{R}$, let $\Phi^{sX} : \mathscr{D}^{sX} \to M$ be the flow of the vector field sX. Then, for all $t \in \mathbb{R}$,*

$$\mathscr{D}^{sX}_t = \mathscr{D}_{st}, \qquad \Phi^{sX}_t = \Phi_{st}. \tag{3.2.15}$$

Since every flow satisfies $\Phi_0 = \mathrm{id}_M$, point 3 states that a complete flow defines an action of the additive group $\mathbb{R}$ on M.[7] Complete flows are therefore also called one-parameter groups of local diffeomorphisms. Generalizing this terminology, flows are sometimes referred to as local one-parameter groups of diffeomorphisms.

Proof We prove assertion 1. By the defining property 4 of flows, we have $\Phi_t(\mathscr{D}_t) \subset \mathscr{D}_{-t}$ and $\Phi_{-t}(\mathscr{D}_{-t}) \subset \mathscr{D}_t$. Applying Φ_t to the second relation and using property 2, we obtain $\mathscr{D}_{-t} \subset \Phi_t(\mathscr{D}_t)$. Thus, $\Phi_t(\mathscr{D}_t) = \mathscr{D}_{-t}$ and $\Phi_{-t}(\mathscr{D}_{-t}) = \mathscr{D}_t$. By the properties 1 and 2, for every $m \in \mathscr{D}_t$, we have $\Phi_{-t}(\Phi_t(m)) = \Phi_0(m) = m$. The proof of assertions 2 and 4 is left to the reader (Exercise 3.2.3). Assertion 3 is obvious. □

Proposition 3.2.11 *Let X be a vector field on M, let $\Phi : \mathscr{D} \to M$ be its flow and let $m \in M$.*

1. *If $X_m = 0$, then $\mathscr{D}_m = \mathbb{R}$ and $\Phi_m(t) = m$ for all $t \in \mathbb{R}$.*
2. *If $X_m \neq 0$ and Φ_m is injective, $(\mathscr{D}_m, \Phi_m)$ is a submanifold of M diffeomorphic to $\mathbb{R}$.*

[7] Group actions on manifolds will be treated in detail in Chap. 6.

3. *If* $X_m \neq 0$ *and* Φ_m *is not injective,* $\mathscr{D}_m = \mathbb{R}$ *and the image of* Φ_m *is an embedded submanifold of* M *diffeomorphic to* S^1.

In particular, the images of the maximal integral curves of X carry natural submanifold structures. It will follow by a general argument in Section 3.5 that these submanifold structures are initial and hence unique, cf. Remark 3.5.16. Of course, this is relevant in the situation of point 2 only.

Proof 1. If $X_m = 0$, the curve $\gamma : \mathbb{R} \to M$, $\gamma(t) := m$, is a maximal integral curve of X through m. Hence, by uniqueness, $\mathscr{D}_m = \mathbb{R}$ and $\mathscr{D}_m = \gamma$.

2. and 3. If $X_m \neq 0$, then $X_{\Phi_m(t)} \neq 0$ for all $t \in \mathscr{D}_m$, because otherwise Φ_m would be constant by point 1. It follows that Φ_m is an immersion. Now, if Φ_m is injective, it is a submanifold diffeomorphic to the open interval $\mathscr{D}_m$ and hence to $\mathbb{R}$. Otherwise, there exists a minimal positive $T \in \mathscr{D}_m$ such that $\Phi_m(T) = m$. Then $\mathscr{D}_{\Phi_T(m)} = \mathscr{D}_m$, hence $2T \in \mathscr{D}_m$ and $\mathscr{D}_{\Phi_{2T}(m)} = \mathscr{D}_m$. Iterating this argument one finds $\mathscr{D}_m = \mathbb{R}$. For every $t \in \mathbb{R}$ and $k \in \mathbb{Z}$,

$$\Phi_m(t + kT) = \Phi_{t+kT}(m) = \Phi_t(m) = \Phi_m(t). \tag{3.2.16}$$

Consider the topological quotient $\mathbb{R}/\mathbb{Z}T$ of $\mathbb{R}$ by the equivalence relation $t \sim s$ iff $(t - s) = kT$ for some $k \in \mathbb{Z}$. By choosing representatives in the open intervals $(0, T)$ or $(-\frac{T}{2}, \frac{T}{2})$, respectively, one obtains two local charts which are smoothly compatible and hence define a smooth structure on $\mathbb{R}/\mathbb{Z}T$. With respect to this smooth structure, $\mathbb{R}/\mathbb{Z}T$ is diffeomorphic to the sphere S^1. Due to (3.2.16), Φ_m descends to an immersion $\tilde{\Phi}_m : \mathbb{R}/\mathbb{Z}T \to M$. Since T was chosen to be minimal, $\tilde{\Phi}_m$ is injective and hence defines a submanifold of M whose image coincides with the image of Φ_m. By Remark 1.6.13/2, this submanifold is embedded. This yields the assertion. □

Proposition 3.2.12 *On a compact manifold, every vector field is complete.*

Proof Let $\Phi : \mathscr{D} \to M$ be the flow of a given vector field and let $m \in M$. Assume that $\mathscr{D}_m = (a, b)$ with $b < \infty$. Choose a sequence $\{t_n\}$ in $\mathscr{D}_m$ converging to b. By compactness of M, the sequence $\{\gamma(t_n)\}$ has a cluster point $\tilde{m} \in M$. Since $\mathscr{D}$ is an open neighbourhood of $\{0\} \times M$ in $\mathbb{R} \times M$, there exists $\varepsilon > 0$ and a neighbourhood U of $\tilde{m}$ in M such that $(-\varepsilon, \varepsilon) \times U \subset \mathscr{D}$. By construction of $\tilde{m}$, there exists $t \in \mathscr{D}_m$ such that $|b - t| < \varepsilon$ and $\Phi_m(t) \in U$ and hence $(b - t, \Phi_t(m)) \in \mathscr{D}$. Since also $(t, m) \in \mathscr{D}$, property 2 of Definition 3.2.5 yields $(b - t) + t = b \in \mathscr{D}_m$ (contradiction). Hence, $b = \infty$ and, analogously, $a = -\infty$. □

Proposition 3.2.13 (Flow and transport) *Let* M *and* N *be manifolds. Let* $\varphi \in C^\infty(M, N)$, $X \in \mathfrak{X}(M)$, $Y \in \mathfrak{X}(N)$ *and* $t \in \mathbb{R}$.

1. *If* X *and* Y *are* φ*-related, then* $\varphi(\mathscr{D}^X_t) \subset \mathscr{D}^Y_t$ *and* $\Phi^Y_t \circ \varphi_{\restriction \mathscr{D}^X_t} = \varphi \circ \Phi^X_t$.
2. *If* φ *is a diffeomorphism, then* $\mathscr{D}^{\varphi_* X}_t = \varphi(\mathscr{D}^X_t)$ *and*

$$\Phi^{\varphi_* X}_t = \varphi \circ \Phi^X_t \circ \left(\varphi^{-1}\right)_{\restriction \mathscr{D}^{\varphi_* X}_t}.$$

3. *If in addition $M = N$ and $\varphi_* X = X$, then $\varphi(\mathscr{D}_t^X) = \mathscr{D}_t^X$ and*

$$\varphi \circ \Phi_t^X = \Phi_t^X \circ \varphi_{\restriction \mathscr{D}_t^X}.$$

Proof 1. Let $m \in \mathscr{D}_t^X$. By assumption, for any $s \in \mathscr{D}_m^X$,

$$\tfrac{\mathrm{d}}{\mathrm{d}s}_{\restriction s} \varphi \circ \Phi_m^X(s) = \varphi' X_{\Phi_m^X(s)} = Y_{\varphi \circ \Phi_m^X(s)},$$

hence $\varphi \circ \Phi_m^X$ is an integral curve of Y with initial condition $\varphi \circ \Phi_m^X(0) = \varphi(m)$. Therefore,

$$\varphi \circ \Phi_s^X(m) = \varphi \circ \Phi_m^X(s) = \Phi_{\varphi(m)}^Y(s) = \Phi_s^Y \circ \varphi(m), \quad s \in \mathscr{D}_m^X.$$

Since $t \in \mathscr{D}_m^X$, this yields the assertion.

2. Apply point 1 to X and $Y = \varphi_* X$. Then, replace X by $\varphi_* X$ and φ by φ^{-1}.

3. This is an immediate consequence of point 2. □

Example 3.2.14 Let (U, κ) be a local chart on M. We determine the flow of the local vector field ∂_i on U. Using the diffeomorphism $\kappa : U \to \kappa(U)$, Proposition 3.2.13/2 and Example 3.2.7, we obtain

$$\mathscr{D}^{\partial_i} = \big\{(t, m) \in \mathbb{R} \times M : \kappa(m) + s\mathbf{e}_i \in \kappa(U) \text{ for all } s \in [0, t]\big\},$$
$$\Phi_t^{\partial_i}(m) = \kappa^{-1}\big(\kappa(m) + t\mathbf{e}_i\big).$$

That is, the maximal integral curve through m is given by the connected component of m of the i-th coordinate line through m, parameterized by the coordinate itself minus $\kappa^i(m)$.

Proposition 3.2.15 (Flows of commuting vector fields) *Let M be a manifold and let $X, Y \in \mathfrak{X}(M)$. We have $[X, Y] = 0$ iff*

$$\Phi_t^X \circ \Phi_s^Y(m) = \Phi_s^Y \circ \Phi_t^X(m) \tag{3.2.17}$$

for all $t = t_0, s = s_0 \in \mathbb{R}$ and $m \in M$ such that both sides are defined for all t between 0 and t_0 and s between 0 and s_0.[8]

Proof First, assume that Φ^X and Φ^Y commute. For every $m \in M$ there exists $\varepsilon > 0$ such that both sides of (3.2.17) are defined for $t, s \in (-\varepsilon, \varepsilon)$. Thus, in view of (3.2.12), Formula (3.2.17) implies $[X, Y]_m(f) = 0$ for all $f \in C^\infty(M)$, hence $[X, Y]_m = 0$. Conversely, assume that $[X, Y] = 0$ and let $m \in M$ and $t_0, s_0 \in \mathbb{R}$ be such that both sides of (3.2.17) are defined for all t between 0 and t_0 and s between 0 and s_0. In the following, t and s are assumed to satisfy these inequalities without further notice. To prove (3.2.17) it suffices to show that, for all t, the curve $s \mapsto \Phi_t^X \circ \Phi_s^Y(m)$ is an integral curve of Y through the point $\Phi_t^X(m)$. This is equivalent to

$$\big(\Phi_t^X\big)' Y_{\Phi_s^Y(m)} = Y_{\Phi_t^X \circ \Phi_s^Y(m)} \tag{3.2.18}$$

[8]One says that the flows of X and Y commute.

for all t and s. Fix t and s and denote $p := \Phi_t^X \circ \Phi_s^Y(m)$. Equation (3.2.18) is equivalent to

$$\left(\Phi_u^X\right)' Y_{\Phi_{-u}^X(p)} = Y_p \tag{3.2.19}$$

for all u between 0 and t. The left hand side of this equation is smooth in u, because its coordinates with respect to the chart on $\mathrm{T}M$ induced by a chart (U, κ) on M are given by $\kappa^i(p)$ and

$$\left(\left(\Phi_u^X\right)' Y_{\Phi_{-u}^X(p)}\right)\left(\kappa^i\right) = \frac{\mathrm{d}}{\mathrm{d}s}_{\restriction 0} \left(\kappa^i \circ \Phi_u^X \circ \Phi_s^Y \circ \Phi_{-u}^X(p)\right).$$

Since the right hand side is a partial derivative of a smooth function of the two variables u and s, it is smooth in u. Hence, we may differentiate the left hand side of (3.2.19) with respect to u, and we can conclude that this equation holds if

$$\frac{\mathrm{d}}{\mathrm{d}v}_{\restriction u} \left(\left(\Phi_v^X\right)' Y_{\Phi_{-v}^X(p)}\right) = \left(\Phi_u^X\right)'_{p_u} \left(\frac{\mathrm{d}}{\mathrm{d}v}_{\restriction 0} \left(\Phi_v^X\right)' Y_{\Phi_{-v}^X(p_u)}\right) = 0,$$

where $p_u := \Phi_{-u}^X(p)$. Since $(\Phi_u^X)'_{p_u}$ is a linear isomorphism from $\mathrm{T}_{p_u}M$ to T_pM, the latter equation holds if

$$\frac{\mathrm{d}}{\mathrm{d}v}_{\restriction 0} \left(\left(\Phi_v^X\right)' Y_{\Phi_{-v}^X(q)}\right) = 0, \tag{3.2.20}$$

for all $q \in M$. Applying the left hand side to $f \in C^\infty(M)$ and using that the evaluation of a tangent vector at q on f is a linear mapping from T_qM to $\mathbb{R}$, as well as the product rule (2.2.8) and Formula (3.2.12), one obtains

$$\left(\frac{\mathrm{d}}{\mathrm{d}v}_{\restriction 0} \left(\left(\Phi_v^X\right)' Y_{\Phi_{-v}^X(q)}\right)\right)(f) = \frac{\mathrm{d}}{\mathrm{d}v}_{\restriction 0} \left(\left(\Phi_v^X\right)' Y_{\Phi_{-v}^X(q)}(f)\right) = [Y, X]_q(f).$$

Since, by assumption, $[X, Y] = 0$, this yields (3.2.20) and hence proves the proposition. □

To formulate the last of the basic properties of flows to be discussed here, we need

Definition 3.2.16 (Flow box chart) Let M be an n-dimensional manifold and let X be a vector field on M. A flow box chart for X is a local chart (U, κ) on M such that $X_{\restriction U} = \partial_1$ and $\kappa(U) = (-a, a) \times V$ for some $a > 0$ and some open neighbourhood V of the origin in $\mathbb{R}^{n-1}$.

Proposition 3.2.17 (Straightening Lemma) *Let M be a manifold of dimension n, let $X \in \mathfrak{X}(M)$ and let Φ be its flow.*

1. *A local chart (U, κ) on M with $\kappa(U) = (-a, a) \times V \subset \mathbb{R} \times \mathbb{R}^{n-1}$ is a flow box chart for X iff for all $s \in (-a, a)$, $y \in V$ and $t \in \mathbb{R}$ such that $|s + t| < a$ we have*
$$\left(\kappa \circ \Phi_t \circ \kappa^{-1}\right)(s, y) = (s + t, y). \tag{3.2.21}$$
2. *For every $m \in M$ such that $X_m \neq 0$ there exists a flow box chart for X at m.*

Proof 1. Let κ be such that Eq. (3.2.21) holds. Differentiation at $t = 0$ yields $X^1 = 1$ and $X^i = 0$ for $i = 2, \ldots, n$, so that $X_{\restriction U} = \partial_1$, indeed. Conversely, if κ is such that $X_{\restriction U} = \partial_1$, then the local system (3.2.3) associated with X by (U, κ) is that of Example 3.2.14. This yields the assertion.

2. Since $X_m \neq 0$, there is a local frame $\{X_1, \ldots, X_n\}$ in $\mathrm{T}M$ at m such that $X_1 = X_{\restriction V}$. There exists $a > 0$ such that the mapping

$$\Psi : (-a, a)^n \to M, \qquad \Psi(\mathbf{t}) := \Phi^{X_1}_{t_1} \circ \cdots \circ \Phi^{X_n}_{t_n}(m)$$

is well-defined. Since $\Psi'_{(0,\ldots,0)}$ coincides with the vector space isomorphism $\mathbb{R}^n \to \mathrm{T}_m M$ induced by the basis $\{X_{1,m}, \ldots, X_{n,m}\}$, according to the Inverse Mapping Theorem, a can be chosen so that $U := \Psi((-a, a)^n)$ is open in M and Ψ is a diffeomorphism onto U. Then, (U, Ψ^{-1}) is a local chart on M at m. Since $X_{\restriction U} = X_{1\,\restriction U}$, for all $s \in (-a, a)$, $\mathbf{t} \in (-a, a)^{n-1}$ and $t \in \mathbb{R}$ such that $|t + s| < a$, we obtain

$$\left(\Psi^{-1} \circ \Phi_t \circ \Psi\right)(s, \mathbf{t}) = \Psi^{-1}\left(\Phi^{X_1}_{s+t} \circ \Phi^{X_2}_{t_1} \circ \cdots \circ \Phi^{X_n}_{t_{n-1}}(m)\right) = (s + t, \mathbf{t}).$$

Hence, by point 1, (U, Ψ^{-1}) is a flow box chart. □

Exercises

3.2.1 Along the lines of Example 3.2.8, determine the flow of the linear vector field on $\mathbb{R}^2$ given by the matrix

$$A = \begin{bmatrix} 0 & \frac{1}{\mu} \\ D & 0 \end{bmatrix}, \qquad \mu, D > 0.$$

Find a mechanical system whose time evolution in phase space is given by this flow.

3.2.2 Determine e^{tA} for a real $n \times n$ matrix A of Jordan normal form, see Example 3.2.8.

3.2.3 Complete the proof of Proposition 3.2.10.
Hint. For proving assertion 2, use the defining property 2 of flows and the fact that $\mathscr{D}_m$ is connected.

3.2.4 Determine the flow of the vector field $X = x^i \partial_i$ on $\mathbb{R}^n$ and give a geometric interpretation.

3.2.5 Consider the vector fields $X_1 = -x^2\partial_1 + x^1\partial_2$ and $X_2 = x^1\partial_1 + x^2\partial_2$ on $\mathbb{R}^2$. Show that X_1 and X_2 commute. Check that X_1 is invariant under the transport by the flow of X_2 and vice versa. Relate X_1 and X_2 to polar coordinates, see Example 1.1.11.

3.2.6 Determine the flows of the vector fields

$$X_1 = x^2\partial_3 - x^3\partial_2, \qquad X_2 = x^3\partial_1 - x^1\partial_3, \qquad X_3 = x^1\partial_2 - x^2\partial_1$$

on $\mathbb{R}^3$ and find a geometric interpretation.

3.2.7 Construct flow box charts for the vector fields of the previous exercises.

3.3 The Lie Derivative

The Lie derivative with respect to a vector field X on M extends the directional derivative along the integral curves of X from functions to tensor fields T on M. The idea behind is to compare the value of $T \in \Gamma(\mathbb{T}^q_p M)$ at a given point m with its values along the integral curve through m, transported back to m by the flow $\Phi : \mathscr{D} \to M$ of X and to pass to the differential quotient. More precisely, for every $(t, m) \in \mathscr{D}$, one can apply the transport operator associated with the diffeomorphism $\Phi_{-t} : \mathscr{D}_{-t} \to \mathscr{D}_t$ to the tensor field $T_{\restriction \mathscr{D}_{-t}}$ induced on the open subset $\mathscr{D}_{-t}$ and evaluate the tensor field on $\mathscr{D}_t$ so obtained at m. Thus, we obtain a mapping[9]

$$\mathscr{D} \to \mathbb{T}^q_p M, \qquad (t, m) \mapsto (\Phi_{-t*} T)_m. \tag{3.3.1}$$

This mapping is smooth: for the variable m, this is obvious. To show smoothness in t, due to (2.5.7) and (2.5.8), it suffices to consider the cases where T is a function or a vector field. The case of a function is obvious. The case of a vector field was explained in the proof of Proposition 3.2.15, cf. Eq. (3.2.19). Thus, we can define a mapping

$$\mathscr{L}_X T : M \to \mathbb{T}^q_p M, \qquad (\mathscr{L}_X T)_m := \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} (\Phi_{-t*} T)_m. \tag{3.3.2}$$

Since $\mathscr{L}_X T$ is given by the partial derivative[10] of (3.3.1) with respect to the first variable, evaluated on the vector field $(\frac{\mathrm{d}}{\mathrm{d}t}, 0)$ and restricted to the submanifold $\{0\} \times M$, it is smooth. Thus, $\mathscr{L}_X T$ is a tensor field of the same type as T.

Definition 3.3.1 The tensor field $\mathscr{L}_X T$ is called the Lie derivative of T with respect to X.

According to (2.5.6),

$$(\mathscr{L}_X T)_m = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} \left(\left((\Phi_{-t})' \right)^{\otimes} T_{\Phi_t(m)} \right) = \lim_{t \to 0} \frac{((\Phi_{-t})')^{\otimes} T_{\Phi_t(m)} - T_m}{t}.$$

This justifies the intuitive picture given in the beginning. For further use, we note that

$$\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction t} \left((\Phi_{-t})_* T \right)_m = \left((\Phi_{-t})_* (\mathscr{L}_X T) \right)_m, \tag{3.3.3}$$

see Exercise 3.3.1. For functions and vector fields, the Lie derivative can be expressed in terms of operations we know already.

Proposition 3.3.2 *For a vector field X on M, one has*

1. $\mathscr{L}_X f = X(f)$ *for all* $f \in C^\infty(M)$,
2. $\mathscr{L}_X Y = [X, Y]$ *for all* $Y \in \mathfrak{X}(M)$.

---

[9] We suppress the restriction of T to $\mathscr{D}_{-t}$.

[10] See Remark 2.2.10.

Proof 1. This is obvious.

2. According to Proposition 3.1.1, it suffices to show that $(\mathscr{L}_X Y)(f) = [X,Y](f)$ for all smooth functions f on M. Using (2.5.6) and the fact that application of tangent vectors at m to f is a linear mapping from $\mathrm{T}_m M$ to $\mathbb{R}$, we calculate

$$(\mathscr{L}_X Y)_m(f) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \Big(\big((\Phi^X_{-t})' Y_{\Phi^X_t(m)}\big)(f)\Big) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \frac{\mathrm{d}}{\mathrm{d}s}\restriction_0 \big(f \circ \Phi^X_{-t} \circ \Phi^Y_s \circ \Phi^X_t(m)\big).$$

By the product rule (2.2.8) and by (3.2.12), the right hand side equals[11]

$$\frac{\mathrm{d}}{\mathrm{d}s}\restriction_0 \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 f\big(\Phi^Y_s\big(\Phi^X_t(m)\big)\big) - \frac{\mathrm{d}}{\mathrm{d}s}\restriction_0 \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 f\big(\Phi^X_t\big(\Phi^Y_s(m)\big)\big) = [X,Y]_m(f). \qquad \square$$

The next proposition collects the basic properties of the Lie derivative.

Proposition 3.3.3 *Let X be a vector field on M.*

1. $\mathscr{L}_X(T \otimes S) = (\mathscr{L}_X T) \otimes S + T \otimes (\mathscr{L}_X S)$ *for all tensor fields T, S.*
2. $\mathscr{L}_X(\alpha \wedge \beta) = (\mathscr{L}_X \alpha) \wedge \beta + \alpha \wedge (\mathscr{L}_X \beta)$ *for all $\alpha, \beta \in \Omega^*(M)$.*
3. $\mathscr{L}_X \circ C = C \circ \mathscr{L}_X$ *for every contraction operator C of tensor fields.*
4. $[\mathscr{L}_X, \mathscr{L}_Y] = \mathscr{L}_{[X,Y]}$ *for all $Y \in \mathfrak{X}(M)$.*
5. $\varphi_* \circ \mathscr{L}_X = \mathscr{L}_{\varphi_* X} \circ \varphi_*$ *for every diffeomorphism $\varphi : M \to N$.*

The first two assertions state that the Lie derivative establishes a derivation of the algebra of tensor fields and of the exterior algebra of differential forms, respectively.

Proof 1. By (2.5.8) and the product rule (2.2.6),

$$\begin{aligned}\big(\mathscr{L}_X(T \otimes S)\big)_m &= \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \big((\Phi_{-t*}T)_m \otimes (\Phi_{-t*}S)_m\big) \\ &= \left(\frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 (\Phi_{-t*}T)_m\right) \otimes S_m + T_m \otimes \left(\frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 (\Phi_{-t*}S)_m\right) \\ &= \big((\mathscr{L}_X T) \otimes S + T \otimes (\mathscr{L}_X S)\big)_m .\end{aligned}$$

A similar calculation yields point 2.

3. Since $\mathscr{L}_X$ and C are linear and local operations, that is, $(\mathscr{L}_X T)_m$ and $C(T)_m$ depend on the values of T in an arbitrarily small neighbourhood of m only, it suffices to prove the assertion for $T = \alpha_1 \otimes \cdots \otimes \alpha_p \otimes Y_1 \otimes \cdots \otimes Y_q$ with $\alpha_i \in \Omega^1(M)$ and $Y_i \in \mathfrak{X}(M)$, cf. Remark 2.5.2. Then, in view of point 1, it suffices to show that $\mathscr{L}_X\langle \alpha, Y\rangle = \langle \mathscr{L}_X \alpha, Y\rangle + \langle \alpha, \mathscr{L}_X Y\rangle$ for all $Y \in \mathfrak{X}(M)$ and $\alpha \in \Omega^1(M)$. Indeed, by (2.5.7) and the product rule,

$$\mathscr{L}_X\langle \alpha, Y\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \big(\Phi_{-t*}\langle \alpha, Y\rangle\big) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \langle \Phi_{-t*}\alpha, \Phi_{-t*}Y\rangle = \langle \mathscr{L}_X \alpha, Y\rangle + \langle \alpha, \mathscr{L}_X Y\rangle.$$

[11] This is the same calculation as in the last step of the proof of Proposition 3.2.15.

4. Again, it suffices to prove the assertion for tensor fields of the form $\alpha_1 \otimes \cdots \otimes \alpha_p \otimes Y_1 \otimes \cdots \otimes Y_q$ with $\alpha_i \in \Omega^1(M)$ and $Y_i \in \mathfrak{X}(M)$. First, we show that if the assertion holds for T and S, then it holds for $T \otimes S$. Since $\mathscr{L}_X$ and $\mathscr{L}_Y$ are derivations of the algebra of tensor fields and since the commutator of derivations is a derivation, we have

$$[\mathscr{L}_X, \mathscr{L}_Y](T \otimes S) = \big([\mathscr{L}_X, \mathscr{L}_Y]T\big) \otimes S + T \otimes \big([\mathscr{L}_X, \mathscr{L}_Y]S\big).$$

Since by assumption the assertion holds for T and S individually, and since $\mathscr{L}_{[X,Y]}$ is a derivation, the right hand side equals $\mathscr{L}_{[X,Y]}(T \otimes S)$, indeed. Thus, it suffices to prove the assertion for vector fields and 1-forms. For vector fields it follows from Proposition 3.3.2/2 and from the Jacobi identity for the commutator. To prove the assertion for 1-forms, we first observe that, due to Proposition 3.3.2/1, it holds for functions. Hence, for $\alpha \in \Omega^1(M)$ and $Z \in \mathfrak{X}(M)$, we have $[\mathscr{L}_X, \mathscr{L}_Y]\langle\alpha, Z\rangle = \mathscr{L}_{[X,Y]}\langle\alpha, Z\rangle$. Using point 3, the right hand side can be rewritten as $\langle\mathscr{L}_{[X,Y]}\alpha, Z\rangle + \langle\alpha, \mathscr{L}_{[X,Y]}Z\rangle$ and the left hand side takes the form $\langle[\mathscr{L}_X, \mathscr{L}_Y]\alpha, Z\rangle + \langle\alpha, [\mathscr{L}_X, \mathscr{L}_Y]Z\rangle$. Since the assertion holds for Z, it then follows for α.

5. Using Proposition 3.2.13/2 and (2.5.6), for $T \in \Gamma(\mathbb{T}^r_s M)$ and $p \in N$ one obtains

$$\big(\mathscr{L}_{\varphi_* X}(\varphi_* T)\big)_p = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \big(\varphi_*\big((\Phi_{-t})_* T\big)\big)_p = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \big((\varphi')^{\otimes}\big((\Phi_{-t})_* T\big)_{\varphi^{-1}(p)}\big).$$

Here, $(\varphi')^{\otimes}$ stands for the linear mapping $(\mathbb{T}^r_s)_{\varphi^{-1}(p)} M \to (\mathbb{T}^r_s)_p N$. Hence, the right hand side equals $(\varphi')^{\otimes}(\frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 (\Phi_{-t*}T)_{\varphi^{-1}(p)}) = (\varphi_*(\mathscr{L}_X T))_p$. □

Exercises

3.3.1 Use (2.5.6) to prove Formula (3.3.3).

3.4 Time-Dependent Vector Fields

Let M be a manifold and let $J \subset \mathbb{R}$ be an open interval.

Definition 3.4.1 (Time-dependent vector field) A time-dependent vector field on M is a smooth mapping $X: J \times M \to \mathrm{T}M$ such that $X(t, m) \in \mathrm{T}_m M$ for all $(t, m) \in J \times M$. A smooth curve $\gamma : I \to M$ is an integral curve of X if $I \subset J$ and

$$\dot\gamma(t) = X\big(t, \gamma(t)\big) \quad \text{for all } t \in I. \tag{3.4.1}$$

The properties of maximality and completeness of an integral curve and of completeness of a time-dependent vector field are defined analogously as for ordinary vector fields. By definition, for every $t \in J$, the induced mapping $X_t : M \to \mathrm{T}M$, $X_t(m) := X(t, m)$, is a vector field on M. It is, therefore, common to denote time-dependent vector fields by X_t, with $t \in J$ understood. Note that the parameter of an integral curve coincides with the parameter of the family. This means that while a point moves along an integral curve, the vector field which this curve is tangent to changes synchronously.

In the sequel, for simplicity we assume $J=\mathbb{R}$. To determine the flow of X one decouples the parameters of the curve and of the vector field by passing to the ordinary vector field $\bar{X}$ on $\mathbb{R}\times M$ defined by

$$\bar{X}_{(t,m)} := \big((t,1), X(t,m)\big), \quad t\in\mathbb{R},\ m\in M, \tag{3.4.2}$$

where we have used the natural identification $\mathrm{T}(\mathbb{R}\times M)\cong(\mathbb{R}\times\mathbb{R})\times\mathrm{T}M$. Let $\bar{\Phi}:\bar{\mathscr{D}}\to\mathbb{R}\times M$ denote the flow of $\bar{X}$ and let $\mathrm{pr}_{\mathbb{R}}:\mathbb{R}\times M\to\mathbb{R}$ and $\mathrm{pr}_M:\mathbb{R}\times M\to M$ denote the natural projections.

Proposition 3.4.2 *For every $m\in M$ and $t_0\in\mathbb{R}$, there exists a unique maximal integral curve $\gamma:I\to M$ of X with initial condition $\gamma(t_0)=m$. It is given by*

$$I=\bar{\mathscr{D}}_{(t_0,m)}+t_0, \qquad \gamma(t)=\mathrm{pr}_M\circ\bar{\Phi}_{(t-t_0)}(t_0,m). \tag{3.4.3}$$

Proof For a smooth curve $\bar{\gamma}:\bar{I}\to\mathbb{R}\times M$, define $\bar{\gamma}^{\mathbb{R}}:\bar{I}+t_0\to\mathbb{R}$ and $\bar{\gamma}^{M}:\bar{I}+t_0\to M$ by

$$\bar{\gamma}^{\mathbb{R}}(t):=\mathrm{pr}_{\mathbb{R}}\circ\bar{\gamma}(t-t_0), \qquad \bar{\gamma}^{M}(t):=\mathrm{pr}_{M}\circ\bar{\gamma}(t-t_0).$$

The curve $\bar{\gamma}$ is an integral curve of $\bar{X}$ through (t_0,m) iff

$$\dot{\bar{\gamma}}^{\mathbb{R}}(t)=1, \qquad \dot{\bar{\gamma}}^{M}(t)=X\big(\bar{\gamma}^{\mathbb{R}}(t),\bar{\gamma}^{M}(t)\big)$$

with initial conditions $\bar{\gamma}^{\mathbb{R}}(t_0)=t_0$ and $\bar{\gamma}^{M}(t_0)=m$, respectively. The equations for $\bar{\gamma}^{\mathbb{R}}$ are solved by $\bar{\gamma}^{\mathbb{R}}(t)=t$ for all $t\in\bar{I}+t_0$. The above system is then equivalent to

$$\dot{\bar{\gamma}}^{M}(t)=X\big(t,\bar{\gamma}^{M}(t)\big), \qquad \bar{\gamma}^{M}(t_0)=m.$$

First, this shows that the curve γ defined by (3.4.3) is an integral curve of X with the appropriate initial condition. Second, this shows that if $\gamma_1:I_1\to M$ is an integral curve of X with initial condition $\gamma_1(t_0)=m$, the curve $\bar{\gamma}_1:I_1+t_0\to\mathbb{R}\times M$ defined by $\bar{\gamma}_1(t):=(t+t_0,\gamma(t+t_0))$ is an integral curve of $\bar{X}$ through (t_0,m). Therefore, γ is maximal and unique. □

Like the maximal integral curves of an ordinary vector field combine to an ordinary flow, the maximal integral curves of a time-dependent vector field will combine to a time-dependent flow. For a set A, let $\Delta_A\subset A\times A$ denote the diagonal.

Definition 3.4.3 (Time-dependent flow) Let $\mathscr{D}$ be an open neighbourhood of $\Delta_{\mathbb{R}}\times M$ in $\mathbb{R}\times\mathbb{R}\times M$ and let $\Phi:\mathscr{D}\to M$ be a smooth mapping. For $m\in M$ and $t_1,t_2\in\mathbb{R}$, denote

$$\mathscr{D}_{t_1,m}:=\big\{t_2\in\mathbb{R}:(t_2,t_1,m)\in\mathscr{D}\big\}, \qquad \mathscr{D}_{t_2,t_1}:=\big\{m\in M:(t_2,t_1,m)\in\mathscr{D}\big\}$$

and let $\Phi_{t_2,t_1}:\mathscr{D}_{t_2,t_1}\to M$ denote the induced mapping, given by $\Phi_{t_2,t_1}(m)=\Phi(t_2,t_1,m)$. The mapping Φ is called a time-dependent flow on M if for all $t_1,t_2,t_3\in\mathbb{R}$ and $m\in M$ the following holds:

1. $\Phi_{t_1,t_1} = \mathrm{id}_M$,
2. if (t_2, t_1, m) and $(t_3, t_2, \Phi_{t_2,t_1}(m)) \in \mathscr{D}$, then $(t_3, t_1, m) \in \mathscr{D}$ and
$$\Phi_{t_3,t_2}\big(\Phi_{t_2,t_1}(m)\big) = \Phi_{t_3,t_1}(m),$$
3. $\mathscr{D}_{t_1,m}$ is connected,
4. $\Phi_{t_2,t_1}(\mathscr{D}_{t_2,t_1}) \subset \mathscr{D}_{t_1,t_2}$.

A time-dependent flow is called maximal if it does not admit an extension, in the sense of mappings, which itself is a time-dependent flow. A time-dependent flow is called complete if $\mathscr{D} = \mathbb{R} \times \mathbb{R} \times M$.

Now, for a time-dependent vector field X, define

$$\mathscr{D} := \big\{(t, t_0, m) \in \mathbb{R} \times \mathbb{R} \times M : \big(t - t_0, (t_0, m)\big) \in \bar{\mathscr{D}}\big\} \tag{3.4.4}$$

and

$$\Phi : \mathscr{D} \to M, \quad \Phi_{t,t_0}(m) := \mathrm{pr}_M \circ \bar{\Phi}_{t-t_0}\big((t_0, m)\big). \tag{3.4.5}$$

Using Proposition 3.4.2 and the fact that $\bar{\mathscr{D}}$ is an ordinary flow, one can prove the following (Exercise 3.4.1).

Proposition 3.4.4 *The assignment of Φ to X given by* (3.4.4) *and* (3.4.5) *defines a bijection between time-dependent vector fields on M and maximal time-dependent flows on M. Complete time-dependent vector fields thereby correspond to complete time-dependent flows.*

Remark 3.4.5

1. Let $\Phi : \mathscr{D} \to M$ be a time-dependent flow on M. Then, for all $t_1, t_2 \in \mathbb{R}$ such that $\mathscr{D}_{t_2,t_1} \neq \emptyset$, we have $\Phi_{t_2,t_1}(\mathscr{D}_{t_2,t_1}) = \mathscr{D}_{t_1,t_2}$ and $\Phi_{t_2,t_1} : \mathscr{D}_{t_2,t_1} \to \mathscr{D}_{t_1,t_2}$ is a diffeomorphism with inverse $\Phi_{t_1,t_2} : \mathscr{D}_{t_1,t_2} \to \mathscr{D}_{t_2,t_1}$. Moreover, connectedness of $\mathscr{D}_{t_1,m}$ implies that for every t between t_1 and t_2 one has $\mathscr{D}_{t_2,t_1} \subset \mathscr{D}_{t,t_1}$ and $\Phi_{t,t_1}(\mathscr{D}_{t_2,t_1}) \subset \mathscr{D}_{t_2,t}$. The proof of these two statements is left to the reader (Exercise 3.4.2).
2. Time-dependent flows are in bijective correspondence with smooth 1-parameter families $\{\Phi_t\}$ of local diffeomorphisms of M: given Φ_{t_2,t_1}, one defines $\Phi_t := \Phi_{t,0}$. Then, $\Phi_{t_2,t_1} = \Phi_{t_2,0} \circ \Phi_{0,t_1} = \Phi_{t_2,0} \circ \Phi_{t_1,0}^{-1}$ and hence
$$\Phi_{t_2,t_1} = \Phi_{t_2} \circ \Phi_{t_1}^{-1}.$$
Conversely, given $\{\Phi_t\}$, this equation defines local diffeomorphisms Φ_{t_2,t_1} for all t_1 and t_2. We leave it to the reader to show that this family forms a time-dependent flow on M.
3. If a time-dependent flow Φ is complete, all the mappings Φ_{t_2,t_1} are diffeomorphisms of M and Definition 3.4.3 reduces to the requirement that Φ is smooth and satisfies
$$\Phi_{t_3,t_2} \circ \Phi_{t_2,t_1} = \Phi_{t_3,t_1} \quad \text{for all } t_1, t_2, t_3 \in \mathbb{R}, \qquad \Phi_{t,t} = \mathrm{id}_M \quad \text{for all } t \in \mathbb{R}.$$

These are the defining properties of an action on M of the groupoid $\mathbb{R} \times \mathbb{R}$ with multiplication $(a, b) \cdot (b, d) = (a, d)$ and inverse $(a, b)^{-1} = (b, a)$. Therefore, a complete time-dependent flow may also be referred to as a two-parameter groupoid of diffeomorphisms.

4. Ordinary flows embed into time-dependent flows as follows. If Φ° is an ordinary flow on M with domain $\mathscr{D}^\circ$, by defining

$$\mathscr{D} := \left\{(t_2, t_1, m) \in \mathbb{R} \times \mathbb{R} \times M : t_2 - t_1 \in \mathscr{D}^\circ_m\right\}, \qquad \Phi_{t_2,t_1} := \Phi^\circ_{t_2-t_1},$$

we obtain a time-dependent flow on M. The latter has the property that $\mathscr{D}_{t_2,t_1}$ and Φ_{t_2,t_1} depend on the difference $t_2 - t_1$ only. Conversely, if $\Phi : \mathscr{D} \to M$ is a time-dependent flow on M with this property, then

$$\mathscr{D}^\circ := \left\{(t, m) \in \mathbb{R} \times M : (t, 0, m) \in \mathscr{D}\right\}, \qquad \Phi^\circ_t := \Phi_{t,0}$$

yields an ordinary flow on M. The proof of both assertions is left to the reader (Exercise 3.4.3).

To conclude this section, we briefly discuss the special case of periodically time-dependent vector fields.

Definition 3.4.6 A time-dependent vector field X on M is said to be periodic if there exists a minimal $T > 0$, called the period, such that for all $t \in \mathbb{R}$ and $m \in M$ one has

$$X(t + T, m) = X(t, m).$$

Proposition 3.4.7 *The flow $\Phi : \mathscr{D} \to M$ of a periodically time-dependent vector field with period T satisfies*

1. *$\mathscr{D}_{t_2+T,t_1+T} = \mathscr{D}_{t_2,t_1}$ and $\Phi_{t_2+T,t_1+T} = \Phi_{t_2,t_1}$ for all $t_1, t_2 \in \mathbb{R}$,*
2. *$\mathscr{D}_{kT,0}$ is nonempty and $\Phi_{kT,0} = \Phi^k_{T,0}$ for all $k \in \mathbb{Z}$.*

Point 2 includes the statement that the domain of $\Phi^k_{T,0}$ coincides with $\mathscr{D}_{kT,0}$.

Proof Let $\bar{X}$ be the vector field on $\mathbb{R} \times M$ defined by (3.4.2) and let $\bar{\Phi} : \bar{\mathscr{D}} \to \mathbb{R} \times M$ be its flow. Then, $\mathscr{D}$ and $\bar{\mathscr{D}}$ are related by (3.4.4) and Φ and $\bar{\Phi}$ are related by (3.4.5).

1. Consider the diffeomorphism

$$\varphi_T : \mathbb{R} \times M \to \mathbb{R} \times M, \qquad \varphi_T(t, m) := (t + T, m).$$

A brief calculation shows that $\varphi_{T*}\bar{X} = \bar{X}$. According to Proposition 3.2.13/3, then $\varphi_T(\bar{\mathscr{D}}_t) = \bar{\mathscr{D}}_t$ and $\bar{\Phi}_t \circ \varphi_T = \varphi_T \circ \bar{\Phi}_t$. First, in view of (3.4.4), this implies $\mathscr{D}_{t_2,t_1} = \mathscr{D}_{t_2+T,t_1+T}$. Second, for every $m \in \mathscr{D}_{t_2,t_1}$, (3.4.5) implies

$$\Phi_{t_2+T,t_1+T}(m) = \mathrm{pr}_M \circ \bar{\Phi}_{t_2-t_1} \circ \varphi_T(t_1, m) = \mathrm{pr}_M \circ \varphi_T \circ \bar{\Phi}_{t_2-t_1}(t_1, m) = \Phi_{t_2,t_1}(m).$$

2. We prove this by induction on k. For $k = 1$, the assertion is obvious. Thus, assume that it holds for k. According to Remark 3.4.5/1 and the first equality in assertion 1,

$$\Phi^k_{T,0}(\mathscr{D}_{(k+1)T,0}) = \Phi_{kT,0}(\mathscr{D}_{(k+1)T,0}) \subset \mathscr{D}_{(k+1)T,kT} = \mathscr{D}_{T,0}.$$

It follows that $\Phi_{T,0}$ can be applied to $\Phi^k_{T,0}(\mathscr{D}_{(k+1)T,0})$, so that $\Phi^{k+1}_{T,0}$ is defined on $\mathscr{D}_{(k+1)T,0}$. Conversely, if $m \in M$ belongs to the domain of $\Phi^{k+1}_{T,0}$, then it belongs to the domain of $\Phi^k_{T,0}$ and $\Phi^k_{T,0}(m)$ belongs to $\mathscr{D}_{(k+1)T,kT}$. Since by the induction assumption, the domain of $\Phi^k_{T,0}$ coincides with $\mathscr{D}_{kT,0}$, it follows that $m \in \mathscr{D}_{(k+1)T,0}$. Finally, using the second equality in assertion 1, on $\mathscr{D}_{(k+1)T,0}$ we calculate

$$\Phi^{k+1}_{T,0} = \Phi_{T,0} \circ \Phi^k_{T,0} = \Phi_{(k+1)T,kT} \circ \Phi_{kT,0} = \Phi_{(k+1)T,0}. \qquad \square$$

Exercises

3.4.1 Use Proposition 3.4.2 and the fact that $\bar{\mathscr{D}}$ is an ordinary flow to prove Proposition 3.4.4.

3.4.2 Prove the assertions of Remark 3.4.5/1.

3.4.3 Show that ordinary flows may be viewed as time-dependent flows, cf. Remark 3.4.5/4.

3.4.4 Let $M = \mathbb{R}$. Determine the maximal integral curves and the time-dependent flow of the time-dependent vector field X on M given by $X(t,x) = (x, \frac{x}{1-t})$, $t \in (-1,1)$.

3.5 Distributions and Foliations

Let M be a manifold. For a subset D of $\mathrm{T}M$, let $\mathfrak{X}^D(M)$ denote the set of vector fields on M taking values in D. Correspondingly, let $\mathfrak{X}^D_{\mathrm{loc}}(M)$ denote the set of local vector fields on M taking values in D.

Definition 3.5.1 (Distribution) A distribution[12] on M is a subset D of $\mathrm{T}M$ such that for all $m \in M$ the following holds.

1. $D_m := D \cap \mathrm{T}_m M$ is a linear subspace of $\mathrm{T}_m M$.
2. For every $Y \in D_m$, there exists $X \in \mathfrak{X}^D(M)$ such that $X_m = Y$.

The function which assigns to $m \in M$ the dimension of D_m is called the rank of D. If the rank is constant, D is called regular. Otherwise, it is called singular.

Concerning distributions and regularity we follow the terminology used e.g. in [181].[13] Beware that it is quite common to include the constant rank requirement into the definition of distribution and to reserve the notion of regularity for additional properties. Then, distributions which are not necessarily of constant rank are referred to as generalized distributions.

[12] To distinguish this notion from a distribution in the sense of analysis one should speak, more precisely, of a geometric distribution. However, it is common to omit the term geometric.

[13] According to [181], our distributions should be referred to, more precisely, as smooth distributions. Since in this book we will meet only smooth distributions, we systematically omit the adjective smooth.

Remark 3.5.2

1. According to Remark 2.3.2/4, condition 2 of Definition 3.5.1 is equivalent to the requirement that for every $Y \in D_m$ there exists $X \in \mathfrak{X}^D_{\mathrm{loc}}(M)$ such that $X_m = Y$.
2. Let $m \in M$ and $r = \dim D_m$. Condition 2 of Definition 3.5.1 implies that there exists a local r-frame $\{X_1, \ldots, X_r\}$ in $\mathrm{T}M$ at m taking values in D, such that $X_{1,m}, \ldots, X_{r,m}$ span D_m (choose a basis in D_m, apply property 2 and restrict the domains of the vector fields so obtained appropriately). This shows that the rank is locally non-decreasing, that is, every $m \in M$ has a neighbourhood where the rank is greater than or equal to the rank at m. Moreover, if the rank is constant, then $X_{1,m}, \ldots, X_{r,m}$ also span $D_{\tilde{m}}$ for all $\tilde{m} \in U$. Thus, Proposition 2.7.5/2 yields that D is regular iff it is a vertical subbundle of $\mathrm{T}M$. Hence, the definition of regular distribution given in Example 2.7.6 is equivalent to the one given here.

Definition 3.5.3 (Integral manifold) Let D be a distribution on M. A connected submanifold (N, ψ) of M is called an integral manifold[14] of D through $m \in M$ if $m \in \psi(N)$ and

$$\psi'_p(\mathrm{T}_p N) = D_{\psi(p)} \tag{3.5.1}$$

for all $p \in N$. D is said to be integrable if for every $m \in M$ there exists an integral manifold of D through m.

Along an integral manifold, D has constant rank. If N is given by a subset, Formula (3.5.1) reads

$$\mathrm{T}N = D_{\restriction N}. \tag{3.5.2}$$

Example 3.5.4

1. Every vector field X on M generates a distribution D by $D_m := \mathbb{R}X_m$. The rank of D at m is 0 if $X_m = 0$ and 1 otherwise. According to Proposition 3.2.11, the images of the maximal integral curves of X are submanifolds of M. Obviously, they are integral manifolds of D. Thus, D is integrable.
2. Every subset $A \subset \mathfrak{X}(M)$ generates a distribution D, with D_m defined to be the linear span of the set $\{X_m : X \in A\}$, and every distribution can be generated this way. A sufficient, but by far not necessary, condition for D to be regular of rank r is that A be an r-frame.
3. Let E be a vector bundle over a manifold M and let $\Phi\colon E \to \mathrm{T}M$ be a vertical vector bundle morphism. The image of Φ is a distribution, because it is locally spanned by $\Phi \circ s_i$, where $\{s_i\}$ is a local frame in E.
4. Let $\pi\colon M \to P$ be a submersion. According to Example 2.7.7, $\ker \pi'$ is a regular distribution on M of rank $\dim M - \dim P$. According to the Level Set Theorem

[14] In part of the literature, for (N, ψ) to be an integral manifold, only $\psi'_p(\mathrm{T}_p N) \subset D_{\psi(p)}$ is required. What we call integral manifold here is referred to there as an integral manifold which is everywhere of maximal dimension.

for manifolds (Corollary 1.8.3), the connected components of the level sets of π are integral manifolds of $\ker \pi'$. Since they make up M, $\ker \pi'$ is integrable. In fact, the connected components of the level sets of π are maximal integral manifolds, because the composition of a curve in an integral manifold with π is an integral curve of the zero vector field on P and is hence constant.

5. Let $\varepsilon > 0$ and $0 \leq r \leq n$. Consider the open cube $(-\varepsilon, \varepsilon)^n \subset \mathbb{R}^n$. The vector fields $\partial_1, \ldots, \partial_r$ span a regular distribution D of rank r on this cube. This distribution is integrable, with the integral manifolds given by the subsets $(-\varepsilon, \varepsilon)^r \times \{(x_{r+1}, \ldots, x_n)\}$, where $(x_{r+1}, \ldots, x_n) \in (-\varepsilon, \varepsilon)^{n-r}$ is fixed. It will be shown below that, locally, every integrable regular distribution of rank r on an n-dimensional manifold is of this form.
6. The distribution on $\mathbb{R}^2$ spanned by the vector fields ∂_x and $y\partial_y$ is singular, because it has rank 1 on the x-axis and rank 2 outside. Similarly, the distribution spanned by the vector fields ∂_x and $x\partial_y$ is singular, because it has rank 1 on the y-axis and rank 2 outside. The first of these distributions is integrable, with integral manifolds being the x-axis and the two open half-planes. In contrast, the reader can easily see that the second one is not integrable.

Next, we are going to derive criteria for integrability. We need the following notions.

Definition 3.5.5 A distribution D on M is called

1. involutive if $\mathfrak{X}^D(M) \subset \mathfrak{X}(M)$ is a Lie subalgebra,
2. homogeneous[15] if for all $X \in \mathfrak{X}^D_{\mathrm{loc}}(M)$ and all (t, m) in the domain of Φ^X, one has

$$\left(\Phi^X_t\right)'_m D_m = D_{\Phi^X_t(m)}.$$

Note that homogeneity means invariance of D under the flow of an arbitrary local vector field taking values in D.

Remark 3.5.6 Each of the following two conditions is equivalent to involutivity (Exercise 3.5.1).

1. For all $X, Y \in \mathfrak{X}^D_{\mathrm{loc}}(M)$ whose domains have nontrivial intersection, $[X, Y] \in \mathfrak{X}^D_{\mathrm{loc}}(M)$.
2. For all $m_0 \in M$, there exists an open neighbourhood U of m_0 in M and local vector fields $X_1, \ldots, X_r$ on U such that
   (a) $X_{1,m}, \ldots, X_{r,m}$ span D_m for all $m \in U$,
   (b) $[X_i, X_j] = c^k_{ij} X_k$ with smooth functions $c^k_{ij} : U \to \mathbb{R}$.

[15] After Stefan [280].

Example 3.5.7

1. The distribution D generated by a vector field X on M is involutive: let $Y, Z \in \mathfrak{X}^D(M)$. If $\dim D_m = 1$, there exists an open neighbourhood V of m such that $X_{\restriction V}$ does not have zeros. Hence, there exist unique $f, g \in C^\infty(V)$ such that $Y_{\restriction V} = f X_{\restriction V}$ and $Z_{\restriction V} = g X_{\restriction V}$. Then, (3.1.1) and (3.1.7) yield
$$[Y, Z]_m = [f X_{\restriction V}, g X_{\restriction V}]_m = \big(f(m) X_m(g) - g(m) X_m(f)\big) X_m \in D_m.$$
If $\dim D_m = 0$, then $Y_m = Z_m = 0$ and hence
$$[Y, Z]_m(f) = Y_m\big(Z(f)\big) - Z_m\big(Y(f)\big) = 0$$
for all $f \in C^\infty(M)$, so that $[Y, Z]_m = 0$.
2. The distribution on $\mathbb{R}^3$ spanned by the vector fields $\partial_x + y\partial_z$ and $(x^2 + 1)\partial_y$ is regular, but not involutive (Exercise 3.5.2).
3. The (singular) distribution on $\mathbb{R}^2$ spanned by the vector fields ∂_x and $y\partial_y$ is involutive, whereas the distribution spanned by ∂_x and $x\partial_y$ is not.

Definition 3.5.8 (Adapted chart) Let D be a distribution on M, let $m \in M$, and let $r = \dim D_m$, $n = \dim M$. A local chart (U, κ) on M is said to be adapted to D at m if

1. $\kappa(m) = 0$ and $\kappa(U) = (-\varepsilon, \varepsilon)^n$ for some $\varepsilon > 0$,
2. $\partial_1, \dots, \partial_r \in \mathfrak{X}^D_{\mathrm{loc}}(M)$,
3. for all $\mathbf{c} \in (-\varepsilon, \varepsilon)^{n-r}$, D has constant rank along $U_{\mathbf{c}} := \kappa^{-1}((-\varepsilon, \varepsilon)^r \times \{\mathbf{c}\})$.

Remark 3.5.9

1. The subsets $U_{\mathbf{c}}$, $\mathbf{c} \in (-\varepsilon, \varepsilon)^{n-r}$, are embedded submanifolds of M. They are referred to as the slices of U. By condition 2, the rank of D at points of U is greater than or equal to the rank at the central point m, i.e., r. If a slice $U_{\mathbf{c}}$ contains a point where D has the minimal rank r then, by counting dimensions, conditions 2 and 3 imply that $U_{\mathbf{c}}$ is an integral manifold of D. In particular, U_0 is an integral manifold of D through m.
2. If D is regular, condition 3 is redundant. In this case, condition 2 means that $\partial_1, \dots, \partial_r$ span D on U and that every slice $U_{\mathbf{c}}$ is an integral manifold. Thus, in an adapted chart, the integral manifolds are given by the equations
$$x_{r+1} = c_1, \dots, x_n = c_{n-r}.$$
3. In case D is generated by a vector field X and $X_m \neq 0$, every flow box chart for X at m is adapted to D at m.

Theorem 3.5.10 (Stefan and Sussmann) *Let M be a manifold of dimension n and let D be a distribution on M. The following statements are equivalent.*

1. *D is integrable.*
2. *D is involutive and has constant rank along integral curves of elements of $\mathfrak{X}^D_{\mathrm{loc}}(M)$.*

3. *D is homogeneous.*
4. *For every $m \in M$, there exists a local chart adapted to D at m.*

Proof See [280, 283] for the original sources. Our proof follows [181, App. 3].

1 ⇒ 2: Let $X, Y \in \mathfrak{X}^D(M)$. By integrability, every $m \in M$ is contained in some integral manifold (N_m, ψ_m) of D. Due to Corollary 3.1.6, since X and Y are tangent to (N_m, ψ_m), so is their commutator. It follows $[X, Y]_m \in D_m$ for all $m \in M$. Since the integral curves through m of the elements of $\mathfrak{X}^D_{\mathrm{loc}}(M)$ are contained in $\psi_m(N_m)$ and since the rank of D is constant along integral manifolds, the second assertion follows, too.

2 ⇒ 3: Let $X \in \mathfrak{X}^D_{\mathrm{loc}}(M)$ with domain U, let $\Phi : \mathscr{D} \to U$ be the flow of X and let $(t, m) \in \mathscr{D}$. Assume that we are given open neighbourhoods of $\Phi_m(s)$ for all $s \in [0, t]$. Since the preimages under Φ_m of these neighbourhoods provide an open covering of the compact interval $[0, t]$, finitely many of them already cover the integral curve of X from m to $\Phi_t(m)$. This shows that it suffices to prove the assertion for m being arbitrary but fixed, U being an arbitrarily small neighbourhood of m, and all $t \in \mathscr{D}_m$. Let $r = \dim D_m$. We choose U so that there exists a local r-frame $\{Y_1, \dots, Y_r\}$ taking values in D and spanning D_m. Define smooth curves $Z_i : \mathscr{D}_m \to \mathrm{T}_m M$ by

$$Z_i(t) := (\Phi_{-t*} Y_i)_m, \quad i = 1, \dots, r,$$

where by an abuse of notation we have omitted the restriction of Y_i to $\mathscr{D}_{-t}$. By construction,

$$(\Phi_t)' Z_i(t) = (Y_i)_{\Phi_t(m)}, \quad i = 1, \dots, r. \tag{3.5.3}$$

In particular, the $Z_i(t)$ are linearly independent. It suffices to show that $Z_i(t) \in D_m$ for all $t \in \mathscr{D}_m$, because then (3.5.3) implies $(\Phi_t)' D_m \subset D_{\Phi_t(m)}$ and by the assumption on the rank of D there holds equality. According to Formula (3.3.3) and Proposition 3.3.2/2,

$$\dot{Z}_i(t) = (\Phi_{-t*} \mathscr{L}_X Y_i)_m = \big(\Phi_{-t*}[X, Y_i]\big)_m = (\Phi_{-t})'\big([X, Y_i]_{\Phi_t(m)}\big). \tag{3.5.4}$$

Since D is involutive, by Remark 3.5.6, $[X, Y_i]_{\Phi_t(m)} \in D_{\Phi_t(m)}$ for all $t \in \mathscr{D}_m$. Since the rank of D along the integral curve $t \mapsto \Phi_t(m)$ is constant, $(Y_1)_{\Phi_t(m)}, \dots, (Y_r)_{\Phi_t(m)}$ span $D_{\Phi_t(m)}$ for all $t \in \mathscr{D}_m$. Hence there exist unique smooth functions $\lambda^j_i : \mathscr{D}_m \to \mathbb{R}$ such that $[X, Y_i]_{\Phi_t(m)} = \lambda^j_i(t)(Y_j)_{\Phi_t(m)}$ (summation convention). Then, (3.5.4) implies that the curves $Z_i(t)$ satisfy the ordinary differential equation

$$\dot{Z}_i(t) = \lambda^j_i(t) Z_j(t)$$

so that, due to the initial condition $Z_i(0) \in D_m$, they stay in D_m, as asserted.

3 ⇒ 4: The proof generalizes the construction of a flow box chart in the proof of Proposition 3.2.17. Let $m \in M$ and let $r = \dim D_m$. Choose a local r-frame $\{X_1, \dots, X_r\}$ at m taking values in D. By Proposition 2.3.15, in a neighbourhood of m this local r-frame can be complemented to a local frame $\{X_1, \dots, X_n\}$ in $\mathrm{T}M$. There exists $\varepsilon > 0$ such that the mapping

$$\Psi : (-\varepsilon, \varepsilon)^n \to M, \qquad \Psi(\mathbf{t}) := \Phi^{X_1}_{t_1} \circ \cdots \circ \Phi^{X_n}_{t_n}(m)$$

is defined. Since Ψ'_0 amounts to the natural vector space isomorphism $\mathbb{R}^n \to \mathrm{T}_m M$ induced by the basis $\{X_{1,m}, \dots, X_{n,m}\}$, according to the Inverse Mapping Theorem, ε can be chosen so that $U := \Psi((-\varepsilon, \varepsilon)^n)$ is open in M and Ψ is a diffeomorphism onto U. Then $(U, \kappa := \Psi^{-1})$ is a local chart on M at m. It remains to check that (U, κ) satisfies conditions 1–3 of Definition 3.5.8. Condition 1 holds by construction. For condition 2, let $i \in \{1, \dots, r\}$ and $\tilde{m} \in U$. Then, denoting $\Phi_i \equiv \Phi^{X_i}_{\kappa^i(\tilde{m})}$, we obtain

$$\partial^\kappa_{i,\tilde{m}}(f) = \big(\Phi'_1 \circ \cdots \circ \Phi'_{i-1}(X_i)_{\Phi_i \circ \cdots \circ \Phi_n(m)}\big)(f).$$

By homogeneity, the vector on the right hand side belongs to $D_{\tilde{m}}$, and condition 2 holds, indeed. Condition 3 follows from homogeneity and the fact that, by construction, any element of the subset $\kappa^{-1}((-\varepsilon, \varepsilon)^r \times \{\mathbf{x}\})$ can be joined to $\Psi(0, \mathbf{x})$ by a composition of integral curves of $X_1, \dots, X_r$.

4 ⇒ 1: This follows from Remark 3.5.9/1. □

Remark 3.5.11

1. The proofs of the implications 2 ⇒ 3 and 3 ⇒ 4 show that it suffices that the constant rank condition of point 2 or the homogeneity condition of point 3 hold for a neighbourhood of every point and a family of local vector fields spanning D over that neighbourhood.
2. Consider the proof of the implication 3 ⇒ 4. From the formula for $\partial^\kappa_{i,\tilde{m}}$ it follows that $\partial_1 = X_{1\restriction U}$, that is, (U, κ) is a flow box chart for X_1. For the remaining local vector fields X_i, $i = 2, \dots, r$, this is not true in general. However, it is true in the special case where the X_i commute, because then their flows commute, see Proposition 3.2.15. For later use, let us formulate this observation as follows. Let D_m be spanned by the values at m of a local r-frame $\{X_1, \dots, X_r\}$ in $\mathrm{T}M$ over V. If the X_i commute, there exists a local chart (U, κ) adapted to D at m such that $U \subset V$ and $X_{i\restriction U} = \partial_i$, $i = 1, \dots, r$.

Corollary 3.5.12 (Frobenius Theorem) *Let M be a manifold and let D be a regular distribution on M. If D is involutive, then for every $m \in M$, there exists a local chart adapted to D at m. In particular, D is integrable iff it is involutive.*

Example 3.5.13

1. The distributions discussed in points 1, 3 and 4 of Example 3.5.4 are involutive and homogeneous, because they are integrable. The distribution on $\mathbb{R}^3$ spanned by the vector fields $\partial_x + y\partial_z$ and $(x^2 + 1)\partial_y$ is neither integrable nor homogeneous, because it is not involutive, see Example 3.5.7/2. For the first of the distributions of Example 3.5.4/6, Theorem 3.5.10 yields homogeneity and confirms integrability. On the other hand, the second one is neither homogeneous nor integrable.

2. In view of Remark 3.5.11/1, point 2 of Theorem 3.5.10 yields that a distribution D spanned by commuting vector fields $X_1, \ldots, X_r$ is always integrable. Indeed, since the flows commute, Proposition 3.2.13/2 yields

$$X_i\big(\Phi_t^{X_j}(m)\big) = \big(\Phi_t^{X_j}\big)'\big(X_i(m)\big)$$

for all i and j, hence D has constant rank along the integral curves of the X_i.

Next, we prove that integral manifolds of integrable distributions are initial and that there exist maximal integral manifolds. Without loss of generality, for convenience we may assume any integral manifold (N, φ) to be given by its image $\varphi(N) \subset M$, endowed with the topology and differentiable structure induced from N via φ.

Lemma 3.5.14 *Let D be a distribution on M.*

1. *Let $N_1 \subset M$ and $N_2 \subset M$ be integral manifolds of D. If $N_1 \cap N_2 \neq \varnothing$, then $N_1 \cap N_2$ is open in N_1 and N_2 and the smooth structures on $N_1 \cap N_2$ induced from N_1 and N_2 coincide. There is a unique smooth structure on $N_1 \cup N_2$ such that N_1 and N_2 are open submanifolds. With respect to this structure, $N_1 \cup N_2$ is an integral manifold of D.*
2. *Let $N \subset M$ be an integral manifold of D, let $m \in M$ such that* $\dim D_m = \dim N$ *and let (U, κ) be a local chart adapted to D at $m \in M$. If a slice $U_{\mathbf{c}}$ of U intersects N, then $U_{\mathbf{c}}$ is an integral manifold of D and $N \cap U_{\mathbf{c}}$ is open and closed*[16] *in $N \cap U$ with respect to the relative topology induced from N. In particular, the number of slices of U which intersect N is at most countable.*

The number of slices of U intersected by N in assertion 2 may happen to be (countably) infinite, indeed: for $M = \mathrm{S}^1 \times \mathrm{S}^1$ with angle coordinates ϕ_1 and ϕ_2 and D generated by the vector field $X = \partial_{\phi_1} + \omega \partial_{\phi_2}$ with ω irrational, the intersection of any open subset with a maximal integral curve of X has infinitely many slices.

Proof 1. The assertion on $N_1 \cup N_2$ follows from that on $N_1 \cap N_2$, hence it suffices to prove the latter. The dimensions of N_1 and N_2 coincide with the rank of D on $N_1 \cap N_2$ and hence are equal. Denote this number by r. For $m \in N_1 \cap N_2$, choose a local r-frame $\{X_1, \ldots, X_r\}$ in $\mathrm{T}M$ over some neighbourhood U of m taking values in D. Since X_j is tangent to N_i, by restriction it induces a local vector field $X_j^{(i)}$ on $U \cap N_i$, $j = 1, \ldots, r$, $i = 1, 2$. By the same argument as in the proof of the implication $3 \Rightarrow 4$ of Theorem 3.5.10, there exists $\varepsilon > 0$ such that the mappings

$$\Psi^{(i)} : (-\varepsilon, \varepsilon)^r \to U \cap N_i, \qquad \Psi^{(i)}(\mathbf{t}) := \Phi_{t_1}^{X_1^{(i)}} \circ \cdots \circ \Phi_{t_r}^{X_r^{(i)}}(m)$$

are diffeomorphisms onto open subsets V_i of N_i, $i = 1, 2$. By construction, $\Psi^{(1)}(\mathbf{t}) = \Psi^{(2)}(\mathbf{t})$ as elements of M for all $\mathbf{t} \in (-\varepsilon, \varepsilon)^r$, cf. Remark 3.2.9/2. First,

[16] It is in fact a connected component of $N \cap U$. We do not need this though.

this implies $V_1 = V_2 =: V$ and hence $V \subset N_1 \cap N_2$, where V is open in both N_1 and N_2. Second, this implies that with respect to the differentiable structure induced from either N_i, V is diffeomorphic to $(-\varepsilon, \varepsilon)^r$. Since the construction works for every $m \in N_1 \cap N_2$, the assertion follows.

2. Let $r = \dim N \equiv \dim D_m$. If N intersects a slice $U_{\mathbf{c}}$, then $U_{\mathbf{c}}$ contains a point where D has rank r. By Remark 3.5.9/1, then $U_{\mathbf{c}}$ is an integral manifold of D. Now, assertion 1 implies that $N \cap U_{\mathbf{c}}$ is open in N and hence in $N \cap U$, where the topology of $N \cap U$ is assumed to be induced from N. Since this holds for all slices which intersect N and since

$$N \cap U_{\mathbf{c}} = (N \cap U) \Big\backslash \Big(\bigcup_{\tilde{\mathbf{c}} \neq \mathbf{c}} N \cap U_{\tilde{\mathbf{c}}} \Big),$$

the intersection $N \cap U_{\mathbf{c}}$ is also closed in $N \cap U$. It follows that $N \cap U_{\mathbf{c}}$ is a union of connected components of $N \cap U$. Since the slices are disjoint, this implies that the number of slices which intersect N is at most as large as the number of connected components of $N \cap U$. Since the topology of $N \cap U$ is induced from the manifold N, it is second countable, hence the number of connected components is at most countable. □

Proposition 3.5.15 *Integral manifolds of integrable distributions are initial submanifolds.*

Proof Let M be a manifold, let D be an integrable distribution and let N be an integral manifold of D. To show that N is initial, we must prove that, for every smooth manifold P and every smooth mapping $\varphi : P \to M$ such that $\varphi(P) \subset N$, the restriction in range $\varphi^{\restriction N} : P \to N$ is continuous. For that purpose, we will show that for every $p \in P$ and for every open neighbourhood V of $\varphi(p)$ in N, the preimage $\varphi^{-1}(V)$ is an open neighbourhood of p in P.

Choose a local chart (U, κ) adapted to D at $\varphi(p)$. Since the slice U_0 is an integral manifold through $\varphi(p)$ and since $\varphi(p) \in N$, Lemma 3.5.14/1 implies that $U_0 \cap N$ is an open submanifold of N and of U_0. Hence, without loss of generality we may restrict attention to open neighbourhoods V of $\varphi(p)$ in $U_0 \cap N$. Since $U_0 \cap N$ is open in U_0 and since U_0 is embedded, there exists an open subset $\tilde{V}$ of U such that $\tilde{V} \cap U_0 \cap N = V$. Now, let W be the connected component of p in $\varphi^{-1}(\tilde{V})$. By construction, W is an open neighbourhood of p in P. It remains to show that $W \subset \varphi^{-1}(V)$, that is, $\varphi(W) \subset V$. Since W is a connected subset of P and since φ is continuous as a mapping to M, $\varphi(W)$ is a connected subset of M which contains $\varphi(p)$ and which is contained in $N \cap U$. Hence, the image of $\varphi(W)$ under the mapping $(\kappa^{r+1}, \dots, \kappa^n)$ is a connected subset of $\mathbb{R}^{n-r}$ which contains 0 and which is contained in the image of $N \cap U$ under this mapping. Since the latter image labels the slices of U intersecting N, it is countable by Lemma 3.5.14/2. Since a nonempty countable connected subset of $\mathbb{R}^{n-r}$ consists of a single point, it follows that $\kappa^{r+1}(\varphi(W)) = \cdots = \kappa^n(\varphi(W)) = 0$ and hence $\varphi(W) \subset N \cap U_0$. Then, $\varphi(W) \subset \tilde{V} \cap U_0 \cap N = V$. This completes the proof. □

Remark 3.5.16 It follows, in particular, that the images of the maximal integral curves of a vector field on M are initial submanifolds of M, cf. the remark after Proposition 3.2.11. This can also be proved directly by means of the Straightening Lemma 3.2.17 and the above countability argument.

Theorem 3.5.17 *Let M be a manifold and let D be an integrable distribution on M. For every $m \in M$ there exists a unique integral manifold (N_m, ψ_m) of D through m which is maximal in the following sense. For every integral manifold (N, ψ) of D such that $\psi(N) \cap \psi_m(N_m) \neq \varnothing$ there holds $\psi(N) \subset \psi_m(N_m)$, and $(N, \psi^{\restriction N_m})$ is an open submanifold of N_m.*

Proof As mentioned above, for convenience, in the proof we assume all integral manifolds to be given by subsets. Moreover, we will repeatedly use the statements of Lemma 3.5.14, mostly without explicitly spelling that out. Let $m \in M$ be given and let $r = \dim D_m$. Every integral manifold of D through m has dimension r. Define a subset $N_m \subset M$ by taking the union over all integral manifolds of D through m. Equip N_m with the topology generated by the open subsets of the integral manifolds through m. Then, the union over the maximal atlases of the integral manifolds through m defines an atlas on N_m. By Lemma 3.5.14/1, this atlas is smooth. By the same argument as in the case of a regular distribution, see e.g. [302], one can show that N_m is second countable. Then, N_m is a manifold of dimension r. By construction, the local representatives of the natural inclusion mapping $N_m \to M$ are smooth, hence N_m is a smooth submanifold of M. Also by construction, every $\tilde{m} \in N_m$ belongs to an integral manifold N through m and N is an open submanifold of N_m. Hence, $\mathrm{T}_{\tilde{m}} N_m = \mathrm{T}_{\tilde{m}} N = D_{\tilde{m}}$, so that N_m is an integral manifold through m. It has the maximality property stated in the theorem, because if some integral manifold N of D intersects N_m, then $N \cup N_m$ is an integral manifold through m, hence it is contained in N_m. There follows $N \cap N_m = N$ and hence N is an open submanifold of N_m. To show uniqueness, let $\tilde{N}_m$ be an integral manifold through m which has the above maximality property. Then, $\tilde{N}_m = N_m$ as sets and each of them is an open submanifold of the other one, hence they are equal as manifolds. □

The properties of the family of maximal integral manifolds of an integrable distribution can be conveniently summarized in the notion of foliation. Our definition follows [280].

Definition 3.5.18 (Foliation) Let M be a manifold of dimension n. A foliation of M is a family $\mathscr{N}$ of connected submanifolds of M, called the leaves of the foliation, such that

1. the leaves are pairwise disjoint and $\bigcup_{N \in \mathscr{N}} N = M$,[17]
2. for every $m \in M$ there exists a local chart (U, κ) of M satisfying

[17] That is, $\mathscr{N}$ is a partition of M.

(a) $\kappa(m) = 0$ and $\kappa(U) = (-\varepsilon, \varepsilon)^n$ for some $\varepsilon > 0$,
(b) the leaves are invariant under the flows of the local vector fields $\partial_1, \ldots, \partial_r$, where r denotes the dimension of the leaf containing m.

The function which assigns to $m \in M$ the dimension of its leaf is called the rank or the dimension of $\mathscr{N}$ at m. $\mathscr{N}$ is called regular if this function is constant, that is, if all leaves have the same dimension. Otherwise, $\mathscr{N}$ is called singular.

Concerning terminology, the remark made after Definition 3.5.1 applies accordingly to foliations. Local charts at m satisfying the two conditions of point 2 will be called adapted to $\mathscr{N}$.

Remark 3.5.19

1. If $\mathscr{N}$ is regular of dimension r, condition 2b is equivalent to the condition that for every leaf N, the image $\kappa(N \cap U)$ is a union of subsets $(-\varepsilon, \varepsilon)^r \times \{\mathbf{c}\}$, where $\mathbf{c} \in (-\varepsilon, \varepsilon)^{n-r}$. This union is necessarily at most countable. Moreover, in this case, if a leaf N is an embedded submanifold of M, it is closed in M. To see this, let $\{m_n\}$ be a sequence in N converging to some $m \in M$. Let N_m denote the leaf of m. For every local chart (U, κ) at m adapted to $\mathscr{N}$, U intersects N. Since N is embedded, (U, κ) can be chosen so that $\kappa(N \cap U) = (-\varepsilon, \varepsilon)^r \times \{\mathbf{c}\}$ for some $\mathbf{c} \in (-\varepsilon, \varepsilon)^{n-r}$. Since $\kappa(m_n) \to \kappa(m) = 0$, it follows that $\mathbf{c} = 0$ and hence $m \in N$.
2. By condition 2b, the dimension of leaves is locally non-decreasing, that is, for every $m \in M$ there exists a neighbourhood U such that the dimension of any leaf intersecting U is equal to or greater than the dimension of the leaf containing m.

Example 3.5.20

1. The family of connected components of a manifold M is a regular foliation of dimension n. The family of all one-point subsets is the only regular foliation of dimension 0.
2. If $\pi : M \to P$ is a smooth submersion with the property that the preimage $\pi^{-1}(p)$ is connected for all $p \in P$, the Constant Rank Theorem implies that the preimages form a regular foliation of dimension $\dim M - \dim P$. A foliation of this form is called simple. By Remark 1.5.16/4, for a given simple foliation of M, the manifold P is unique up to diffeomorphisms. It is referred to as the space of leaves of the foliation. Let us add that Theorem 3.5.10/4 says that every integrable regular foliation is locally simple.
3. According to Proposition 3.2.17, the images of the integral curves of a vector field on M form a foliation. It is regular of dimension 1 if the vector field does not have zeros.
4. The partition of $\mathbb{R}^2$ consisting of the single points of the x-axis and the two open half-planes separated by the x-axis is a smooth foliation, whereas the partition consisting of this axis and the single points outside is not.

Proposition 3.5.21 *The assignment of the family of maximal integral manifolds to an integrable distribution defines a bijection between smooth integrable distributions on M and smooth foliations on M. Regular distributions of dimension r thereby correspond to regular foliations of dimension r.*

Proof The family of maximal integral manifolds of an integrable distribution is a foliation: Theorem 3.5.17 implies that condition 1 of Definition 3.5.18 is satisfied. Theorem 3.5.10/4 and the invariance of maximal integral manifolds under the flows of the elements of $\mathfrak{X}^D_{\mathrm{loc}}(M)$ imply that condition 2 is satisfied. Conversely, let $\mathscr{N}$ be a smooth foliation. For $m \in M$, let D_m be the tangent space at m of the leaf containing m. This defines a subset $D \subset \mathrm{T}M$. By condition 2(b) of Definition 3.5.18, D_m is spanned by the values at m of elements of $\mathfrak{X}^D_{\mathrm{loc}}(M)$. Hence, D is a distribution. By construction, every leaf of $\mathscr{N}$ is an integral manifold of D. First, this implies that D is integrable. Second, in view of Theorem 3.5.17, this implies that every leaf of $\mathscr{N}$ is contained as an open subset in a maximal integral manifold N of D. Thus, N is a union of leaves. Since the leaves are disjoint and since N is connected, N must coincide with a single leaf. This shows that $\mathscr{N}$ is the family of maximal integral manifolds of D, and the assignment is bijective, indeed. The assertion about regular distributions and foliations then follows by observing that for every $m \in M$, the dimension of the maximal integral manifold of an integrable distribution through m is equal to the rank of this distribution at m. □

Exercises

3.5.1 Use Remark 2.3.2/4 and Formula (3.1.1) to prove that the conditions given in Remark 3.5.6 are equivalent to involutivity.

3.5.2 Show that the vector fields $X = \partial_x + y\partial_z$ and $Y = (x^2 + 1)\partial_y$ generate a non-involutive regular distribution of rank 2 on $\mathbb{R}^3$.

3.5.3 Show that the distribution on $\mathbb{R}^3 \setminus \{0\}$, generated by the vector fields

$$X_1 = x_2\partial_3 - x_3\partial_2, \qquad X_2 = x_3\partial_1 - x_1\partial_3, \qquad X_3 = x_1\partial_2 - x_2\partial_1,$$

is regular of rank 2 and involutive. Find the maximal integral manifolds. How are they related to the action of the rotation group SO(3) on $\mathbb{R}^3$?

3.5.4 Show that the distribution on the unit sphere $\mathrm{S}^3 \subset \mathbb{R}^4$, generated by the vector fields

$$X = \left(1 - x_4 - x_1^2\right)\partial_1 - x_1x_2\partial_2 - x_1x_3\partial_3 + x_1(1 - x_4)\partial_4,$$
$$Y = -x_1x_2\partial_1 + \left(1 - x_4 - x_2^2\right)\partial_2 - x_2x_3\partial_3 + x_2(1 - x_4)\partial_4,$$

is integrable. Find the maximal integral manifolds. Is this distribution regular?

3.6 Critical Integral Curves

In the last four sections of this chapter, we discuss qualitative aspects of the flows of vector fields. Usually, this is treated as a part of the theory of dynamical systems.[18] Let M be a smooth manifold, let X be a vector field on M and let $\Phi : \mathscr{D} \to M$ be the flow of X. Recall from Proposition 3.2.11 that the image of a maximal integral curve of X is a submanifold of M which consists of a single point or is diffeomorphic either to S^1 or to $\mathbb{R}$.

Definition 3.6.1 (Critical integral curve) A maximal integral curve γ of X is called critical if its image is compact, that is, if the image consists of a single point or if it is diffeomorphic to the sphere S^1. In the first case, γ is called an equilibrium of X. In the second case, it is called a periodic integral curve.

In what follows, by an abuse of terminology and notation, by an integral curve γ of X we will mean the mapping itself or its image in M, depending on the context.

Let γ be a periodic integral curve of X and let $m \in \gamma$. The minimal positive real number $T \in \mathbb{R}$ satisfying $\Phi_T(m) = m$ is called the period of γ and Φ_T is called the period mapping. Due to $\Phi_T \circ \Phi_t(m) = \Phi_t \circ \Phi_T(m)$, the period does not depend on the choice of m. Since a considerable part of the analysis of the flow near a periodic integral curve is reduced to the study of fixed points of local diffeomorphisms, parallely to periodic integral curves of flows we will discuss fixed points of local diffeomorphisms.

A fundamental tool in the study of critical integral curves is the linearized flow. Given a critical integral curve γ of X, we consider the normal bundle

$$\mathrm{N}\gamma = (\mathrm{T}M)_{\restriction\gamma}/\mathrm{T}\gamma$$

of the submanifold γ of M. Recall that $\mathrm{N}\gamma$ is a real vector bundle over γ of dimension $\dim M - \dim \gamma$. Since γ is invariant under the flow Φ, the tangent mappings $(\Phi_t)'$, $t \in \mathbb{R}$, induce vector bundle automorphisms

$$(\Phi_t)'^{\gamma} : \mathrm{N}\gamma \to \mathrm{N}\gamma, \tag{3.6.1}$$

which project to the diffeomorphisms of γ induced by Φ_t. The family $\{(\Phi_t)'^{\gamma} : t \in \mathbb{R}\}$ is a one-parameter group of diffeomorphisms and hence defines a complete flow on $\mathrm{N}\gamma$.

Definition 3.6.2 (Linearized flow) The flow on $\mathrm{N}\gamma$ defined by $\{(\Phi_t)'^{\gamma} : t \in \mathbb{R}\}$ is called the linearization of Φ along γ.

We are going to determine the linearization explicitly. First, consider the case of an equilibrium. Here, $\gamma = \{m\}$ and $\mathrm{N}\gamma = \mathrm{T}_m M$. The linearized flow $(\Phi_t)'^{\gamma}$ is given by the one-parameter group of vector space automorphisms $(\Phi_t)'_m : \mathrm{T}_m M \to \mathrm{T}_m M$,

---

[18] A (continuous time) dynamical system consists of a manifold M and a vector field on M. We do not use this terminology here.

$t \in \mathbb{R}$. Hence, according to Example 3.2.8, this flow is generated by the linear vector field on $\mathrm{T}_m M$ which corresponds to the vector space endomorphism

$$\mathrm{Hess}_m(X) := \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} (\Phi_t)'_m \tag{3.6.2}$$

of $\mathrm{T}_m M$ and one has

$$(\Phi_t)'_m = \mathrm{e}^{t\,\mathrm{Hess}_m(X)}, \quad t \in \mathbb{R}. \tag{3.6.3}$$

Definition 3.6.3 The mapping $\mathrm{Hess}_m(X) : \mathrm{T}_m M \to \mathrm{T}_m M$ is called the Hessian endomorphism of X at m.

Remark 3.6.4

1. Let (U, κ) be a local chart at m and let X^i denote the coefficient functions of X with respect to the local frame of $\mathrm{T}M$ induced by κ. A brief computation (Exercise 3.6.1) shows that with respect to the basis $\{\partial_{1,m}, \ldots, \partial_{n,m}\}$ of $\mathrm{T}_m M$, the endomorphism $\mathrm{Hess}_m(X)$ is represented by the $(n \times n)$-matrix

$$\big[\kappa'_m \circ \mathrm{Hess}_m(X) \circ \big(\kappa'_m\big)^{-1}\big]^i_j = \frac{\partial (X^i \circ \kappa^{-1})}{\partial x^j}\big(\kappa(m)\big). \tag{3.6.4}$$

2. $\mathrm{Hess}_m(X)$ can be characterized by X alone, without using the flow, as follows. Let s_0 be the zero section of $\mathrm{T}M$. Its image $s_0(M)$ is an embedded submanifold of $\mathrm{T}M$, diffeomorphic to M. Consider the tangent mapping $X'_m : \mathrm{T}_m M \to \mathrm{T}_{s_0(m)}(\mathrm{T}M)$. Since $s_0(M)$ and $\mathrm{T}_m M$ intersect transversally in $s_0(m)$, one has

$$\mathrm{T}_{s_0(m)}(\mathrm{T}M) = \mathrm{T}_{s_0(m)} s_0(M) + \mathrm{T}_{s_0(m)}(\mathrm{T}_m M).$$

By counting dimensions one sees that the sum is direct, hence it defines a projection to $\mathrm{T}_{s_0(m)}(\mathrm{T}_m M)$. Composing this projection with the natural identification of $\mathrm{T}_{s_0(m)}(\mathrm{T}_m M)$ with $\mathrm{T}_m M$, one obtains a linear mapping $\tau_m : \mathrm{T}_{s_0(m)}(\mathrm{T}M) \to \mathrm{T}_m M$. Using τ_m, one can express $\mathrm{Hess}_m(X)$ as

$$\mathrm{Hess}_m(X) = \tau_m \circ X'_m. \tag{3.6.5}$$

The easiest way to prove this is to use (3.6.4). This is left to the reader (Exercise 3.6.2).

Next, consider the case of a periodic integral curve. In contrast to equilibria, periodic integral curves come in a great variety of shapes. Hence, we cannot expect to obtain an explicit formula for the linearized flow along γ like (3.6.3) for the case of an equilibrium. Rather, we will construct a normal form for the linearized flow along γ. This requires the following notion.

Definition 3.6.5 (Conjugacy) Let $k \geq 0$.

1. Flows $\Phi^{(i)} : \mathscr{D}^{(i)} \to M_i$, $i = 1, 2$, as well as the corresponding vector fields, are said to be conjugate of class C^k if there exists a diffeomorphism $h \colon M_1 \to M_2$ of

class C^k such that $(\mathrm{id}_{\mathbb{R}} \times h)(\mathscr{D}^{(1)}) = \mathscr{D}^{(2)}$ and, for all $(t, m) \in \mathscr{D}^{(1)}$,

$$h\big(\Phi_t^{(1)}(m)\big) = \Phi_t^{(2)}\big(h(m)\big).$$

2. Local diffeomorphisms $\varphi^{(i)} : U_i \to M_i$, $i = 1, 2$, are said to be conjugate of class C^k if there exists a diffeomorphism $h: M_1 \to M_2$ of class C^k such that $U_2 = h(U_1)$ and

$$h \circ \varphi^{(1)} = \varphi^{(2)} \circ h_{\restriction U_1}.$$

Recall that a diffeomorphism of class C^0 is just a homeomorphism. In this case, one speaks of topological conjugacy.

As a second input, we need the suspension of a vector space automorphism. Let V be a finite-dimensional real vector space, let $a \in \mathrm{GL}(V)$ and let $T > 0$. We define an equivalence relation on the direct product $V \times \mathbb{R}$ by $(v_1, t_1) \sim (v_2, t_2)$ iff there exists $k \in \mathbb{Z}$ such that $(v_2, t_2) = (a^k v_1, t_1 - kT)$. Let V^a denote the set of equivalence classes. By constructing local charts in an analogous way as for the Möbius strip in Example 1.1.12, one can equip V^a with a smooth structure (Exercise 3.6.4). With respect to this structure, the natural projection $V \times \mathbb{R} \to V^a$ is a local diffeomorphism. In particular, V^a has the same dimension as $V \times \mathbb{R}$, i.e., $\dim V + 1$. The flow on $V \times \mathbb{R}$ induced by the standard vector field $(0, \frac{\mathrm{d}}{\mathrm{d}t})$ is given by $(t, (v, s)) \mapsto (v, s + t)$. Hence, it maps equivalence classes to equivalence classes and thus induces a flow Σ^a on V^a. By definition, the projected flow is given by

$$\Sigma_t^a\big(\big[(v, s)\big]\big) = \big[(v, s + t)\big], \quad v \in V, \ s, t \in \mathbb{R}. \tag{3.6.6}$$

Definition 3.6.6 (Suspension) V^a is called the suspension with period T of V relative to a. The flow Σ^a on V^a is called the suspension of a with period T. The integral curve through $[(0, 0)]$ is referred to as the central integral curve of Σ^a.

The flow Σ^a is complete, that is, it is a one-parameter group of diffeomorphisms of V^a. It does not have fixed points. Due to

$$\Sigma_{kT}^a\big(\big[(v, s)\big]\big) = \big[(v, s + kT)\big] = \big[\big(a^k v, s\big)\big], \quad v \in V, \ s \in \mathbb{R}, \ k \in \mathbb{Z},$$

an integral curve of Σ^a is periodic iff it passes through a point $[(v, s)]$ satisfying $a^k(v) = v$ for some $k \in \mathbb{Z} \setminus \{0\}$. In this case, the integral curve has period kT. In particular, the central integral curve is periodic with period T. Let us add that V^a can be viewed as a vector bundle over $\mathbb{R}/T\mathbb{Z} \cong \mathrm{S}^1$ with typical fibre V and transition functions with respect to a covering of S^1 by two open intervals U_1, U_2 given by the constant mappings $x \mapsto \mathrm{id}_V$ for one of the connected components of $U_1 \cap U_2$ and $x \mapsto a$ for the other one.

Example 3.6.7

1. For $V = \mathbb{R}$ and a being multiplication by -1, the suspension V^a with period $T = 1$ is diffeomorphic to the Möbius strip, cf. Examples 1.1.12 and 2.2.4. All integral curves of Σ^a are periodic. The central integral curve has period 1 and the other integral curves have period 2.

2. For $V = \mathbb{C}$ and a being multiplication by $e^{i\alpha}$, the suspension with period $T = 1$ is diffeomorphic to $\mathbb{C} \times S^1$. If $\frac{\alpha}{2\pi}$ is rational, all integral curves are periodic, where those apart from the central one have the same period (a certain integer). Otherwise, the central integral curve is the only one which is periodic. For $r > 0$, the submanifold $\{[(z,t)] : |z| = r\}$ is invariant under the suspended flow and the induced flow is conjugate to the rational or irrational torus flow, cf. Example 1.6.6/3.

Now, we can construct the desired normal form for the linearized flow along a periodic integral curve γ. For that purpose, we observe that the mapping

$$P_m^\gamma := (\Phi_T)_m'^\gamma : \mathrm{N}_m\gamma \to \mathrm{N}_m\gamma \tag{3.6.7}$$

is a vector space automorphism. It is called the period automorphism of γ at m.

Proposition 3.6.8 *Let γ be a periodic integral curve of X with period T. For every $m \in \gamma$, the linearized flow is smoothly conjugate to the suspension with period T of the period automorphism P_m^γ.*

Proof Denote $a = P_m^\gamma$ and define a mapping

$$\tilde{h}: \mathrm{N}_m\gamma \times \mathbb{R} \to \mathrm{N}\gamma, \qquad \tilde{h}\big([X], t\big) := (\Phi_t)'^\gamma [X].$$

For simplicity, we write $v \equiv [X]$. Due to $\tilde{h}(a^k v, t - kT) = \tilde{h}(v,t)$ for all $k \in \mathbb{Z}$, the mapping $\tilde{h}$ is constant on equivalence classes and hence induces a mapping

$$h: (\mathrm{N}_m\gamma)^a \to \mathrm{N}\gamma, \qquad h\big([(v,t)]\big) := (\Phi_t)'^\gamma v.$$

By construction,

$$h \circ \Sigma_t^a = (\Phi_t)'^\gamma \circ h.$$

It remains to show that h is a diffeomorphism. Since the natural projection $\mathrm{N}_m\gamma \times \mathbb{R} \to (\mathrm{N}_m\gamma)^a$ is a local diffeomorphism, it suffices to show that h is bijective and that $\tilde{h}'_{(v,t)}$ is a bijection for all $(v,s) \in \mathrm{N}_m\gamma \times \mathbb{R}$. Surjectivity of h follows from the fact that the integral curves of the linearized flow project to γ and hence each of them passes through $\mathrm{N}_m\gamma$. To prove injectivity, let $(v_1,t_1), (v_2,t_2) \in \mathrm{N}_m\gamma \times \mathbb{R}$ such that $\tilde{h}(v_1,t_1) = \tilde{h}(v_2,t_2)$. Then, $v_2 = (\Phi_{t_1-t_2})'^\gamma v_1$. Since v_1 and v_2 are both in $\mathrm{N}_m\gamma$, there exists $k \in \mathbb{Z}$ such that $t_1 - t_2 = kT$. It follows that $v_2 = a^k v_1$ and hence $[(v_2,t_2)] = [(v_1,t_1)]$. Finally, since $\mathrm{N}_m\gamma \times \mathbb{R}$ and $\mathrm{N}\gamma$ have the same dimension, bijectivity of $\tilde{h}'_{(v,t)}$ follows from surjectivity. The latter follows by observing that for every $t \in \mathbb{R}$, the mapping $v \mapsto \tilde{h}(v,t)$ is a vector space isomorphism of fibres of $\mathrm{N}\gamma$, and that for every $v \in \mathrm{N}_m\gamma$, the mapping $t \mapsto \tilde{h}(v,t)$ is a curve in $\mathrm{N}\gamma$ projecting to γ and hence intersecting each fibre transversally. □

Remark 3.6.9 Differentiation of the obvious equality $\Phi_T = \Phi_t \circ \Phi_T \circ \Phi_{-t}$ at $\Phi_t(m)$ yields

$$P_{\Phi_t(m)}^\gamma = (\Phi_t)_m'^\gamma \circ P_m^\gamma \circ (\Phi_{-t})_{\Phi_t(m)}'^\gamma.$$

Hence, all period automorphisms are conjugate under the linearized flow.

As a result of the previous discussion, the linearized flow along a critical integral curve γ is completely determined by a certain characteristic linear mapping. In case γ is an equilibrium, this is the Hessian endomorphism. Its eigenvalues are called the characteristic exponents of γ. In case γ is periodic this is the period automorphism at a chosen point $m \in \gamma$. Its eigenvalues are called the characteristic multipliers of γ. Extending this terminology, by the characteristic linear mapping associated with a fixed point m of a local diffeomorphism φ we mean the tangent mapping φ'_m. Its eigenvalues are called the characteristic multipliers of m. Note that, being eigenvalues of linear mappings, characteristic exponents and multipliers come with geometric and algebraic multiplicities.[19]

Next, we recall the following terminology from linear algebra. An endomorphism of a finite-dimensional real vector space V is called non-degenerate if all eigenvalues are nonzero. A non-degenerate endomorphism of V is called elliptic (hyperbolic) if every (no) eigenvalue lies on the imaginary axis. An automorphism of V is called non-degenerate if all eigenvalues are distinct from 1. A non-degenerate automorphism of V is called elliptic (hyperbolic) if every (no) eigenvalue lies on the unit circle. Via the characteristic linear mapping, this terminology carries over to critical integral curves of vector fields and to fixed points of local diffeomorphisms. Thereby, the characteristic linear mapping of an equilibrium has to be treated as an endomorphism, because it is genuinely infinitesimal, whereas the characteristic linear mapping of a periodic integral curve or of a fixed point of a local diffeomorphism has to be treated as an automorphism, because it is genuinely a transformation.

Definition 3.6.10 (Ellipticity and hyperbolicity) A critical integral curve of a vector field or a fixed point of a local diffeomorphism is called non-degenerate, elliptic or hyperbolic if the associated characteristic linear mapping is, respectively, non-degenerate, elliptic or hyperbolic.

From the separate analysis of hyperbolic and elliptic critical integral curves, conclusions on the behaviour of the flow near an arbitrary critical integral curve can be drawn.

Remark 3.6.11

1. Let γ be a periodic integral curve of X. According to Remark 3.6.9, the characteristic multipliers of X at γ do not depend on the choice of the point $m \in \gamma$. Moreover, due to

$$(\Phi_T)'_m X_m = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} \Phi_T\big(\Phi_t(m)\big) = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} \Phi_t(m) = X_m,$$

X_m is an eigenvector of the automorphism $(\Phi_T)'_m$ of $\mathrm{T}_m M$ with eigenvalue 1. It follows that the characteristic multipliers of X at γ are given by the eigenvalues

[19]The geometric multiplicity of an eigenvalue is the number of linearly independent eigenvectors. The algebraic multiplicity is defined as the multiplicity of the corresponding root of the characteristic polynomial.

of $(\Phi_T)'_m$ with the multiplicity of the eigenvalue 1 of $(\Phi_T)'_m$ reduced by one (which may make this eigenvalue disappear). In particular, γ is non-degenerate iff the eigenvalue 1 of $(\Phi_T)'_m$ has multiplicity one.

2. As a consequence of the Inverse Mapping Theorem, non-degenerate equilibria of vector fields and non-degenerate fixed points of local diffeomorphisms are isolated (Exercise 3.6.5). Moreover, since X is continuous and periodic integral curves are compact, each of them possesses a neighbourhood which is free of equilibria.
3. Let M be an open subset of a finite-dimensional real vector space V and let γ be a periodic integral curve of a vector field X on M. According to Remark 2.3.4/1, under the natural identification of $\mathrm{T}M$ with $M \times V \subset V \times V$, X is represented by a smooth mapping $X : M \to V$ and the tangent mapping $(\Phi_t)'$ can be written in the form $(\Phi_t)'_v(v, u) = (\Phi_t(v), A_v(t)u)$ with $v, u \in V$ and $A_v(t) \in \mathrm{GL}(V)$. For every $v \in \gamma$, this defines a smooth curve $A_v : \mathbb{R} \to \mathrm{GL}(V)$. Differentiation of this curve with respect to t yields the ordinary first order differential equation
$$\dot{A}_v(t) = X'\big(\Phi_t(v)\big)A_v(t) \tag{3.6.8}$$
with initial condition $A_v(0) = \mathbb{1}$. The solution is known as the path-ordered exponential[20,21] and is usually written in the form
$$A_v(t) = \mathrm{T}\exp \int_0^t X'\big(\Phi_s(v)\big)\,\mathrm{d}s.$$
The linearized flow $(\Phi_t)'^{\gamma} : \mathrm{N}\gamma \to \mathrm{N}\gamma$ along γ is thus represented by the family of mappings
$$V/\mathbb{R}X(v) \to V/\mathbb{R}X\big(\Phi_t(v)\big), \quad v \in \gamma,\ t \in \mathbb{R},$$
given by
$$(\Phi_t)'^{\gamma}_v\big(u + \mathbb{R}X(v)\big) = A_v(t)u + \mathbb{R}X\big(\Phi_t(v)\big), \quad u \in V.$$
Hence, the characteristic multipliers of γ coincide with the eigenvalues of $A_v(T)$ with the multiplicity of the eigenvalue 1 reduced by 1 (so that this eigenvalue may disappear).

Example 3.6.12 The vector field $X = 2y\partial_x + (4x - 4x^3)\partial_y$ on $\mathbb{R}^2$ has the equilibrium points $m_1 = (-1, 0)$, $m_2 = (1, 0)$ and $m_3 = (0, 0)$. According to (3.6.4), with respect to the standard chart, the Hessian endomorphism of X at m_i is represented by the matrix
$$\frac{\partial X^i}{\partial x^j} = \begin{bmatrix} 0 & 2 \\ 4 - 12x^2 & 0 \end{bmatrix}.$$

---

[20] For negative t, $\int_0^t X'(\Phi_s(v))\,\mathrm{d}s$ has to be replaced by $-\int_t^0 X'(\Phi_s(v))\,\mathrm{d}s$.

[21] Since (3.6.8) has T-periodic coefficients, by a theorem of Floquet, the solution $A_v(t)$ can be written as $A_v(t) = A_{v,0}(t)\mathrm{e}^{tB_v}$ with T-periodic smooth $A_{v,0} : \mathbb{R} \to \mathrm{GL}(V)$ and constant $B_v \in \mathrm{End}(V)$.

The corresponding eigenvalues are $\lambda_{1\pm} = \lambda_{2\pm} = \pm 4\mathrm{i}$ and $\lambda_{3\pm} = \pm 2\sqrt{2}$. Hence, m_1 and m_2 are elliptic, whereas m_3 is hyperbolic. In particular, all equilibria are non-degenerate. Consider the function $f(x, y) = y^2 - 2x^2 + x^4$. Due to $X(f) = 0$, the level sets of f are invariant under the flow of X. A function with this property is called a first integral or a constant of motion of X. Since we are in two dimensions, it follows that the level set components of f coincide with the images of the integral curves. We leave it to the reader (Exercise 3.6.7) to visualize the flow of X using this observation.

Example 3.6.13 (Linear vector fields) Let V be a finite-dimensional real vector space and let X be a linear vector field on V, defined by the endomorphism A. Then, 0 is an equilibrium point of X. Under the natural identification of $\mathrm{T}_0 V$ with V, the linearized flow on $\mathrm{T}_0 V$ coincides with the flow of X on V. According to Example 3.2.8, the latter is given by $\Phi_t = \mathrm{e}^{tA}$. There follows $\mathrm{Hess}_0(X) = A$. Hence,

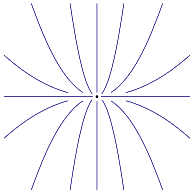
- the characteristic exponents of X at 0 are given by the eigenvalues of A,
- 0 is non-degenerate iff A is invertible; if so, it is the only equilibrium point,
- 0 is hyperbolic or elliptic iff so is A.

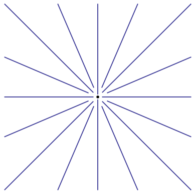
For illustration, we determine the flow in the neighbourhood of the equilibrium point 0 for the case $\dim V = 2$. We choose a basis such that the matrix representing A has Jordan normal form and use the results of Example 3.2.8. Up to the choice of basis, one can distinguish the following types.[22]

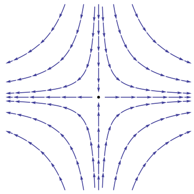
1. A has a zero eigenvalue.
   (a) $A = 0$, $\mathrm{e}^{tA} = \mathrm{id}_V$. Every point is an equilibrium point.
   (b) $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\mathrm{e}^{tA} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$.
   The x_1-axis consists of equilibria. The other maximal integral curves are parallel to the x_2-axis. Their velocity is constant in time and given by the x_2-coordinate.
   (c) $A = \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix}$, $\mathrm{e}^{tA} = \begin{bmatrix} \mathrm{e}^{\lambda t} & 0 \\ 0 & 1 \end{bmatrix}$, $\lambda \neq 0$.
   The x_2-axis consists of equilibria. The other maximal integral curves are parallel to the x_1-axis. Their velocity grows exponentially in time.
2. $A = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}$, $\mathrm{e}^{tA} = \begin{bmatrix} \mathrm{e}^{\lambda t} & 0 \\ 0 & \mathrm{e}^{\mu t} \end{bmatrix}$, $\lambda, \mu \in \mathbb{R}$, $\lambda \neq 0 \neq \mu$.
   The integral curves are given by the graphs of the functions
   $$x_2 = f(x_1) = \pm C|x_1|^{\mu/\lambda}, \quad x_1 \neq 0,\ C > 0, \tag{3.6.9}$$
   and the positive and negative coordinate semiaxes. Three cases can be distinguished.

[22] In the literature, linear vector fields X_1, X_2 defined by A_1, A_2 are said to be linearly equivalent if $A_2 = aA_1a^{-1}$ for some $a \in \mathrm{GL}(V)$. In this terminology, our choice of basis amounts to discussing the linear vector fields on $\mathbb{R}^2$ up to linear equivalence.

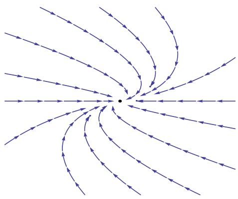
(a) $\lambda > \mu > 0$ or $\lambda < \mu < 0$. The curves (3.6.9) are branches of parabolas. In the first case, all integral curves run outwards, hence 0 is called an unstable knot. In the second case, they run inwards, hence 0 is called a stable knot.
(b) $\lambda = \mu$. The curves (3.6.9) are semi-infinite line segments. As in case 2(a), the running direction is outwards for $\lambda > 0$ and inwards for $\lambda < 0$. Correspondingly, 0 is called a degenerate unstable or stable knot.
(c) $\lambda > 0 > \mu$. The curves (3.6.9) are branches of hyperbolas. The running directions are shown in the figure below. The equilibrium point 0 is called a saddle here.

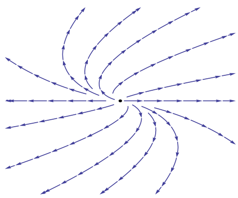

2(a) knot


2(b) degenerate knot

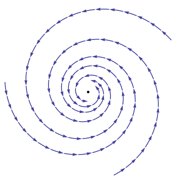

2(c) saddle

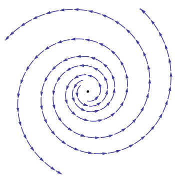
3. $A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$, $e^{tA} = e^{\lambda t} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$, $\lambda \in \mathbb{R}$, $\lambda \neq 0$.
For $\lambda < 0$ (case 3(a)), the maximal integral curves run inwards, whereas for $\lambda > 0$ (case 3(b)) they run outwards. Correspondingly, 0 is called a stable or unstable improper knot.

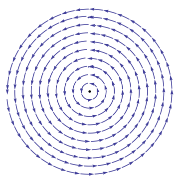

3(a) stable improper knot


3(b) unstable improper knot

4. $A = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}$, $e^{tA} = e^{\alpha t} \begin{bmatrix} \cos(\beta t) & -\sin(\beta t) \\ \sin(\beta t) & \cos(\beta t) \end{bmatrix}$, $\alpha, \beta \in \mathbb{R}$, $\beta > 0$.
This is the case of complex eigenvalues $\lambda_\pm = \alpha \pm i\beta$. If $\alpha < 0$ (case 4(a)), the maximal integral curves are logarithmic spirals approaching 0. If $\alpha > 0$ (case 4(b)), they are spirals running away from 0. If $\alpha = 0$ (case 4(c)), they form circles. Correspondingly, 0 is called a stable or an unstable spiral or a centre.


4(a) stable spiral


4(b) unstable spiral


4(c) centre

In the cases 2, 3, 4(a) and 4(b), the endomorphism A and hence the equilibrium $\{0\}$ is hyperbolic. In the case 4(c), A and hence $\{0\}$ is elliptic. Further critical integral curves besides $\{0\}$ are present in the cases 1(a)–1(c), where they form additional equilibria, and 4(c), where they form periodic integral curves. The latter have the common period $T = \frac{2\pi}{\beta}$. Due to $e^{TA} = \mathbb{1}$ there holds $(\Phi_T)'_v = (e^{TA})'_v = \mathbb{1}$ for all $v \neq 0$. Hence, the single characteristic multiplier is 1, so that all periodic integral curves are degenerate.

Examples for types 2(a), 2(b), 4(a) and 4(c) are provided by the harmonic oscillator in one dimension with linear friction, see Exercise 3.6.9.

Example 3.6.14 (The planar pendulum) Let ϕ, β denote the coordinates induced on $TS^1 \cong S^1 \times \mathbb{R}$ by the standard angle coordinate ϕ on S^1. The vector fields ∂_ϕ and ∂_β provide a global frame in $T(TS^1)$. The planar pendulum with length l, gravitational acceleration g, elongation ϕ and angular velocity β is modelled by the (nonlinear) vector field[23]

$$X = \beta\partial_\phi - \omega^2 \sin\phi\,\partial_\beta, \quad \omega^2 = \frac{g}{l},$$

on TS^1. The equilibrium points are given by $m_s = (0, 0)$ and $m_u = (\pi, 0)$. According to Remark 3.6.4/1, with respect to the basis $\{(\partial_\phi)_{m_i}, (\partial_\beta)_{m_i}\}$ in $T_{m_i}M$, $i = s, u$, the Hessian endomorphisms are represented by the matrices

$$\mathrm{Hess}_{m_s}(X) = \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix}, \qquad \mathrm{Hess}_{m_u}(X) = \begin{bmatrix} 0 & 1 \\ \omega^2 & 0 \end{bmatrix}.$$

The eigenvalues are $\pm i\omega$ for m_s and $\pm\omega$ for m_u, hence m_s is elliptic and m_u is hyperbolic. For the linearized flow on $T_{m_s}M$, the origin is a centre (type 4(c)), whereas for the linearized flow on $T_{m_u}M$, it is a saddle (type 2(c)).[24] The flow of X is shown in Fig. 3.2(a), where the left and right boundaries of the picture must be glued. There are contractible and non-contractible periodic integral curves. The first ones correspond to motions where the pendulum swings forth and back. The latter ones correspond to motions where the pendulum rotates. Altogether they form three 1-parameter families, each of which is diffeomorphic to a cylinder $S^1 \times \mathbb{R}$.[25] Finally, there are two maximal integral curves which are not critical, represented by the fat curves in the picture. They approach m_u for $t \to \pm\infty$ and correspond to the two motions where the pendulum asymptotically approaches the upward position from the left or the right, respectively. Together with m_u, these two integral curves divide M into three domains each of which is filled with periodic integral curves. Since periodic integral curves of different domains cannot be deformed into

---

[23] This is an example of a Hamiltonian system with one degree of freedom.

[24] We will see later that it is exactly these two types of equilibria that can occur in a Hamiltonian system with one degree of freedom.

[25] In Chap. 9 we will learn that periodic integral curves of Hamiltonian systems typically come in such “orbit cylinders”.

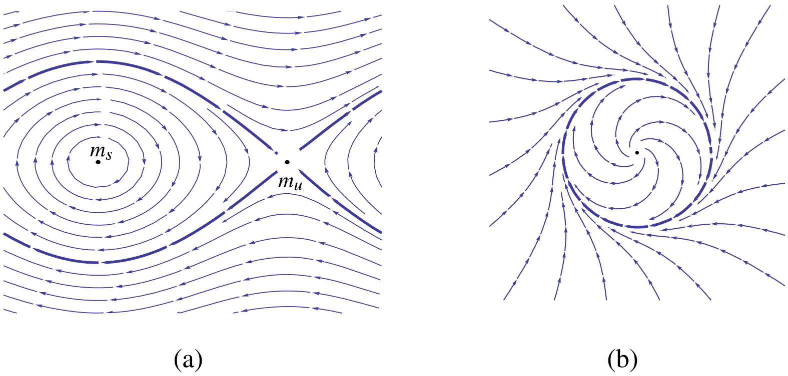


Fig. 3.2 The flows of (**a**) the planar pendulum and (**b**) the modified harmonic oscillator

one another through periodic integral curves, the subset[26] of M made up by the two non-critical integral curves and m_u is called the separatrix.

Example 3.6.15 (The modified harmonic oscillator) As a last example, consider the vector field

$$X = \left(y + \omega x\left(1 - x^2 - \frac{y^2}{\omega^2}\right)\right)\partial_x + \left(-\omega^2 x + \omega y\left(1 - x^2 - \frac{y^2}{\omega^2}\right)\right)\partial_y$$

on $\mathbb{R}^2$. This system is usually referred to as the modified harmonic oscillator. In the sequel we state results about the critical integral curves of this system and leave it to the reader to provide arguments or to carry out calculations.

X possesses the single equilibrium $\{0\}$. The characteristic exponents are $\omega(1 \pm \mathrm{i})$, hence 0 is hyperbolic and the linearized flow about 0 is an unstable spiral (type 4(b)), see Example 3.6.13. Outside the equilibrium, i.e., on $\mathbb{R}^2 \setminus \{0\}$, one has the global chart $\kappa = (r, \phi)$, defined by

$$x = r\cos\phi, \qquad y = \omega r \sin\phi.$$

We use it to identify $\mathbb{R}^2 \setminus \{0\}$ with $\mathbb{R}_+ \times \mathrm{S}^1$ and to move with our discussion to the latter. This way, the equations for the integral curves of X separate,

$$\dot{r} = \omega r\left(1 - r^2\right), \qquad \dot{\phi} = \omega,$$

and for the flow we obtain

$$\kappa \circ \Phi_t \circ \kappa^{-1}(r, \phi) = \left(\left(1 + \left(r^{-2} - 1\right)\mathrm{e}^{-2\omega t}\right)^{-\frac{1}{2}}, \phi + \omega t\right), \quad t \in \mathbb{R}, \qquad (3.6.10)$$

see Fig. 3.2(b). In particular, X is complete. From (3.6.10) one reads off that X possesses the single periodic integral curve γ, given in the coordinates r, ϕ by

[26] In fact, this subset is a figure eight submanifold, see Example 1.6.6/2.

$t \mapsto \gamma(t) = (1, \omega t)$. It has period $T = \frac{2\pi}{\omega}$. The tangent mapping of the flow at the point $(1, \phi)$ of γ is

$$\left(\kappa \circ \Phi_t \circ \kappa^{-1}\right)'_{(1,\phi)} = \begin{bmatrix} e^{-2\omega t} & 0 \\ 0 & 1 \end{bmatrix}.$$

In particular, the characteristic multiplier of γ is $e^{-4\pi}$, hence γ is hyperbolic, too. The normal bundle $N\gamma$ admits the global section $(1, \phi) \mapsto [(\partial_r)_{(1,\phi)}]$. With respect to the corresponding global trivialization $\chi : N\gamma \to \gamma \times \mathbb{R}$, the linearized flow along γ is given by

$$\chi \circ (\Phi_t)'^{\gamma} \circ \chi^{-1}\big((1, \phi), v\big) = \big((1, \phi + \omega t), e^{-2\omega t} v\big). \tag{3.6.11}$$

In particular, the period automorphism $P^{\gamma}_{(1,\phi)}$ amounts to multiplication by $e^{-4\pi}$. Identifying $N_{(1,\phi)}\gamma$ with $\mathbb{R}$ via χ, for the suspension of $N_{(1,\phi)}\gamma$ relative to $P^{\gamma}_{(1,\phi)}$ one obtains

$$(N_{(1,\phi)}\gamma)^{e^{-4\pi}} = (\mathbb{R} \times \mathbb{R})/\sim,$$

where $(s_1, v_1) \sim (s_2, v_2)$ iff $s_1 - s_2 = kT$ and $v_1 = e^{4\pi k} v_2$ for some $k \in \mathbb{Z}$. The mapping $\mathbb{R} \times \mathbb{R} \to \gamma \times \mathbb{R}$ defined by $(s, v) \mapsto ((1, \omega s), e^{-4\pi s} v)$ descends to a diffeomorphism $\tau : (N_{(1,\phi)}\gamma)^{e^{-4\pi}} \to \gamma \times \mathbb{R}$ which satisfies

$$\tau \circ \Sigma_t^{e^{-4\pi}} \circ \tau^{-1}\big((1, \phi), v\big) = \big((1, \phi + \omega t), e^{-4\pi t} v\big).$$

Thus, $\chi^{-1} \circ \tau$ conjugates the linearized flow along γ to the suspension of the period automorphism. In fact, it yields the diffeomorphism h constructed in the proof of Proposition 3.6.8.

Exercises

3.6.1 Prove Eq. (3.6.4).

3.6.2 Use (3.6.4) to prove Formula (3.6.5).

3.6.3 Let m be an equilibrium point of a vector field X on M and let (U, κ) be a local chart at m. Convince yourself by an explicit calculation that the eigenvalues of the matrix $\partial_j(X^i \circ \kappa^{-1})(\kappa(m))$ are independent of the choice of κ.

3.6.4 Along the lines of Example 1.1.12, construct an atlas on the suspension V^a with period T of a finite-dimensional real vector space V relative to an automorphism $a \in GL(V)$, see Definition 3.6.6.

3.6.5 Show that non-degenerate equilibria of vector fields and non-degenerate fixed points of local diffeomorphisms are isolated, cf. Remark 3.6.11.

3.6.6 Consider the Möbius strip as a real one-dimensional vector bundle over S^1 and let ϕ denote the angle coordinate on S^1. Show that every maximal integral curve of the vector field ∂_ϕ is periodic and determine the characteristic multiplier for each of them.

3.6.7 Draw a picture of the flow of the vector field of Example 3.6.12. Compare the behaviour of the flow near the elliptic equilibria with that near the hyperbolic equilibrium.

3.6.8 Prove Eq. (3.6.9).

3.6.9 The harmonic oscillator in one dimension with frequency ω, linear friction coefficient $\alpha \geq 0$, elongation x and velocity y is modelled in phase space by the linear vector field

$$X = y\partial_x - \big(\omega^2 x + 2\alpha y\big)\partial_y \tag{3.6.12}$$

on $\mathbb{R}^2$. Determine the type of flow about the equilibrium point 0 according to Example 3.6.13 for the cases $\alpha = 0$, $0 < \alpha < \omega$, $\alpha = \omega$ and $\alpha > \omega$, and discuss how the integral curves change with increasing α.

3.6.10 The index of a non-degenerate equilibrium $\{m\}$ of a vector field X is defined to be the sign of the determinant of $\mathrm{Hess}_m(X)$. Determine the index of the equilibrium $\{0\}$ for the vector field $X = -x^i\partial_i$ (summation convention) on $\mathbb{R}^n$.

3.7 The Poincaré Mapping

Let M be a smooth manifold and let X be a vector field on M with flow $\Phi : \mathscr{D} \to M$. Let γ be a periodic integral curve of X with period T. As mentioned before, a large part of the analysis of the flow of X near γ can be reduced to the study of a certain local diffeomorphism. This will be constructed now. The idea is that instead of continuously watching the flow along the integral curves near γ it suffices to record where these integral curves hit a certain transversal submanifold. This leads to the notion of Poincaré mapping.

Let $m_0 \in \gamma$ and let $\mathscr{P}$ be a submanifold of codimension 1 which is transversal[27] to the integral curves of X and which satisfies $\mathscr{P} \cap \gamma = \{m_0\}$. For dimensional reasons, transversality implies

$$X_m \notin \mathrm{T}_m\mathscr{P} \quad \text{for all } m \in \mathscr{P}. \tag{3.7.1}$$

The subset of points $m \in \mathscr{P}$ for which there exists $t > 0$ such that $\Phi_t(m) \in \mathscr{P}$ will be called the returning subset of $\mathscr{P}$ and will be denoted by $\mathscr{P}_{\mathrm{ret}}$. Due to (3.7.1), integral curves of X intersect $\mathscr{P}$ at isolated values of t. Thus, for every $m \in \mathscr{P}_{\mathrm{ret}}$, there exists a minimal flow parameter value $\tau(m) > 0$ such that $\Phi_{\tau(m)}(m) \in \mathscr{P}$. The assignment $m \mapsto \tau(m)$ defines a function $\tau : \mathscr{P}_{\mathrm{ret}} \to \mathbb{R}_+$, called the first return time function. The corresponding mapping

$$\Theta : \mathscr{P}_{\mathrm{ret}} \to \mathscr{P}, \qquad \Theta(m) := \Phi_{\tau(m)}(m), \tag{3.7.2}$$

is called the first return mapping. Obviously, $\tau(m_0) = T$ and $\Theta(m_0) = m_0$. Neither τ nor Θ need be continuous: for example, view the Möbius strip as the suspension of the automorphism $s \mapsto -s$ of $\mathbb{R}$, cf. Example 3.6.7/1. Let γ be the central integral curve of the corresponding flow and let $\mathscr{P}$ be given by the equivalence classes of pairs $(x, 0)$ with $b < x < a$, where $b < 0$ and $a > 0$. Unless $b = -a$, the mappings τ and Θ are not continuous.

[27] Recall that, in particular, $\mathscr{P}$ is transversal to all integral curves it does not intersect.

Lemma 3.7.1

1. *$\mathscr{P}_{\text{ret}}$ is an open neighbourhood of m_0 in $\mathscr{P}$.*
2. *Θ is smooth iff τ is smooth. In this case, for all $Y \in \mathrm{T}_{m_0}\mathscr{P}$,*

$$\Theta'(Y) = \tau'(Y)X_{m_0} + (\Phi_T)'(Y) \equiv (\Phi_T)'\big(\tau'(Y)X_{m_0} + Y\big). \qquad (3.7.3)$$

Proof 1. Let $m \in \mathscr{P}_{\text{ret}}$. Due to (3.7.1), $X_{\Theta(m)} \neq 0$. According to Proposition 3.2.17, there exists a flow box chart (U, κ) for X at $\Theta(m)$. Since $\mathscr{P}$ is transversal to the integral curves of X, $\kappa(U \cap \mathscr{P})$ is transversal to the maximal integral curves of the local representative of X. Thus, by possibly shrinking U transversally to the direction of the flow we can achieve that $\kappa(U \cap \mathscr{P})$ intersects all of these integral curves. Then, $\mathrm{pr}_M(\Phi^{-1}(U \cap \mathscr{P}) \setminus (\{0\} \times M)) = \mathrm{pr}_M(\Phi^{-1}(U) \setminus (\{0\} \times M))$. Since pr_M is an open mapping, this is an open subset of M. Hence, intersection with $\mathscr{P}$ yields an open subset of $\mathscr{P}$ which contains m and which is contained in $\mathscr{P}_{\text{ret}}$.

2. Obviously, if τ is smooth then so is Θ. Conversely, assume that Θ is smooth and let $m \in \mathscr{P}_{\text{ret}}$ be arbitrary. As before, there exists a flow box chart (U, κ) for X at $\Theta(m)$. By possibly shrinking U transversally to the direction of the flow we can achieve that for every $\tilde{m} \in \Theta^{-1}(U \cap \mathscr{P})$ one has $\tau(m) \in \mathscr{D}_{\tilde{m}}$ and $\Phi_{\tau(m)}(\tilde{m}) \in U$. Let $\mathrm{pr}_1 : \mathbb{R}^n \to \mathbb{R}$ denote the projection to the first coordinate (the one which corresponds to the flow parameter). Since $\Theta(\tilde{m}) = \Phi_{\tau(\tilde{m})}(\tilde{m})$ and $\Phi_{\tau(m)}(\tilde{m})$ are both in $U \cap \mathscr{P}$, the difference $\tau(\tilde{m}) - \tau(m)$ is given by

$$\tau(\tilde{m}) - \tau(m) = \mathrm{pr}_1 \circ \kappa\big(\Theta(\tilde{m})\big) - \mathrm{pr}_1 \circ \kappa\big(\Phi_{\tau(m)}(\tilde{m})\big).$$

Since the right hand side is smooth in the variable $\tilde{m}$, τ is smooth on the open neighbourhood $\Theta^{-1}(U \cap \mathscr{P})$ of m in $\mathscr{P}_{\text{ret}}$. Finally, (3.7.3) follows by a straightforward calculation using the product rule. □

Definition 3.7.2 (Poincaré mapping) Let γ be a periodic integral curve and let $m_0 \in \gamma$. A Poincaré mapping for γ at m_0 is a triple $(\mathscr{P}, \mathscr{W}, \Theta)$, where

1. $\mathscr{P}$ is an embedded submanifold of M of codimension 1 which is transversal to every integral curve of X and which satisfies $\mathscr{P} \cap \gamma = \{m_0\}$,
2. $\mathscr{W}$ is an open neighbourhood of m_0 in $\mathscr{P}_{\text{ret}}$ and $\Theta : \mathscr{W} \to \mathscr{P}$ is the restriction[28] to $\mathscr{W}$ of the first return mapping associated with $\mathscr{P}$,

such that Θ is a diffeomorphism onto its image.

Theorem 3.7.3 (Existence) *For every periodic integral curve γ of a vector field and for every $m_0 \in \gamma$, there exists a Poincaré mapping at m_0.*

Proof As before, denote the manifold by M, the vector field under consideration by X and let $\Phi : \mathscr{D} \to M$ be the flow of X. Since γ is periodic, $X_{m_0} \neq 0$, hence there exists $f \in C^\infty(M)$ such that $X_{m_0}(f) \neq 0$. Choose f such that $f(m_0) = 0$.

[28]Denoted by the same letter.

According to the Level Set Theorem, the subset $N \subset f^{-1}(0)$ of regular points is an embedded submanifold of M of codimension 1 containing m_0. There exists an open neighbourhood $\mathscr{P}$ of m_0 in N such that $X_m(f) \neq 0$ for all $m \in \mathscr{P}$, that is, $\mathscr{P}$ is transversal to the integral curves of X. Consequently, (3.7.1) holds and the integral curves intersect $\mathscr{P}$ in isolated points. Thus, by possibly shrinking $\mathscr{P}$ we can achieve that $\mathscr{P} \cap \gamma = \{m_0\}$. Due to the existence of flow box charts at m_0, we may furthermore assume that there exists $\varepsilon_1 > 0$ such that $(-\varepsilon_1, \varepsilon_1) \times \mathscr{P} \subset \mathscr{D}$ and $\Phi_t(m) \notin \mathscr{P}$ for all $m \in \mathscr{P}$ and $0 < |t| < \varepsilon_1$. Since $\mathscr{D}_{m_0} = \mathbb{R}$, there exists an open neighbourhood $\mathscr{W}$ of m_0 in $\mathscr{P}$ and $a > 0$ such that $(0, T + a) \times \mathscr{W} \subset \mathscr{D}$. Consider the smooth function $f \circ \Phi$ on $(0, T + a) \times \mathscr{W}$. We have

$$f \circ \Phi(T, m_0) = 0, \qquad \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction T} f \circ \Phi(t, m_0) = X_{m_0}(f) \neq 0.$$

Hence, after possibly shrinking $\mathscr{W}$, the Implicit Mapping Theorem 1.5.10 yields $0 < \varepsilon_2 < a$ and a smooth function $\tilde{\tau} : \mathscr{W} \to (T - \varepsilon_2, T + \varepsilon_2)$ such that for every $m \in \mathscr{W}$ and $t \in (T - \varepsilon_2, T + \varepsilon_2)$ there holds $f \circ \Phi(t, m) = 0$ iff $t = \tilde{\tau}(m)$. Define[29]

$$\tilde{\Theta} : \mathscr{W} \to N, \qquad \tilde{\Theta}(m) := \Phi_{\tilde{\tau}(m)}(m).$$

Since $\mathscr{P}$ is open in N and since $\tilde{\Theta}$ is continuous, by possibly further shrinking $\mathscr{W}$ we can achieve $\tilde{\Theta}(\mathscr{W}) \subset \mathscr{P}$. Then, $\tilde{\Theta}$ is a smooth mapping from $\mathscr{W}$ to $\mathscr{P}$ satisfying (3.7.3). Since $X_{m_0} \notin \mathrm{T}_{m_0}\mathscr{P}$ and $(\Phi_T)'_{m_0}$ is injective, the latter implies that $\tilde{\Theta}'_{m_0}$ is injective and hence bijective. According to the Inverse Mapping Theorem, by possibly shrinking $\mathscr{W}$ once more we can achieve that $\tilde{\Theta}$ is a diffeomorphism onto an open subset of $\mathscr{P}$.

Finally, we show that $\mathscr{P}$ and $\mathscr{W}$ can be shrunk so that $\tilde{\tau}$ becomes the first return time mapping τ and $\tilde{\Theta}$ becomes the restriction to $\mathscr{W}$ of the first return mapping Θ associated with $\mathscr{P}$. Assume, on the contrary, that this is impossible. Then, every neighbourhood V of m_0 in $\mathscr{P}$ contains a point m such that $\Phi_t(m) \in V$ for some $t \in (0, \tilde{\tau}(m))$. By our choice of $\mathscr{P}$ and by the uniqueness property of $\tilde{\tau}$, then $\varepsilon \leq t \leq T - \varepsilon$, where $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$. Thus, we can find sequences $\{m_n\}$ in $\mathscr{P}$ and $\{t_n\}$ in $[\varepsilon, T - \varepsilon]$ such that $m_n \to m_0$ and $\Phi_{t_n}(m_n) \to m_0$. On the other hand, the sequence $\{t_n\}$ has a cluster point t_0 in $[\varepsilon, T - \varepsilon]$ and, by passing to a subsequence converging to t_0, we obtain $\Phi_{t_n}(m_n) \to \Phi_{t_0}(m_0)$. Since $\mathscr{P}$ is Hausdorff, the two limits coincide, that is, there holds $\Phi_{t_0}(m_0) = m_0$ (contradiction). This completes the proof. □

Remark 3.7.4 Let $(\mathscr{P}, \mathscr{W}, \Theta)$ be a Poincaré mapping for γ at m_0. One can show that, by possibly shrinking $\mathscr{P}$ and $\mathscr{W}$, one may always assume that there exists $\varepsilon > 0$ such that

[29] Note that $\tilde{\tau}$ need not be the first return time mapping and $\tilde{\Theta}$ need not be the first return mapping associated with $\mathscr{P}$.

1. $(-\varepsilon, T+\varepsilon) \times \mathscr{W} \subset \mathscr{D}$,
2. the first return time function τ takes values in $(T-\varepsilon, T+\varepsilon)$,
3. if $\Phi_t(m) \in \mathscr{P}$ for some $(t, m) \in (-\varepsilon, T+\varepsilon) \times \mathscr{W}$, then $t = 0$ or $t = \tau(m)$,
4. for any neighbourhood U of m_0 in $\mathscr{W}$, $\Phi((-\varepsilon, T+\varepsilon) \times U)$ is a neighbourhood of γ in M.

The proof is left to the reader (Exercise 3.7.1).

Next, we study the relation between Poincaré mappings at different points of γ.

Proposition 3.7.5 (Uniqueness) *Any two Poincaré mappings $(\mathscr{P}_i, \mathscr{W}_i, \Theta_i)$ for γ at $m_i \in \gamma$, $i = 1, 2$, are equivalent in the following sense.*[30] *There exist open neighbourhoods $\tilde{\mathscr{P}}_i$ of m_i in $\mathscr{P}_i$ and $\tilde{\mathscr{W}}_i \subset \Theta_i^{-1}(\tilde{\mathscr{P}}_i)$ of m_i in $\mathscr{W}_i$ as well as a smooth mapping $\lambda : \tilde{\mathscr{P}}_1 \to \mathbb{R}$ such that the assignment $m \mapsto \Phi_{\lambda(m)}(m)$ defines a diffeomorphism $\varphi : \tilde{\mathscr{P}}_1 \to \tilde{\mathscr{P}}_2$ satisfying $\varphi(\tilde{\mathscr{W}}_1) = \tilde{\mathscr{W}}_2$ and*

$$\Theta_2 \circ \varphi(m) = \varphi \circ \Theta_1(m) \quad \textit{for all } m \in \tilde{\mathscr{W}}_1. \tag{3.7.4}$$

Proof There exists a unique $t_1 \in [0, T)$ such that $\Phi_{t_1}(m_1) = m_2$. According to Remark 1.8.4, there exists $f_2 \in C^\infty(M)$ such that $\mathscr{P}_2 = f_2^{-1}(0)$. Use f_2 and the initial solution (t_1, m_1) of the equation $f_2 \circ \Phi(t, m) = 0$ to construct the smooth function $\lambda : \tilde{\mathscr{P}}_1 \to \mathbb{R}$ and the diffeomorphism $\varphi : \tilde{\mathscr{P}}_1 \to \tilde{\mathscr{P}}_2$ in exactly the same way as τ and Θ, respectively, in the proof of Theorem 3.7.3. The resulting open neighbourhoods $\tilde{\mathscr{P}}_i$ of m_i in $\mathscr{P}_i$ can be shrunk in a compatible way such that there exists $\varepsilon > 0$ satisfying $(-\varepsilon, \varepsilon) \times \tilde{\mathscr{P}}_2 \subset \mathscr{D}$ and $\Phi_t(\tilde{\mathscr{P}}_2) \cap \tilde{\mathscr{P}}_2 = \varnothing$ for all $0 < |t| < \varepsilon$, as well as $|\lambda(m) - t_1| < \frac{\varepsilon}{6}$ for all $m \in \tilde{\mathscr{P}}_1$. Then, choose $\tilde{\mathscr{W}}_i$ so that $m_i \in \tilde{\mathscr{W}}_i \subset \Theta_i^{-1}(\tilde{\mathscr{P}}_i)$, $\varphi(\tilde{\mathscr{W}}_1) = \tilde{\mathscr{W}}_2$ and $|\tau_i(m) - T| < \frac{\varepsilon}{3}$ for all $m \in \tilde{\mathscr{W}}_i$. After these adjustments, (3.7.4) holds. Indeed, for any $m \in \tilde{\mathscr{W}}_1$, both $\Theta_2 \circ \varphi(m)$ and $\varphi \circ \Theta_1(m)$ are elements of $\tilde{\mathscr{P}}_2$ and $\Theta_2 \circ \varphi(m) = \Phi_\Delta \circ \varphi \circ \Theta_1(m)$, where

$$\Delta = \lambda(m) - \lambda\bigl(\Theta_1(m)\bigr) + \tau_2\bigl(\varphi(m)\bigr) - \tau_1(m).$$

Due to

$$|\Delta| \leq \bigl|\lambda(m) - \lambda\bigl(\Theta_1(m)\bigr)\bigr| + \bigl|\tau_2\bigl(\varphi(m)\bigr) - T\bigr| + \bigl|\tau_1(m) - T\bigr| < \varepsilon,$$

it follows $\Delta = 0$ and hence (3.7.4). This proves the proposition. □

Finally, we examine the relation between Poincaré mappings and the period automorphisms.

Proposition 3.7.6 *Let γ be a periodic integral curve of X and let $m_0 \in \gamma$. Let $(\mathscr{P}, \mathscr{W}, \Theta)$ be a Poincaré mapping for γ at m_0. By restriction, the natural projection* $\mathrm{pr}\colon \mathrm{T}_m M \to \mathrm{N}_m\gamma$ *induces an isomorphism $\chi : \mathrm{T}_m\mathscr{P} \to \mathrm{N}_m\gamma$ which satisfies*

$$\chi \circ \Theta'_m = P_m^\gamma \circ \chi. \tag{3.7.5}$$

[30] Briefly, their germs at m_i are conjugate in the sense of Definition 3.6.5.

In particular, the eigenvalues of Θ'_m coincide with the characteristic multipliers of X at γ.

Proof For dimensional reasons, (3.7.1) implies that χ is bijective and hence an isomorphism. Since $\mathrm{pr}(X_{m_0}) = 0$, (3.7.5) follows from (3.7.3). □

Example 3.7.7

1. Consider the modified harmonic oscillator of Example 3.6.15. As a submanifold transversal to the integral curves of X we may choose, in the coordinates x and y, $\mathscr{P} = \{(x, 0) : x > 0\}$. Denote the points of $\mathscr{P}$ by x. From (3.6.10) we read off the first return time function τ and the first return mapping Θ,
$$\tau(x) = T = \frac{2\pi}{\omega}, \qquad \Theta(x) = \big(1 + (x^{-2} - 1)\mathrm{e}^{-4\pi}\big)^{-\frac{1}{2}}.$$
Obviously, Θ is a diffeomorphism of $\mathscr{P}$ onto itself, hence, if we choose $\mathscr{W} = \mathscr{P}$, then $(\mathscr{P}, \mathscr{W}, \Theta)$ is a Poincaré mapping for γ. Let us add that the iterates of Θ are given by
$$\Theta^k(x) = \big(1 + (x^{-2} - 1)\mathrm{e}^{-4\pi k}\big)^{-\frac{1}{2}}, \tag{3.7.6}$$
which for large k is given asymptotically by $\Theta^k(x) \sim 1 - \frac{1}{2}(x^{-2} - 1)\mathrm{e}^{-4\pi k}$.
2. Let V be a finite-dimensional real vector space, let $a \in \mathrm{GL}(V)$ and let $T \in \mathbb{R}_+$. Consider the flow on the suspension V^a of V relative to a with period T which is given by the suspension Σ^a of a with period T. Let $m_0 := [(0, 0)] \in V^a$ and let γ be the central integral curve of Σ^a (which passes through m_0). Under the identification of V with the embedded submanifold $\{[(v, 0)] : v \in V\}$ of V^a, the triple $(\mathscr{P}, \mathscr{W}, \Theta) = (V, V, a)$ is a Poincaré mapping for γ at m_0 (Exercise 3.7.4).

Exercises

3.7.1 Show that every Poincaré mapping can be shrunk so that it satisfies conditions 1–4 of Remark 3.7.4.

3.7.2 Provide the details for the construction of the smooth function λ and the diffeomorphism φ in the proof of Proposition 3.7.5.

3.7.3 Verify Formula (3.7.6) of Example 3.7.7.

3.7.4 Prove that the mapping given in Example 3.7.7/2 is a Poincaré mapping for the central integral curve of the suspension of a vector space automorphism.

3.8 Stability

As already mentioned in the introduction, in the study of stability of a dynamical system one is concerned with the long-time behaviour of a flow, with emphasis both on returning properties and on attraction properties of integral curves. In this context a variety of concepts and a lot of subtle techniques exist. In our elementary introduction to this field, we limit our attention exclusively to orbital stability, which we merely call stability here. Only at the end of this section we will make some remarks on other concepts.

We give parallel definitions of stability for flows and for local diffeomorphisms. First, let X be a vector field on M and let $\Phi : \mathscr{D} \to M$ be its flow. A subset A of M is said to be invariant under Φ if $\Phi_t(m) \in A$ for all $m \in A$ and $t \in \mathscr{D}_m$. Of course, critical integral curves provide important examples for invariant subsets.

Definition 3.8.1 (Stability for flows) A Φ-invariant subset $A \subset M$ is called

1. stable under Φ if for every neighbourhood U of A in M there exists a neighbourhood V of A in M such that $\mathbb{R}_+ \times V \subset \mathscr{D}$ and $\Phi_t(m) \in U$ for all $m \in V$ and $t \in \mathbb{R}_+$,
2. asymptotically stable under Φ if it is stable under Φ and if there exists a neighbourhood V of A in M with the following properties:
   (a) $\mathbb{R}_+ \times V \subset \mathscr{D}$,
   (b) for every $m \in V$ and every neighbourhood U of A there exists $t_0 \in \mathbb{R}_+$ such that $\Phi_t(m) \in U$ for all $t \geq t_0$.

   In this case, V is called a basin of attraction[31] for A under Φ.

Now, let φ be a local diffeomorphism of M. A subset A of M is said to be invariant under φ if $\varphi(m) \in A$ for all $m \in A$ which are in the domain of φ. For a point m in the domain of φ and $k \in \mathbb{Z}_+$ we say that $\varphi^k(m)$ is defined if, successively, $\varphi^1(m), \dots, \varphi^{k-1}(m)$ is in the domain of φ.

Definition 3.8.2 (Stability for local diffeomorphisms) A φ-invariant subset $A \subset M$ is called

1. stable under φ if for every neighbourhood U of A in M there exists a neighbourhood V of A in M such that for all $m \in V$ and $k \in \mathbb{Z}_+$, $\varphi^k(m)$ is defined and lies in U,
2. asymptotically stable under φ if it is stable under φ and if there exists a neighbourhood V of A in M with the following properties:
   (a) $\varphi^k(m)$ is defined for all $m \in V$ and all $k \in \mathbb{Z}_+$,
   (b) for every $m \in V$ and every neighbourhood U of A there exists $k_0 \in \mathbb{Z}_+$ such that $\varphi^k(m) \in U$ for all $k \geq k_0$.

   In this case, V is called a basin of attraction for A under φ.

Remark 3.8.3

1. Stability and asymptotic stability are local properties. They do not change if the vector field or the local diffeomorphism is modified outside a neighbourhood of A. As a consequence, if A is compact, without loss of generality one may assume that X is complete. Similarly, in the case of a local diffeomorphism φ,

[31] Beware that in the literature it is common to use this name for the largest such set.

if A is compact and contained in the domain of φ, without loss of generality one may assume that φ is a global diffeomorphism (Exercise 3.8.1).

2. If A is compact, hence in particular if A is a critical integral curve, the definition of (asymptotic) stability may be formulated in terms of a metric ρ compatible with the topology of M. The proof is left to the reader (Exercise 3.8.3). For $B \subset M$, define

$$\tilde{\rho}(x, B) := \inf\{\rho(x, y) : y \in B\}.$$

A Φ-invariant subset A of M is

(a) stable if for every $\varepsilon > 0$ there is $\delta > 0$ such that $\tilde{\rho}(m, A) < \delta$ implies $\mathbb{R}_+ \subset \mathscr{D}_m$ and $\tilde{\rho}(\Phi_t(m), A) < \varepsilon$ for all $t \in \mathbb{R}_+$.

(b) asymptotically stable if there exists $\delta > 0$ such that $\tilde{\rho}(m, A) < \delta$ implies $\mathbb{R}_+ \subset \mathscr{D}_m$ and $\tilde{\rho}(\Phi_t(m), A) \to 0$ for $t \to \infty$.

The primary aim of this section is to derive stability criteria for critical integral curves. This includes

(a) stability criteria for periodic integral curves in terms of the period mapping Φ_T and a Poincaré mapping,

(b) stability criteria for the linearized flow and the discussion of how stability under the linearized flow is related to stability under the flow itself,

(c) a stability criterion in terms of a Lyapunov function.

We start with deriving stability criteria for periodic integral curves in terms of the period mapping and a Poincaré mapping. For that purpose we need

Lemma 3.8.4 *Let A be a Φ-invariant subset and let $t > 0$ such that* $[0, t] \times A \subset \mathscr{D}$. *For every neighbourhood U of A in M there exists a neighbourhood V of A in M such that* $[0, t] \times V \subset \mathscr{D}$ *and* $\Phi_s(V) \subset U$ *for all* $s \in [0, t]$.

Proof Since $\mathscr{D}$ is open in $\mathbb{R} \times M$, for every $m \in A$, there exists an open neighbourhood V_{0m} of m in M such that $\{t\} \times V_{0m} \subset \mathscr{D}$. Then, $V_0 = \bigcup_{m\in A} V_{0m}$ is an open neighbourhood of A in M satisfying $\{t\} \times V_0 \subset \mathscr{D}$ and Proposition 3.2.10/2 implies $[0, t] \times V_0 \subset \mathscr{D}$. By restriction, Φ induces a continuous mapping $\tilde{\Phi}$: $[0, t] \times V_0 \to M$. Let $m \in A$. Due to $\tilde{\Phi}([0, t] \times \{m\}) \subset A \subset U$, $\tilde{\Phi}^{-1}(U)$ is a neighbourhood of $[0, t] \times \{m\}$ in $[0, t] \times V_0$. Since $[0, t]$ is compact, the Tube Lemma of elementary topology yields a neighbourhood V_m of m in V_0 such that $[0, t] \times V_m \subset \tilde{\Phi}^{-1}(U)$. Then, V_m is a neighbourhood of m in M and for $s \in [0, t]$ there holds $\Phi_s(V_m) = \tilde{\Phi}(\{s\} \times V_m) \subset U$. Then, $V := \bigcup_{m\in A} V_m$ has the desired properties. □

Proposition 3.8.5 (Stability of periodic integral curves) *Let γ be a periodic integral curve with period T. Let $m_0 \in \gamma$ and let $(\mathscr{P}, \mathscr{W}, \Theta)$ be a Poincaré mapping for γ at m_0. The following statements are equivalent.*

1. *γ is stable under the flow of X.*
2. *γ is stable under the local diffeomorphism Φ_T of M.*
3. *m_0 is stable under the local diffeomorphism Θ of $\mathscr{P}$.*

The equivalence remains valid if stable is replaced by asymptotically stable.

Proof Since stability is a local concept, we may shrink $\mathscr{P}$ and $\mathscr{W}$ so that there exists $0 < \varepsilon < T$ satisfying conditions 1–4 of Remark 3.7.4. For $U \subset \mathscr{W}$, denote $U^\varepsilon := \Phi_{(-\varepsilon,T+\varepsilon)}(U)$. Since for $m \in \mathscr{W}$ and $k \in \mathbb{Z}_+$ such that $\Theta^k(m)$ is defined there holds $\Theta^1(m), \ldots, \Theta^{k-1}(m) \in \mathscr{W}$, one can put

$$\tau_k(m) := \tau(m) + \tau\big(\Theta(m)\big) + \cdots + \tau\big(\Theta^{k-1}(m)\big).$$

Then, $\Theta^k(m) = \Phi_{\tau_k(m)}(m)$ and $\tau_k(m) \geq k(T-\varepsilon)$.

1 ⇒ 3: Let γ be stable under Φ. Let U be a neighbourhood of m_0 in $\mathscr{P}$. Define $U_0 := \mathscr{W} \cap U$ and $U_1 := U_0 \cap \Theta^{-1}(U_0)$. Condition 3 of Remark 3.7.4 implies

$$U_1^\varepsilon \cap \mathscr{P} \subset U_0. \tag{3.8.1}$$

According to condition 4 of this remark, U_1^ε is a neighbourhood of γ in M. Hence, by assumption, there exists a neighbourhood V_1 of γ such that $\mathbb{R}_+ \times V_1 \subset \mathscr{D}$ and $\Phi_t(m) \in U_1^\varepsilon$ for all $m \in V_1$ and $t \in \mathbb{R}_+$. Define $V := \mathscr{W} \cap V_1$ and let $m \in V$. We show by induction that for all $k \in \mathbb{Z}_+$, $\Theta^k(m)$ is defined and lies in U_0. Due to $U_0 \subset U$, this implies that m_0 is stable under Θ. For $k = 1$, the assertion holds due to $V \subset U_1$ and $\Theta(U_1) \subset U_0$. Thus, assume that $\Theta^k(m)$ is defined and lies in U_0. Due to $U_0 \subset \mathscr{W}$, then $\Theta^{k+1}(m)$ is defined, hence so is $\tau_{k+1}(m)$. Since $\tau_{k+1}(m) \geq 0$, we have $\Theta^{k+1}(m) \equiv \Phi_{\tau_{k+1}(m)}(m) \in U_1^\varepsilon$, hence $\Theta^{k+1}(m) \in U_0$ by (3.8.1).

3 ⇒ 2: Assume that m_0 is stable under Θ. Let U be a neighbourhood of γ in M. According to Lemma 3.8.4, U contains a neighbourhood U_0 of γ in M such that $\Phi_s(U_0) \subset U$ for all $s \in [0, T+\varepsilon]$. Then, $U_1 = U_0 \cap \mathscr{P}$ is a neighbourhood of m_0 in $\mathscr{P}$. By assumption, there exists a neighbourhood V_1 of m_0 in $\mathscr{P}$ such that for all $m \in V_1$ and $l \in \mathbb{Z}_+$, $\Theta^l(m)$ is defined and lies in U_1. By condition 4 of Remark 3.7.4, $V := V_1^\varepsilon$ is a neighbourhood of γ in M. Since Θ^l is defined on V_1 for all $l \in \mathbb{Z}_+$, there holds $\mathbb{R}_+ \times V_1 \subset \mathscr{D}$ and hence $\mathbb{R}_+ \times V \subset \mathscr{D}$, so that Φ_T^k is defined on V for all $k \in \mathbb{Z}_+$. Let $m \in V$ and $k \in \mathbb{Z}_+$. Write $m = \Phi_t(\tilde{m})$ with $\tilde{m} \in V_1$ and $t \in (-\varepsilon, T+\varepsilon)$. There exist $l \in \mathbb{Z}_+$ and $s \in [0, T+\varepsilon]$ such that $kT + t = \tau_l(\tilde{m}) + s$. Then, $\Phi_T^k(m) = \Phi_s(\Theta^l(\tilde{m}))$ and hence $\Phi_T^k(m) \in U$. Thus, γ is stable under Φ_T.

2 ⇒ 1: Let γ be stable under Φ_T. Let U be a neighbourhood of γ in M. According to Lemma 3.8.4, there is a neighbourhood U_0 of γ in M such that $\Phi_s(U_0) \subset U$ for all $s \in [0, T]$. By assumption, for U_0 there exists a neighbourhood V of γ in M such that for all $m \in V$ and $k \in \mathbb{Z}_+$, $\Phi_T^k(m)$ is defined and lies in U_0. That Φ_T^k is defined on V for all $k \in \mathbb{Z}_+$ implies, in particular, $\mathbb{R}_+ \times V \subset \mathscr{D}$. Let $m \in V$ and $t \in \mathbb{R}_+$. Write $t = kT + s$ with $k \in \mathbb{Z}_+$ and $s \in [0, T]$. Then, $\Phi_t(m) = \Phi_s(\Phi_T^k(m)) \in U$. Hence, γ is stable under Φ.

The proof for asymptotic stability is analogous and, therefore, left to the reader. □

Next, we derive stability criteria for the linearized flow along a critical integral curve.

Definition 3.8.6 (Linear stability)

1. A critical integral curve γ of a vector field X is called linearly stable or linearly asymptotically stable if the zero section[32] of $\mathrm{N}\gamma$ is stable or asymptotically stable, respectively, under the linearized flow of X.
2. A fixed point m of a local diffeomorphism φ of M is called linearly stable or linearly asymptotically stable if the origin of $\mathrm{T}_m M$ is stable or asymptotically stable, respectively, under φ'_m.

First, we study the stability of the origin of an abstract finite-dimensional real vector space V under the flow of a linear vector field and under a vector space automorphism. The key ingredient is an appropriate decomposition of the corresponding spectrum. For a linear mapping L on V, let $\mathrm{spec}_0^+(L)$, $\mathrm{spec}_0^-(L)$ and $\mathrm{spec}_0(L)$ denote, respectively, the subsets of eigenvalues with positive, negative and zero real part. Analogously, let $\mathrm{spec}_1^+(L)$, $\mathrm{spec}_1^-(L)$ and $\mathrm{spec}_1(L)$ denote, respectively, the subsets of eigenvalues with absolute value larger, smaller and equal to 1. In addition, for $i = 0, 1$, let $\mathrm{spec}_i^d(L)$ denote the subset of $\mathrm{spec}_i(L)$ of eigenvalues whose geometric and algebraic multiplicities coincide[33] and let $\mathrm{spec}_i^n(L) = \mathrm{spec}_i(L) \setminus \mathrm{spec}_i^d(L)$. For $i = 0, 1$ and $a = +, -, d, n$, let $E_i^a(L)$ and $E_i(L)$ denote the subspace of V spanned by the algebraic eigenspaces[34] of the eigenvalues in $\mathrm{spec}_i^a(L)$ and $\mathrm{spec}_i(L)$, respectively. Thus, one has the disjoint decompositions

$$\begin{aligned} \mathrm{spec}(L) &= \mathrm{spec}_0^-(L) \cup \mathrm{spec}_0(L) \cup \mathrm{spec}_0^+(L), \\ \mathrm{spec}_0(L) &= \mathrm{spec}_0^d(L) \cup \mathrm{spec}_0^n(L), \end{aligned} \tag{3.8.2}$$

$$\begin{aligned} \mathrm{spec}(L) &= \mathrm{spec}_1^-(L) \cup \mathrm{spec}_1(L) \cup \mathrm{spec}_1^+(L), \\ \mathrm{spec}_1(L) &= \mathrm{spec}_1^d(L) \cup \mathrm{spec}_1^n(L), \end{aligned} \tag{3.8.3}$$

and the corresponding direct sum decompositions

$$V = E_0^-(L) \oplus E_0(L) \oplus E_0^+(L), \qquad E_0(L) = E_0^d(L) \oplus E_0^n(L), \tag{3.8.4}$$

$$V = E_1^-(L) \oplus E_1(L) \oplus E_1^+(L), \qquad E_1(L) = E_1^d(L) \oplus E_1^n(L), \tag{3.8.5}$$

where all subspaces in (3.8.4) and (3.8.5) are invariant under L. We need the following two facts from Euclidean geometry.

Lemma 3.8.7 *Let V be a finite-dimensional real vector space.*

1. *Let $A \in \mathrm{End}(V)$ and let $c, d \in \mathbb{R}$ such that $c < \mathrm{Re}(\lambda) < d$ for all $\lambda \in \mathrm{spec}(A)$. There exists a scalar product $\langle\cdot,\cdot\rangle$ on V such that $c\langle v, v\rangle < \langle Av, v\rangle < d\langle v, v\rangle$ for all $v \in V \setminus \{0\}$.*

---

[32] Recall that for an equilibrium $\gamma = \{m\}$, the zero section of $\mathrm{N}\gamma$ reduces to the origin of $\mathrm{T}_m M$.

[33] This means that L is diagonalizable on the corresponding algebraic eigenspace.

[34] That is, $L_{\restriction V_i} - \lambda\, \mathrm{id}_{V_i}$ is nilpotent.

2. *Let* $a \in \mathrm{GL}(V)$ *and let* $c, d > 0$ *such that* $c < |\lambda| < d$ *for all* $\lambda \in \mathrm{spec}(a)$. *There exists a scalar product on* V *whose norm satisfies* $c\|v\| < \|av\| < d\|v\|$ *for all* $v \in V \setminus \{0\}$.

Proof 1. It suffices to show that each algebraic eigenspace of A admits a scalar product with the desired property and to take the orthogonal direct sum of these scalar products. Thus, we may assume that A has a single Jordan block with eigenvalue λ. Choose a basis in V such that the corresponding matrix of A has Jordan normal form, cf. Example 3.2.8. Let $\langle \cdot, \cdot \rangle_0$ denote the scalar product on V making this basis orthonormal. Let $r = \dim V$ in case λ is real and $r = \frac{\dim V}{2}$ otherwise. Choose $0 < \varepsilon < \min\{d - \mathrm{Re}(\lambda), \mathrm{Re}(\lambda) - c\}$ and let T denote the upper triangular $(r \times r)$-matrix with entries $T_{kl} := \varepsilon^{r-l}$ for $k \leq l$. Then,

$$\langle v, w \rangle := \begin{cases} \langle Tv, Tw \rangle_0 & \lambda \in \mathbb{R}, \\ \langle \mathfrak{R}(T)v, \mathfrak{R}(T)w \rangle_0 & \lambda \notin \mathbb{R} \end{cases}$$

defines a scalar product on V with the desired property (Exercise 3.8.4). Here, $\mathfrak{R}(T)$ denotes the realification of complex matrices defined in Example 3.2.8.

2. Let $\rho(a) = \max\{|\lambda| : \lambda \in \mathrm{spec}(a)\}$ denote the spectral radius of a. Due to $\rho(a^{-1})^{-1} = \min\{|\lambda| : \lambda \in \mathrm{spec}(a)\}$, there exist $c_0, d_0 \in \mathbb{R}$ such that $c < c_0 < \rho(a^{-1})^{-1}$ and $\rho(a) < d_0 < d$. Choose some scalar product $\langle \cdot,\cdot \rangle_0$ on V and define

$$\langle v, w \rangle = \sum_{k=1}^{\infty} \big\langle c_0^k a^{-k} v, c_0^k a^{-k} w \big\rangle_0 + \sum_{k=0}^{\infty} \big\langle d_0^{-k} a^k v, d_0^{-k} a^k w \big\rangle_0.$$

To see that the series on the right hand side converge, choose $c_1, d_1 \in \mathbb{R}$ such that $c_0 < c_1 < \rho(a^{-1})^{-1}$ and $\rho(a) < d_1 < d_0$. Use the Cauchy-Schwarz inequality and the fact that for the operator norm associated with the scalar product $\langle \cdot,\cdot \rangle_0$ one has $\rho(a^{-1}) = \lim_{k\to\infty} \|a^{-k}\|^{\frac{1}{k}}$ to estimate these series by geometric series with ratios $q_c = \frac{c_0}{c_1}$ and $q_d = \frac{d_1}{d_0}$, respectively. Finally, renaming summation indices one finds

$$\|av\|^2 = c_0^2 \sum_{k=1}^{\infty} \big\| c_0^k a^{-k} v \big\|_0^2 + c_0^2 \|v\|_0^2 + d_0^2 \sum_{k=1}^{\infty} \big\| d_0^{-k} a^k v \big\|_0^2.$$

Replacing d_0 by c_0 or c_0 by d_0, one obtains a lower estimate or an upper estimate, respectively. Thus, $\|av\|^2 \geq c_0\|v\|^2 > c\|v\|^2$ and $\|av\|^2 \leq d_0\|v\|^2 < d\|v\|^2$. □

Lemma 3.8.8 *Let* V *be a finite-dimensional real vector space.*

1. *For every* $A \in \mathrm{End}(V)$ *there exists an adapted scalar product on* $E_0^{\pm}(A)$ *whose norm satisfies, for all* $v \in E_0^{\pm}(A) \setminus \{0\}$,

$$\begin{aligned} &\big\| e^{\mp tA} v \big\|_{\pm} < \|v\|_{\pm}, \quad t \in \mathbb{R}_+, \\ &\lim_{t\to\pm\infty} \big\| e^{tA} v \big\|_{\pm} = \infty, \qquad \lim_{t\to\mp\infty} \big\| e^{tA} v \big\|_{\pm} = 0. \end{aligned} \tag{3.8.6}$$

2. *For every* $a \in \mathrm{GL}(V)$ *there exists an adapted scalar product on* $E_1^{\pm}(a)$ *whose norm satisfies, for* $v \in E_1^{\pm}(a) \setminus \{0\}$,

$$\big\| a^{\mp 1} v \big\|_{\pm} < \|v\|_{\pm}, \qquad \lim_{k\to\pm\infty} \big\| a^k v \big\|_{\pm} = \infty, \qquad \lim_{k\to\mp\infty} \big\| a^k v \big\|_{\pm} = 0. \tag{3.8.7}$$

Proof 1. First, consider $E_0^-(A)$. There is $d < 0$ such that $\operatorname{Re}\lambda < d$ for all $\lambda \in \operatorname{spec}_0^-(A)$. According to Lemma 3.8.7/1, there is a scalar product $\langle\cdot,\cdot\rangle$ on $E_0^-(A)$ such that $\langle Av, v\rangle < d\|v\|^2$ for all $v \neq 0$. Then,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathrm{e}^{tA}v\| = \frac{\frac{\mathrm{d}}{\mathrm{d}t}\|\mathrm{e}^{tA}v\|^2}{2\|\mathrm{e}^{tA}v\|} = \frac{\langle A\mathrm{e}^{tA}v, \mathrm{e}^{tA}v\rangle}{\|\mathrm{e}^{tA}v\|} < d\|\mathrm{e}^{tA}v\|.$$

There follows $\|\mathrm{e}^{tA}v\| < \mathrm{e}^{dt}\|v\|$ for all $t > 0$ and $\|\mathrm{e}^{tA}v\| > \mathrm{e}^{dt}\|v\|$ for all $t < 0$. Due to $d < 0$, this yields (3.8.6) with the lower signs. The assertion for $E_0^+(A)$ follows by replacing A by $-A$.

2. First, consider $E_1^-(a)$. There is $d < 1$ such that $|\lambda| < d$ for all $\lambda \in \operatorname{spec}_1^-(a)$. According to Lemma 3.8.7/2, there exists a scalar product on $E_1^-(a)$ whose norm satisfies $\|av\| < d\|v\|$ for all $v \neq 0$. Then, $\|a^k v\| < d^k\|v\|$ for all $k \in \mathbb{Z}_+$ and $\|a^k v\| > d^k\|v\|$ for all $k \in \mathbb{Z}_-$. Since $d < 1$, this implies (3.8.7). The assertion for $E_1^+(a)$ follows by replacing a by a^{-1}. □

Lemma 3.8.8 yields the following stability criterion for the origin under a linear flow or a vector space automorphism.

Proposition 3.8.9 (Stability under linear mappings) *Let V be a finite-dimensional real vector space.*

1. *Let $A \in \operatorname{End}(V)$. The origin of V is asymptotically stable under* e^{tA} *if and only if* $\operatorname{spec}(A) = \operatorname{spec}_0^-(A)$. *It is stable under* e^{tA} *iff* $\operatorname{spec}(A) = \operatorname{spec}_0^-(A) \cup \operatorname{spec}_0^d(A)$.
2. *Let $a \in \mathrm{GL}(V)$. The origin of V is asymptotically stable under a if and only if* $\operatorname{spec}(a) = \operatorname{spec}_1^-(a)$. *It is stable under a iff* $\operatorname{spec}(a) = \operatorname{spec}_1^-(a) \cup \operatorname{spec}_1^d(a)$.

Proof 1. Choose scalar products on $E_0^\pm(A)$ according to Lemma 3.8.8. Choose a basis in $E_0(A)$ such that the restriction of A to $E_0(A)$ has Jordan normal form and define the scalar product on $E_0(A)$ such that this basis is orthonormal. Using (3.2.9) and (3.2.10) one can check that

(a) e^{tA} acts isometrically on $E_0^d(A)$,
(b) if $E_0^n(A) \neq 0$, there exists a nonzero $v \in V$ such that $\|\mathrm{e}^{tA}v\| = \sqrt{1+t^2}\|v\|$.

These two observations and (3.8.6) imply the following.

– For all $v \in E_0^-(A)$, $\lim_{t\to\infty}\|\mathrm{e}^{tA}v\| = 0$. Hence, if $V = E_0^-(A)$, 0 is asymptotically stable.
– For all $v \in E_0^-(A) \oplus E_0^d(A)$, $\|\mathrm{e}^{tA}v\| \leq \|v\|$. Hence, if $V = E_0^-(A) \oplus E_0^d(A)$, 0 is stable.
– If $V \neq E_0^-(A)$, there is $v \in V$ with $\|\mathrm{e}^{tA}v\| \geq \|v\|$. Hence, 0 is not asymptotically stable.
– If $V \neq E_0^-(A) \oplus E_0^d(A)$, there is $v \in V$ with $\lim_{t\to\infty}\|\mathrm{e}^{tA}v\| = \infty$. Hence, 0 is not stable.

2. In the argument for point 1, replace the lower index 0 by 1, $t \in \mathbb{R}_+$ by $k \in \mathbb{Z}_+$, e^{tA} by a^k and the reference to (3.8.6) by that to (3.8.7). We have to check that (a)

and (b) remain valid under these replacements. For (a), this is obvious. For (b), it follows by observing that $a^k = \mathrm{diag}(B_1^k, \dots, B_r^k)$, where B_i are the Jordan blocks of a, and by calculating the k-th power of a Jordan block of complex dimension $m_i \geq 2$. If the corresponding eigenvalue λ_i is real, this yields

$$B_i^k = \begin{bmatrix} \lambda_i^k & k\lambda_i^{k-1} & \cdots & \binom{k}{m_i-1}\lambda_i^{k-m_i+1} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & k\,\lambda_i^{k-1} \\ 0 & \cdots & 0 & \lambda_i^k \end{bmatrix}.$$

Otherwise, one has to replace the numbers 0 and λ_i by the 2×2-matrices 0 and $\Re(\lambda_i)$, respectively; see Example 3.2.8 for the notation. In either case, (b) holds true. □

Corollary 3.8.10 (Linear stability)

1. *An equilibrium of a vector field is linearly asymptotically stable if and only if all characteristic exponents have negative real part. It is linearly stable if and only if all characteristic exponents have nonpositive real part and for those with zero real part, the geometric and algebraic multiplicities coincide.*
2. *A periodic integral curve of a vector field or a fixed point of a local diffeomorphism is linearly asymptotically stable if and only if all characteristic multipliers lie in the open unit disk. It is linearly stable if and only if all characteristic multipliers lie in the closed unit disk and for those lying on the unit circle, the geometric and algebraic multiplicities coincide.*

Proof For equilibria of vector fields and fixed points of local diffeomorphisms, the assertion follows directly from Proposition 3.8.9. To prove it for a periodic integral curve γ of a vector field X, choose $m \in \gamma$ and consider the suspension of $\mathrm{N}_m\gamma$ relative to the period automorphism P_m^γ. According to Proposition 3.6.8, the (asymptotic) stability of the zero section of $\mathrm{N}\gamma$ under the linearized flow is equivalent to that of the central integral curve of the suspension of P_m^γ. According to Example 3.7.7/2 and Proposition 3.8.5, the latter is equivalent to the (asymptotic) stability of the origin of $\mathrm{N}_m\gamma$ under P_m^γ. Hence, the assertion follows from Proposition 3.8.9/2 by setting $V = \mathrm{N}_m\gamma$ and $a = P_m^\gamma$. □

Next, we discuss how linear stability is related to stability. A partial answer follows from the Grobman-Hartman Theorem. We cite versions for flows and for local diffeomorphisms.

Theorem 3.8.11 (Grobman-Hartman)

1. *Let X be a vector field on M and let m be an equilibrium point. There exist open neighbourhoods U of m in M and V of the origin in $\mathrm{T}_m M$ such that the flow of $X\restriction_U$ is topologically conjugate to the restriction of the linearized flow $(\Phi_t)'_m$ to V.*

2. *Let φ be a local diffeomorphism of M and let $m \in M$ be a hyperbolic fixed point of φ. There exist open neighbourhoods U of m in M and V of the origin in $\mathrm{T}_m M$ such that the restrictions of φ to U and of φ' to V are topologically conjugate.*

Proof See [236] or [248]. Detailed proofs of assertion 1 can also be found in [123, §IX.7] and [207, §3.6]. □

Let us add that under certain additional conditions on the characteristic exponents or multipliers, respectively, the conjugacy is of class C^r for some r. See [257, §5.8] for examples and further references. The Grobman-Hartman Theorem does not extend beyond hyperbolic critical integral curves, see Example 3.8.14 for a counterexample. It has, however, a generalization to arbitrary critical integral curves, known as the principle of reduction to the centre manifold, see Remark 3.9.13 in the next section.

Corollary 3.8.12 (Stability in the hyperbolic case) *A hyperbolic critical integral curve of a vector field or a hyperbolic fixed point of a local diffeomorphism is stable if and only if it is linearly stable. If it is stable, it is asymptotically stable.*

Proof Since stability is a local property, for equilibria of vector fields and fixed points of local diffeomorphisms, the assertion follows immediately from Theorem 3.8.11. To prove it for a hyperbolic periodic integral curve γ of a vector field X, let $m \in \gamma$ and let us choose a Poincaré mapping $(\mathscr{P}, \mathscr{W}, \Theta)$ for γ at m. According to Proposition 3.8.5, the stability and asymptotic stability of γ under the flow of X is equivalent to the stability and asymptotic stability, respectively, of m under Θ. According to Proposition 3.7.6, the critical multipliers of γ coincide with the eigenvalues of Θ'_m. In particular, m is a hyperbolic fixed point of Θ and the Grobman-Hartman Theorem 3.8.11 implies that m is stable or asymptotically stable under Θ iff so is the origin of $\mathrm{T}_m\mathscr{P}$ under Θ'_m. Then, linear stability follows from Proposition 3.8.9/2 and Corollary 3.8.10/2. □

Remark 3.8.13

1. The Grobman-Hartman Theorem holds for periodic integral curves as well, provided one generalizes the notion of conjugacy by allowing for a reparameterization of the flow parameter.
2. Besides the study of the stability of hyperbolic critical integral curves, there are many other important applications of the Grobman-Hartman Theorem, so that it is actually not quite adequate to treat it in a section on stability. For example, combining Theorem 3.8.11/1 with the classification of linear vector fields generated by a hyperbolic endomorphism up to topological conjugacy, one finds that in the vicinity of a hyperbolic equilibrium point m, the flow of a vector field X is topologically conjugate to the linear flow on $\mathbb{R}^n$ generated by the matrix $\mathbb{1}_{n_+,n_-}$ with $n_\mp = \dim E_0^\mp(\mathrm{Hess}_m(X))$; see [21, §22]. Similarly, combining the Grobman-Hartman Theorem for periodic integral curves with Proposition 3.6.8

and the classification of hyperbolic linear automorphisms up to topological conjugacy, one finds that in the vicinity of a hyperbolic periodic integral curve γ with period T, up to a rescaling of the time parameter, the flow Φ of a vector field is topologically conjugate to the suspension with period 1 of the diagonal real $(n-1)\times(n-1)$-matrix with n_+ entries e and n_- entries e^{-1}, where $n_\pm = \dim E_1^\pm((\Phi_T)'_m)$, $m \in \gamma$, and where the first entry carries the sign of the determinant of $(\Phi_T)'_m$; see [148]. These results provide normal forms for the flow near a hyperbolic critical integral curve.

As a consequence of Corollary 3.8.12, for a critical integral curve of a vector field, linear asymptotic stability implies asymptotic stability, because it implies both linear stability and hyperbolicity. However, neither does linear stability imply stability, nor does asymptotic stability imply linear asymptotic stability. Indeed, while a critical integral curve which is linearly stable but not linearly asymptotically stable (and hence not hyperbolic) may be stable, as in the case of the equilibrium and the periodic integral curves of the frictionless harmonic oscillator, it may as well be asymptotically stable or unstable, as is shown by the following example.

Example 3.8.14 Let $M = \mathbb{R}^2$ and

$$X = \big(-y + xf\big(x^2+y^2\big)\big)\partial_x + \big(x + yf\big(x^2+y^2\big)\big)\partial_y,$$

where $f:\mathbb{R}\to\mathbb{R}$ is a smooth function satisfying $f(0) = f(1) = f'(1) = 0$ and $(s-1)f(s^2) \neq 0$ for $s \neq 0, 1$.[35] The calculations necessary for the following are left to the reader (Exercise 3.8.5).

There exists one equilibrium, the origin, and one periodic integral curve, given by $\gamma(t) = (\cos(t), \sin(t))$. The origin has characteristic exponents $\pm\mathrm{i}$ (hence it is elliptic) and γ has characteristic multiplier 1. In case $(s-1)f(s^2) < 0$ for all $s \neq 0, 1$, the origin is unstable and γ is asymptotically stable. Vice versa, in case $(s-1)f(s^2) > 0$ for all $s \neq 0, 1$, the origin is asymptotically stable and γ is unstable. Let us add that in either case, according to the characteristic exponents, the linearized flow on T_0M is that of a centre (type 4(c) of Example 3.6.13) and hence consists of periodic integral curves. Therefore, it cannot be topologically conjugate to the flow itself.

This completes the discussion of linear stability criteria for critical integral curves. Since these criteria are formulated in terms of characteristic exponents and multipliers, that is, in terms of the spectrum of the associated characteristic linear mapping, they are often referred to as spectral stability criteria. In contrast to that, in the rest of this section, we discuss a functional stability criterion, that is, a criterion which relates stability to the existence of certain functions. Let X be a vector field on M, let $\Phi : \mathscr{D} \to M$ be its flow and let γ be a critical integral curve of X.

Definition 3.8.15 (Lyapunov function) A Lyapunov function for γ consists of an open neighbourhood W of γ and a continuous function $f : W \to \mathbb{R}$ such that

[35] To compare with Example 3.6.15, replace x by $-x$ and set $\omega = 1$ there.

1. $\mathbb{R}_+ \times W \subset \mathscr{D}$,
2. $f(m) = 0$ for all $m \in \gamma$ and $f(m) > 0$ for all $m \in W \setminus \gamma$,
3. $f(\Phi_t(m)) \leq f(m)$ for all $m \in W \setminus \gamma$ and all $t \in \mathbb{R}_+$.

If the inequality in condition 3 holds strictly, (W, f) is called a strong Lyapunov function.

If f is differentiable, for condition 3 to hold it suffices that $X_m(f) \leq 0$ for a Lyapunov function and that $X_m(f) < 0$ for a strong Lyapunov function.

Theorem 3.8.16 (Lyapunov) *Let γ be a critical integral curve of a vector field. If there exists a Lyapunov function for γ, then γ is stable. If there exists a strong Lyapunov function for γ, then γ is asymptotically stable.*

In the case of an asymptotically stable critical integral curve, the converse holds, too: if γ is asymptotically stable, then there exists a strong Lyapunov function, see [44, Thm. V.2.2].

Proof Let $f: W \to \mathbb{R}$ be a Lyapunov function. Let U be a neighbourhood of γ. Since γ is compact, it possesses an open neighbourhood U_0 whose closure $\overline{U_0}$ is compact and satisfies $\overline{U_0} \subset W \cap U$. Let

$$\beta := \min\bigl\{f(m) : m \in \overline{U_0} \setminus U_0\bigr\}, \qquad V := \bigl\{m \in U_0 : f(m) < \beta\bigr\}.$$

Since $\overline{U_0} \setminus U_0$ is compact and since f does not vanish there, β is nonzero. Therefore, $\gamma \subset V$, so that, by continuity of f, V is an open neighbourhood of γ. Since $V \subset W$, by point 1 of Definition 3.8.15, $\mathbb{R}_+ \times V \subset \mathscr{D}$. Let $m \in V$ and $t \in \mathbb{R}_+$. In order to show that $\Phi_t(m) \in U$, it suffices to show that $\Phi_t(m) \in V$. Assume, on the contrary, that $\Phi_t(m) \notin V$. Then, $\Phi_t(m) \notin U_0$, because $f(\Phi_t(m)) < \beta$ by point 3 of the definition. Let $t_0 = \inf\{s \in \mathbb{R}_+ : \Phi_s(m) \notin U_0\}$. Since U_0 is open, there holds $t_0 > 0$ and $\Phi_{t_0}(m) \in \overline{U_0} \setminus U_0$. Then, $f(\Phi_{t_0}(m)) \geq \beta > f(m)$, in contradiction to point 3. Hence, $\Phi_t(m) \in V$ and γ is stable.

Next, assume that f is a strong Lyapunov function. As was just shown, then γ is stable. Construct U_0, β and V as above, choosing $U \equiv W$. We show that V is a basin of attraction for γ. As noted above, $\mathbb{R}_+ \times V \subset \mathscr{D}$. Let $m \in V$. Since $\Phi_t(m) \in V$ for all $t \in \mathbb{R}_+$ and since $\overline{V}$ is compact, the sequence $\{\Phi_n(m)\}$, $n = 1, 2, \ldots$, has a cluster point m_0 in $\overline{V}$. Hence, there exists a strictly increasing sequence $\{n_k\}$, $k = 1, 2, \ldots$, in $\mathbb{N}$ such that $\Phi_{n_k}(m) \to m_0$ for $k \to \infty$. By continuity of f, then $f(\Phi_{n_k}(m)) \to f(m_0)$. Moreover, for $t \in (0, 1)$, $\Phi_{t+n_k}(m_0) \to \Phi_t(m_0)$ and, accordingly, $f(\Phi_{t+n_k}(m)) \to f(\Phi_t(m_0))$. Since $n_k < t + n_k < n_{k+1}$ for all k, property 3 implies

$$\lim_{k\to\infty} f\bigl(\Phi_{n_k}(m)\bigr) \geq \lim_{k\to\infty} f\bigl(\Phi_{t+n_k}(m)\bigr) \geq \lim_{k\to\infty} f\bigl(\Phi_{n_k}(m)\bigr)$$

and hence $f(m_0) = f(\Phi_t(m_0))$. By the strict inequality in property 3, this implies $m_0 \in \gamma$. Thus, for any neighbourhood U of γ, there exists $t_0 \in \mathbb{R}_+$ such that $\Phi_t(m) \in U$ for all $t \geq t_0$. □

Further functional criteria for stability can be found in [44, Ch. V]. In a similar way, functional criteria for instability can be formulated. Functions whose existence implies instability are usually referred to as Chetaev functions.

Example 3.8.17

1. Consider the harmonic oscillator in one dimension with frequency ω, linear friction coefficient $\alpha \geq 0$, elongation x and velocity y, modelled by the vector field (3.6.12) on $\mathbb{R}^2$, see Exercise 3.6.9. Let
$$f(x,y) := \frac{1}{2}\big(y^2 + \omega^2 x^2\big).$$
This corresponds to the mechanical energy per unit mass of the oscillator. There holds $f(0,0) = 0$ and $f(x,y) > 0$ for all $(x,y) \neq (0,0)$. Moreover,
$$X_{(x,y)}(f) = -2\alpha y^2.$$
Thus, in case $\alpha = 0$, f is constant along the integral curves of X and hence a Lyapunov function. In case $\alpha > 0$, f is strictly decreasing and hence a strong Lyapunov function. Hence, the origin is stable for $\alpha = 0$ and asymptotically stable for $\alpha > 0$.
2. Consider the flow given by the suspension with period T of an automorphism a of a finite-dimensional real vector space V. Choose a scalar product on V adapted to a as in the proof of Proposition 3.8.9 and let $\|\cdot\|$ denote the corresponding norm. Define a function $f : V \times \mathbb{R} \to \mathbb{R}$ by
$$f(v,t) := \big(T - \big(t - [t]\big)\big)\big\|a^{[t]}(v)\big\| + \big(t - [t]\big)\big\|a^{[t]+T}(v)\big\|,$$
where $[t]$ denotes the largest integer multiple of T which is smaller than or equal to t. One can show that f descends to a Lyapunov function for the central integral curve γ of the suspension if $\operatorname{spec}(a) = \operatorname{spec}_1^-(a) \cup \operatorname{spec}_1^d(a)$ and to a strong Lyapunov function if $\operatorname{spec}(a) = \operatorname{spec}_1^-(a)$ (Exercise 3.8.6). Then, Theorem 3.8.16 implies that γ is stable in the first case and asymptotically stable in the second case. This result was used in the proof of Corollary 3.8.10/2, where it was obtained from Proposition 3.8.5 and Example 3.7.7/2.

Remark 3.8.18 Let us discuss the relation between the concept of orbital stability presented here and the concept of pointwise stability which is usually referred to as Lyapunov stability. Let X be a vector field on M with flow $\Phi : \mathscr{D} \to M$ and let ρ be a metric on M compatible with the topology. A point $m_0 \in M$ is called

1. Lyapunov stable under Φ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that, for all $m \in M$ and $t \in \mathbb{R}_+$, $\rho(m, m_0) < \delta$ implies $\rho(\Phi_t(m), \Phi_t(m_0)) < \varepsilon$,
2. asymptotically Lyapunov stable if there exists $\delta > 0$ such that, for all $m \in M$, $\rho(m, m_0) < \delta$ implies $\rho(\Phi_t(m), \Phi_t(m_0)) \to 0$ for $t \to \infty$.

While orbital stability is a topological concept, Lyapunov stability, in general, depends on the choice of the metric ρ. For the points of a critical integral curve, however, it is independent of this choice (Exercise 3.8.8). Using the characterization of stability in terms of ρ in Remark 3.8.3/2 one finds that

1. an equilibrium is (asymptotically) stable iff the corresponding equilibrium point is (asymptotically) Lyapunov stable,
2. a periodic integral curve γ is stable if all of its points are Lyapunov stable (Exercise 3.8.9).

The converse of the second statement does not hold in general. A counterexample is given by the planar pendulum, where every periodic integral curve has another period, so that points which are arbitrarily close will separate under time evolution. Moreover, the second statement does not carry over to asymptotic (Lyapunov) stability. In fact, a point of a periodic integral curve cannot be asymptotically Lyapunov stable, because the distance to any other point of this curve is the same after one period. However, here the following converse version of the second statement holds. If a periodic integral curve γ is asymptotically stable and if V is a basin of attraction for γ, then for every $m \in V$ there exists $m_0 \in \gamma$ such that $\rho(\Phi_t(m), \Phi_t(m_0)) \to 0$ for $t \to \infty$, see [9, §I.5.4].

Remark 3.8.19 The concept of stability of a subset of M under the flow of a vector field should be clearly distinguished from the concept of structural stability. A vector field X is said to be structurally stable within a family $\mathscr{X}$ of vector fields and relative to a certain property (P) if there is a neighbourhood[36] of X in $\mathscr{X}$ whose members have property (P). The use of this concept is that property (P) remains valid under small perturbations of the vector field. Analogously, there is a concept of structural stability for local diffeomorphisms.

Exercises

3.8.1 Prove the following.
 (a) If X is a vector field on M with flow $\Phi : \mathscr{D} \to M$ and if $A \subset M$ is a compact subset, then $\mathbb{R} \times A \subset \mathscr{D}$ and there exists an open neighbourhood U of A in M and a complete vector field $\tilde{X}$ on M such that $\tilde{X}_{\restriction U} = X_{\restriction U}$.
 (b) If φ is a local diffeomorphism of M and if A is a compact subset of the domain of φ, there exists an open neighbourhood U of A and a (global) diffeomorphism $\tilde{\varphi}$ of M such that $\tilde{\varphi}_{\restriction U} = \varphi_{\restriction U}$.

3.8.2 Show that a subset A of M which is stable or asymptotically stable under the flow of a vector field on M remains stable if the vector field is modified outside a neighbourhood of A. Prove the analogous assertion for local diffeomorphisms.

3.8.3 Prove the stability criteria in terms of a metric stated in Remark 3.8.3/2.

3.8.4 Complete the proof of Lemma 3.8.7/1.

3.8.5 Carry out the necessary calculations for Example 3.8.14.
Hints. Use Remark 3.6.11/3 to determine the characteristic multipliers of the periodic integral curve γ. Use polar coordinates to show that, for the initial condition $r(0) > 0$, $r(t)$ converges to the following limits when $t \to \infty$.

[36] This requires to have a topology on $\mathscr{X}$. Most often, the family $\mathscr{X}$ is defined by certain parameters, and the topology is thus inherited from the space of parameters.

(a) If $(s-1)f(s^2)<0$ for all $s\neq 0,1$, then $r(t)\to 1$.
(b) If $(s-1)f(s^2)>0$ for all $s\neq 0,1$, then $r(t)\to 0$ for $r(0)<1$ and $r(t)\to\infty$ for $r(0)>1$.

3.8.6 Show that the function f defined in Example 3.8.17/2 descends to a (strong) Lyapunov function on the suspension.

3.8.7 Let $M=\mathbb{R}^2$ and $X=y\partial_x - x(1+xy)\partial_y$. Use the radius function $f(x,y)=x^2+y^2$ to show that the origin is asymptotically stable. In addition, determine the characteristic exponents. (This illustrates once more that asymptotic stability does not imply linear asymptotic stability, see also Example 3.8.14.)

3.8.8 Show that for the points of a critical integral curve, due to compactness, the definition of (asymptotic) Lyapunov stability does not depend on the choice of metric.

3.8.9 Show that a periodic integral curve is stable provided all of its points are Lyapunov stable.

3.8.10 Prove that linear vector fields on $\mathbb{R}^2$ generated by endomorphisms with purely imaginary eigenvalues are topologically conjugate if and only if the eigenvalues coincide. Use this and Remark 3.8.13/2 to classify the linear vector fields on $\mathbb{R}^2$ by topological equivalence. As an application, assign to each of the linear vector fields of Example 3.6.13 its topological equivalence class.

3.9 Invariant Manifolds

In the preceding section we have studied the stability properties of a critical integral curve γ with respect to the flow on a full neighbourhood U of γ. However, in the case where γ is unstable, it is not quite satisfactory to merely know that there exist points in U escaping from U under time evolution. There might as well be points staying in U or approaching γ. Therefore, in this section, we are going to refine the preceding analysis by constructing submanifolds about γ whose points behave in a certain distinguished way under the flow. The starting point of the construction is the decomposition of the tangent space at $m\in\gamma$, given by (3.8.4) and (3.8.5):

$$\mathrm{T}_m M = E_0^-\big(\mathrm{Hess}_m(X)\big)\oplus E_0\big(\mathrm{Hess}_m(X)\big)\oplus E_0^+\big(\mathrm{Hess}_m(X)\big) \tag{3.9.1}$$

if $\gamma=\{m\}$ is an equilibrium and

$$\mathrm{T}_m M = E_1^-\big((\Phi_T)'_m\big)\oplus E_1\big((\Phi_T)'_m\big)\oplus E_1^+\big((\Phi_T)'_m\big) \tag{3.9.2}$$

if γ is periodic. One first establishes the corresponding submanifold structure locally around γ and then extends it globally by means of the flow. Since the local part is intricate and would go beyond the scope of this book, we refer to the literature here. In this section, submanifolds are viewed as subsets, irrespective of whether they are embedded or not, and are allowed to be of class C^k with $k=1,2,\dots,\infty$.

Let X be a vector field on M with flow $\Phi:\mathscr{D}\to M$ and let γ be a critical integral curve of X. Recall that a submanifold S is said to be invariant under Φ if $\Phi_t(m)\in S$ for all $m\in S$ and $t\in\mathscr{D}_m$.

Definition 3.9.1 (Invariant and local invariant manifold) A C^k-submanifold of M is called an invariant manifold for γ if it contains γ and if it is invariant under Φ. It is called a local invariant manifold for γ if it is embedded, contains γ and if X is tangent to it.

Remark 3.9.2 An invariant manifold need not be a local invariant manifold, as it need not be embedded. However, every invariant manifold for γ possesses an open submanifold which is a local invariant manifold for γ. This follows from Remark 1.6.13/1.

Proposition 3.9.3 (Generating invariant manifolds) *For every local invariant manifold S_0 of class C^k for γ, there exists a unique invariant manifold S of class C^k for γ such that*

1. *S_0 is an open submanifold of S,*
2. *for every $m \in S$ there exists $t \in \mathscr{D}_m$ such that $\Phi_t(m) \in S_0$.*

We say that the invariant manifold S is generated by the local invariant manifold S_0. Note that S need not be embedded; this is the price one has to pay for achieving invariance under Φ.

Proof Define $S := \{m \in M : \text{there exists } t \in \mathscr{D}_m \text{ such that } \Phi_t(m) \in S_0\}$. By construction, S is invariant under Φ and satisfies condition 2. To prove the existence of a C^k-submanifold structure, we show that S satisfies condition (S) of Proposition 1.7.1. First, assume that X is complete. Since S_0 satisfies condition (S), there exists a countable family $\{S_{0,i}\}$ of subsets of S_0 covering S_0 and a corresponding family of local C^k-charts $\{(U_i, \kappa_i)\}$ on M such that conditions (S1) and (S2) of Proposition 1.7.1 hold. For every i and every rational t, denote $S_{i,t} := \Phi_t(S_{0,i})$ and $U_{i,t} := \Phi_t(U_i)$ and define the mapping

$$\kappa_{i,t} : U_{i,t} \to \mathbb{R}^n, \qquad \kappa_{i,t}(m) := \kappa_i \circ \Phi_{-t}(m).$$

The family $\{S_{i,t}\}$ is countable. Since $S_0 \subset S$ and since S is invariant under Φ, there holds $S_{i,t} \subset S$. The subsets $S_{i,t}$ cover S: to see this, it suffices to check that for every $m \in S$ there exists a rational number $t \in \mathscr{D}_m$ such that $\Phi_t(m) \in S_0$. Now, since X is tangent to S_0, Remark 3.2.9/2 implies that if $\Phi_t(m) \in S_0$, then $\Phi_s(m) \in S_0$ for s in some open interval containing t. This yields the assertion. Finally, every pair $(U_{i,t}, \kappa_{i,t})$ is a local C^k-chart on M, because $U_{i,t}$ is open and, by restriction, Φ_{-t} induces a diffeomorphism $U_{i,t} \to U_i$. It remains to check that conditions (S1) and (S2) of Proposition 1.7.1 hold for the families $S_{i,t}$ and $(U_{i,t}, \kappa_{i,t})$. Condition (S1) carries over immediately. For condition (S2), let i, j and t, s be given. Then,

$$\kappa_{i,t}(U_{i,t} \cap U_{j,s}) = \kappa_i\big(U_i \cap \Phi_{s-t}(U_j)\big).$$

Since $\Phi_{s-t}(U_j)$ is open in M and hence $U_i \cap \Phi_{s-t}(U_j)$ is open in U_i, the right hand side is open in $\kappa_i(U_i)$. This implies condition (S2).

To remove the assumption that X be complete, first enlarge the original atlas of S_0 by choosing a countable basis of the topology for each chart domain and adding the local charts obtained by restricting the original chart mapping to the elements of this basis. Then, combine each local chart $(U_\alpha, \kappa_\alpha)$ of this new atlas with rational $t \in \bigcap_{m \in U_\alpha} \mathscr{D}_m$ to define subsets $S_{\alpha,t}$ of S and local charts $(U_{\alpha,t}, \kappa_{\alpha,t})$ of M as before. We leave it to the reader to check that these data are well defined, cover S and satisfy conditions (S1) and (S2) (Exercise 3.9.1). This completes the proof of existence of the submanifold structure announced.

To prove uniqueness, let $S^{(1)}$, $S^{(2)}$ be submanifolds of M satisfying the conditions 1 and 2. Due to condition 2, $S^{(1)} = S^{(2)}$ as subsets of M. Thus, to prove that $S^{(1)}$ and $S^{(2)}$ coincide as submanifolds, it suffices to show that the identical mapping $\mathrm{id} : S^{(1)} \to S^{(2)}$ is of class C^k. By Proposition 1.6.10, it suffices to prove continuity, which is equivalent to showing that a neighbourhood in $S^{(1)}$ is also a neighbourhood in $S^{(2)}$. Thus, let $m \in S^{(1)}$ and let W be a neighbourhood of m in $S^{(1)}$. Since $S^{(i)}$ is invariant under Φ, Remark 3.2.9/2 implies that the restriction of Φ to $(\mathbb{R} \times S^{(i)}) \cap \mathscr{D}$ is the flow of the vector field on $S^{(i)}$ induced by X. Hence, the local diffeomorphisms Φ_t of M, $t \in \mathbb{R}$, restrict to local diffeomorphisms of $S^{(i)}$. Using this and condition 1, we find that W may be shrunk so that there exists $t \in \mathbb{R}$ with $\Phi_t(W) \subset S_0$. Since, by condition 1, S_0 carries the relative topology induced from $S^{(1)}$, $\Phi_t(W)$ is a neighbourhood of $\Phi_t(m)$ in S_0. Now, the same arguments, applied in the converse order to $S^{(2)}$, yield that $W = \Phi_{-t}(\Phi_t(W))$ is a neighbourhood of m in $S^{(2)}$. □

The tangent space of a (local) invariant manifold at $m \in \gamma$ is a subspace of $\mathrm{T}_m M$ which is invariant under the Hessian endomorphism $\mathrm{Hess}_m(X)$ in case γ is an equilibrium or under $(\Phi_T)'_m$ in case γ is periodic of period T. On the other hand, the corresponding decomposition (3.9.1) or (3.9.2) is invariant. Thus, it is natural to distinguish (local) invariant manifolds for γ whose tangent spaces at $m \in \gamma$ correspond to a factor or a combination of factors in (3.9.1) or (3.9.2), respectively.

Definition 3.9.4 A (local) invariant manifold S for γ is called a (local) stable, unstable, centre, centre-stable or centre-unstable manifold for γ if the following holds. In case $\gamma = \{m\}$ is an equilibrium, $\mathrm{T}_m S$ coincides with, respectively,

$$E_0^-(A), \quad E_0^+(A), \quad E_0(A), \quad E_0^-(A) \oplus E_0(A), \quad E_0^+(A) \oplus E_0(A),$$

where $A = \mathrm{Hess}_m(X)$. In case γ is periodic with period T, for all $m \in \gamma$, $\mathrm{T}_m S$ coincides with, respectively,[37]

$$E_1^-(a) + \mathrm{T}_m\gamma, \quad E_1^+(a) + \mathrm{T}_m\gamma, \quad E_1(a), \quad E_1^-(a) \oplus E_1(a), \quad E_1^+(a) \oplus E_1(a),$$

where $a = (\Phi_T)'_m$.

The critical integral curve γ of X is also a critical integral curve of the vector field $-X$ and a (local) invariant manifold for γ relative to the flow of X is also a (local) invariant manifold for γ relative to the flow of $-X$. Since on the level of the

[37]Since in the periodic case, the subspaces $E_1^\pm((\Phi_T)'_m)$ do not contain $\mathrm{T}_m\gamma$, the latter has to be added here.

factors in (3.9.1) and (3.9.2), passing from X to $-X$ amounts to interchanging the subspaces E^- and E^+, a (local) stable or centre-stable manifold for γ relative to the flow of X is a (local) unstable or centre-unstable manifold for γ relative to the flow of $-X$, and vice versa. Thus, statements about (local) stable and centre-stable manifolds for γ carry over to (local) unstable and centre-unstable manifolds for γ by passing to the vector field $-X$.

Theorem 3.9.5 (Local existence) *For every critical integral curve γ, there exist*

1. *smooth local stable and unstable manifolds*; *they are locally unique in the sense that the intersection of any two local stable* (*unstable*) *manifolds is an open submanifold of both*,
2. *local centre*, *centre-stable and centre-unstable manifolds of class C^k for every finite k.*

Proof A complete down-to-earth proof can be found in Appendix C of [2] by A. Kelley. It uses a local normal form of X, referred to there as a local pseudochart, which is constructed in §25 and §26 of this book. The standard reference nowadays is [132], which treats invariant manifolds from a more general perspective though and goes far beyond what is needed here. Detailed expositions for the cases of equilibria and hyperbolic periodic integral curves can be found, for example, in [123, §IX.6] or [207, Ch. 4]. □

Let us add that from the proof of Theorem 3.9.5 there follow several intersection properties of the local invariant manifolds, the most important of which is that the local stable, the local unstable and the local centre manifold can be chosen so that their mutual intersections coincide with γ.

Extending the local invariant manifolds of Theorem 3.9.5 by means of Proposition 3.9.3, one obtains the following corollary. We say that $m \in M$ converges to γ under Φ if $\mathbb{R}_+ \subset \mathscr{D}_m$ and if for every neighbourhood U of γ in M there exists $t_0 \in \mathbb{R}_+$ such that $\Phi_t(m) \in U$ for all $t \geq t_0$.

Corollary 3.9.6 (Existence) *For every critical integral curve γ, there exists*

1. *a unique stable manifold $S^-(\gamma)$ and a unique unstable manifold $S^+(\gamma)$ of class C^∞ such that every point of $S^\mp(\gamma)$ converges to γ under the flow of $\pm X$.*
2. *a centre*, *a centre-stable and a centre-unstable manifold of class C^k for every finite k.*

Proof The second assertion follows immediately from Proposition 3.9.3 and Theorem 3.9.5. For the first one it suffices to consider the stable manifold. The assertion about the unstable manifold follows by passing to $-X$. First, we prove existence. By Theorem 3.9.5, there exists a smooth local stable manifold S_0 for γ. Since X is tangent to S_0, by Proposition 2.7.16, it induces a vector field X_0 on S_0. Let $\Phi^0 : \mathscr{D}^0 \to S_0$ denote the flow of X_0. In case $\gamma = \{m\}$ is an equilibrium of X, it is an equilibrium of X_0 and $\mathrm{Hess}_m(X_0)$ is given by the restriction of $\mathrm{Hess}_m(X)$ to the invariant subspace $E_0^-(\mathrm{Hess}_m(X))$. In case γ is periodic of period T under Φ, it is so under Φ^0 and for every $m \in \gamma$, $(\Phi_T^0)'_m$ is given by the restriction of $(\Phi_T)'_m$ to the invariant subspace $E_0^-((\Phi_T)'_m)$. By definition of local stable manifold, this implies that γ is a hyperbolic critical integral curve of X_0 and that it is linearly asymptotically stable under Φ^0, according to Corollary 3.8.10. Then, Corollary 3.8.12 yields that γ is asymptotically stable under Φ^0. This means that there exists an open neighbourhood $\tilde{S}_0$ of γ in S_0 whose points converge to γ under Φ^0. Since, then, they converge to γ under Φ in M, $\tilde{S}_0$ is a local stable manifold for γ. Define $S^-(\gamma)$ to be the invariant manifold generated by $\tilde{S}_0$ in the sense of Proposition 3.9.3. By construction, every point of $S^-(\gamma)$ converges to γ under Φ.

To prove uniqueness, let S^- be a smooth stable manifold for γ whose points converge to γ under Φ. According to Remark 3.9.2, S^- contains an open submanifold $\tilde{S}^-$ which is a local invariant manifold for γ. Since $\mathrm{T}_m\tilde{S}^- = \mathrm{T}_m S^-$ for all $m \in \gamma$, $\tilde{S}^-$ is a local stable manifold for γ. The local uniqueness of local stable manifolds, see Theorem 3.9.5/1, implies that $\tilde{S}^- \cap \tilde{S}_0$ is a local stable manifold for γ and an open submanifold of $\tilde{S}^-$ and $\tilde{S}_0$. Then, it is an open submanifold of S^- and $S^-(\gamma)$. In view of Proposition 3.9.3, this implies $S^- = S^-(\gamma)$ as submanifolds of M. □

The stable and unstable manifolds $S^{\mp}(\gamma)$ are minimal in the sense that every stable manifold for γ contains $S^-(\gamma)$ and every unstable manifold for γ contains $S^+(\gamma)$. As mentioned before, they need not be embedded and also not initial. This is illustrated by Example 3.9.9. Moreover, it is clear from Proposition 3.9.3 that any two centre manifolds for γ coincide if there exists a local centre manifold generating both of them in the sense of that proposition. The same is true for centre-stable and centre-unstable manifolds.

The convergence condition of Corollary 3.9.6/1 defines the subsets $S^{\mp}(\gamma)$ of M only in combination with the property to be a stable or an unstable manifold for γ. Points outside $S^{\mp}(\gamma)$ may converge to γ under the flow of $\pm X$ as well, as is shown by Examples 3.9.10 and 3.9.12 below. If, however, γ is hyperbolic, the subsets $S^{\mp}(\gamma)$ of M can be characterized by this property alone:

Proposition 3.9.7 *If γ is hyperbolic, $S^{\mp}(\gamma)$ consists of the points of M converging to γ under the flow of $\pm X$.*

The characterization given in this proposition is often taken as the definition of the stable and unstable manifolds in the hyperbolic case. Beware that this characterizes $S^{\mp}(\gamma)$ as subsets of M only, and not as submanifolds. To fix the submanifold structure one still has to add the information about the tangent spaces at the points of γ. This is again illustrated by Example 3.9.9, where $S^-(\gamma)$ and $S^+(\gamma)$ coincide as subsets, so that it is only the differentiable structure which distinguishes between them.

Proof As before, it suffices to give the proof for $S^-(\gamma)$. Denote the set of points of M which converge to γ under Φ by S^-. According to Corollary 3.9.6/1, $S^-(\gamma) \subset S^-$. To prove the converse inclusion, we first consider the case that $\gamma = \{m\}$ is an equilibrium. Denote $A = \mathrm{Hess}_m(X)$. According to the Grobman-Hartman Theorem 3.8.11/1, there exist open neighbourhoods U of m in M and V of the origin in $\mathrm{T}_m M$ and a homeomorphism $h : U \to V$ such that the flow $\Phi^U : \mathscr{D}^U \to U$ of $X_{\restriction U}$ satisfies

$$(\Phi_t)'_m \circ h(\tilde{m}) = h \circ \Phi^U_t(\tilde{m}) \tag{3.9.3}$$

for all $(t, \tilde{m}) \in \mathscr{D}^U$. Application of Corollary 3.9.6/1 and Remark 3.9.2 to the flow Φ^U implies that there exists a local stable manifold S^-_U for γ relative to Φ^U whose points converge to m under Φ^U. That is, they converge to m under Φ and their integral curves stay in U for $t \geq 0$. In particular, $S^-_U \subset S^-(\gamma)$.

We show that the subset $h(S^-_U)$ of $\mathrm{T}_m M$ is a neighbourhood of 0 in the subspace $E^-_0(A)$. Due to (3.9.3), the points of $h(S^-_U)$ converge to 0 under $(\Phi_t)'_m$. Using the scalar products on $E^{\mp}_0(A)$ provided by Lemma 3.8.8 and invariance of $E^{\mp}_0(A)$ under $(\Phi_t)'_m = \mathrm{e}^{tA}$ one can check that, in the hyperbolic case, a nonzero element of $\mathrm{T}_m M$ converges to 0 under $(\Phi_t)'_m$ iff it is contained in $E^-_0(A)$. Hence, $h(S^-_U) \subset E^-_0(A)$. Since S^-_U is an embedded submanifold of U, h induces a homeomorphism from S^-_U onto the subset $h(S^-_U)$ of $E^-_0(A)$, equipped with the relative topology. Since S^-_U and $E^-_0(A)$ have the same dimension, the theorem on invariance of domain, in its manifold form,[38] implies that $h(S^-_U)$ is open in $E^-_0(A)$ and hence a neighbourhood of 0 in $E^-_0(A)$, as asserted.

Now, let $m_1 \in S^-$. There exists t_1 such that $\Phi_t(m_1) \in U$ for all $t \geq t_1$. That is, $m_2 := \Phi_{t_1}(m_1)$ converges to m under Φ^U. Then, (3.9.3) implies that $h(m_2)$ converges to 0 under $(\Phi_t)'_m$. As was just noticed, then $h(m_2) \in E^-_0(A)$. Hence, there is $t_2 \in \mathbb{R}_+$ such that $(\Phi_{t_2})'_m(h(m_2)) \in h(S^-_U)$. Then, by (3.9.3) again, $\Phi^U_{t_2}(m_2) = \Phi_{t_1+t_2}(m_1) \in S^-_U$. Since $S^-_U \subset S^-(\gamma)$, it follows that $m_1 \in S^-(\gamma)$. Thus, $S^- \subset S^-(\gamma)$. This proves the proposition in the equilibrium case.

For the periodic case, we extend the terminology concerning convergence to local diffeomorphisms. We say that $m \in M$ converges to m_0 under the local diffeomorphism φ of M if $\varphi^k(m)$ is defined for all $k \in \mathbb{Z}_+$ and $\varphi^k(m) \to m_0$ for $k \to \infty$. Now, assume that γ is periodic of period T. Choose $m \in \gamma$. According to Theorem 3.7.3, there exists a Poincaré mapping $(\mathscr{P}, \mathscr{W}, \Theta)$ for γ at m. Since m is a fixed point of the local diffeomorphism Θ of $\mathscr{P}$ which is hyperbolic by Proposition 3.7.6, the Grobman-Hartman Theorem 3.8.11/2 yields a homeomorphism h from an open neighbourhood of m in $\mathscr{P}$ onto an open neighbourhood V of the origin in $\mathrm{T}_m\mathscr{P}$ which conjugates Θ to Θ'_m. By possibly shrinking $\mathscr{P}$ and $\mathscr{W}$ one may assume that the domain of h coincides with $\mathscr{P}$ and hence there holds

$$\Theta'_m \circ h_{\restriction \mathscr{W}} = h \circ \Theta. \tag{3.9.4}$$

---

[38] Let M, N be smooth manifolds of the same dimension and let $A \subset M$ and $B \subset N$ be subsets which are homeomorphic with respect to the relative topologies induced from M and N, respectively. Then, if A is open in M, B is open in N. See e.g. [199], Chapter VIII, Theorem 6.6 and Exercise 6.5. The reader is encouraged to work out that exercise.

Next, choose a neighbourhood U of γ in M such that $U \cap \mathscr{P} \subset \mathscr{W}$. Let $\Phi^U : \mathscr{D}^U \to U$ denote the flow of $X_{\restriction U}$. Let S_U^- be constructed from U as in the equilibrium case and denote $\tilde{S}_U^- := S_U^- \cap \mathscr{P} \equiv S_U^- \cap \mathscr{W}$. We show that $h(\tilde{S}_U^-)$ is a neighbourhood of 0 in $E_1^-(\Theta_m')$. Due to the following two observations, the argument is completely analogous to the equilibrium case.

1. Repeated application of (3.9.4) yields that the points of $h(\tilde{S}_U^-)$ converge to 0 under Θ_m'.
2. By the definition of Poincaré mapping, $\mathscr{P}$ is transversal to all integral curves of X. Since X is tangent to S_U^-, $\mathscr{P}$ and S_U^- are transversal. According to Corollary 1.8.5, then $\tilde{S}_U^-$ is a submanifold of M; it is embedded, because so are $\mathscr{P}$ and S_U^-, and the dimension formula yields $\dim \tilde{S}_U^- = \dim E_1^-((\Phi_T)_m') = \dim E_1^-(\Theta_m')$.

Now let $m_1 \in S^-$. There is $t_1 \in \mathbb{R}_+$ such that $\Phi_t(m_1) \in U$ for all $t \geq t_1$. One may choose t_1 such that $m_2 := \Phi_{t_1}(m_1) \in \mathscr{W}$. Then, m_2 converges to m under the local diffeomorphism Θ of $\mathscr{P}$. Hence, by (3.9.4), $h(m_2)$ converges to 0 under Θ_m'. Then, $h(m_2) \in E_1^-(\Theta_m')$ and there is $k \in \mathbb{Z}_+$ such that $(\Theta_m')^k(h(m_2)) \in h(\tilde{S}_U^-)$. By (3.9.4), we obtain $\Theta^k(m_2) \in \tilde{S}_U^-$. To k there corresponds $t_2 \in \mathbb{R}_+$ such that $\Theta^k(m_2) = \Phi_{t_2}(m_2) = \Phi_{t_1+t_2}(m_1)$. Since $\tilde{S}_U^- \subset S_U^- \subset S^-(\gamma)$, we get $m_1 \in S^-(\gamma)$. Thus, $S^- \subset S^-(\gamma)$ holds also in the periodic case. □

Next, we give examples and discuss some of the phenomena related to invariant manifolds. We start with linear vector fields.

Example 3.9.8 (Linear vector fields) Let V be a finite-dimensional real vector space, let $A \in \operatorname{End}(V)$ and let X be the linear vector field on V generated by A. Then, $S^\mp(\{0\}) = E_0^\mp(A)$. The subspaces $E_0(A)$, $E_0^-(A) \oplus E_0(A)$ and $E_0^+(A) \oplus E_0(A)$ yield, respectively, a centre, a centre-stable and a centre-unstable manifold for the origin.

Let us continue with the hyperbolic case.

Example 3.9.9 (Planar pendulum) As an example of the stable and unstable manifolds of a hyperbolic equilibrium, consider the upper equilibrium $\gamma = \{m_u\}$ of the planar pendulum of Example 3.6.14. Let $\gamma_{\circlearrowleft}$ and $\gamma_{\circlearrowright}$ denote the two non-critical integral curves, see Fig. 3.2(a). Both $S^+(\gamma)$ and $S^-(\gamma)$ consist of γ, $\gamma_{\circlearrowleft}$ and $\gamma_{\circlearrowright}$. Hence, as subsets of M, they coincide with one another and with the separatrix, which forms a figure eight. The submanifold structures, on the other hand, are distinct, because $S^\mp(\gamma)$ has tangent space $E_0^\mp(\mathrm{Hess}_{m_u}(X))$ at m_u. Thus, $S^-(\gamma)$ and $S^+(\gamma)$ are given by the separatrix, equipped with either one of the two inequivalent figure eight submanifold structures, cf. Example 1.6.6/2. This illustrates that $S^\mp(\gamma)$ need neither be embedded nor initial and that their differentiable structure is quite important, because it is the latter which distinguishes between $S^-(\gamma)$ and $S^+(\gamma)$ here.

As Example 3.9.9 shows, the intersection $S^+(\gamma) \cap S^-(\gamma)$ may contain more than just γ. By invariance under Φ, if the complement $(S^+(\gamma) \cap S^-(\gamma)) \setminus \gamma$ is nonempty, it consists of maximal integral curves. These are called homoclinic. Thus, the integral curves $\gamma_{\circlearrowleft}$ and $\gamma_{\circlearrowright}$ of Example 3.9.9 are homoclinic. Similarly, if γ_1 and γ_2 are two distinct critical integral curves the intersections $S^+(\gamma_1) \cap S^-(\gamma_2)$ and $S^-(\gamma_1) \cap S^+(\gamma_2)$ are invariant under Φ as well. If these intersections are nonempty, the maximal integral curves contained are called heteroclinic. Examples for heteroclinic integral curves can be obtained from $\gamma_{\circlearrowleft}$ and $\gamma_{\circlearrowright}$ of Example 3.9.9 by replacing the angle variable of the planar pendulum by an ordinary Cartesian coordinate, thus producing a 2π-periodic vector field in the plane. Physically, this can be realized for example by adding a counter of rotations to the planar pendulum. To summarize, homoclinic integral curves connect a critical integral curve with itself, whereas heteroclinic integral curves connect different critical integral curves with one another.

Example 3.9.10 (Modified harmonic oscillator) As an example of the stable and unstable manifolds of a hyperbolic periodic integral curve, consider

$$X = \big(y + x\big(1 - x^2 - y^2\big)\big)\partial_x + \big(-x + y\big(1 - x^2 - y^2\big)\big)\partial_y + z\partial_z$$

on $M = \mathbb{R}^3$. This is just the modified harmonic oscillator of Example 3.6.15 with $\omega = 1$ in the x–y-plane, combined with a one-dimensional linear vector field in the z-direction. The equations for the integral curves separate into the equation of the modified harmonic oscillator in the variables x and y and the equation $\dot{z} = z$. Taking the solution of the first one from (3.6.10) we find that X is complete and that, in cylindrical coordinates r, ϕ, z, its integral curves are given by

$$\Phi_t(r, \phi, z) = \Big(\big(1 + \big(r^{-2} - 1\big)\mathrm{e}^{-2t}\big)^{-\frac{1}{2}}, \phi + t, z\mathrm{e}^t\Big). \tag{3.9.5}$$

We read off that the critical integral curves are given by the origin 0 and by the single periodic integral curve γ, given in coordinates by $\gamma(t) = (1, t, 0)$. Let E denote the x–y-plane and let $C_<, C_1, C_>$ denote the subsets defined by $r < 1$, $r = 1$ and $r > 1$ (that is, the open cylinder of radius 1 centered around the z-axis, its boundary and the rest). Analyzing the behaviour of (3.9.5) for $t \to \pm\infty$ one finds

$$S^-\big(\{0\}\big) = \{0\}, \qquad S^+\big(\{0\}\big) = C_<, \qquad S^-(\gamma) = E \setminus \{0\}, \qquad S^+(\gamma) = C_1.$$

All four invariant manifolds are embedded. The whole of E is a stable manifold as well, but it does not satisfy the convergence condition of Corollary 3.9.6/1. Since $S^-(\{0\}) \cap S^+(\{0\}) = \{0\}$ and $S^-(\gamma) \cap S^+(\gamma) = \gamma$, there are no homoclinic integral curves. On the other hand, while $S^-(\{0\}) \cap S^+(\gamma)$ is empty, we get

$$S^-(\gamma) \cap S^+\big(\{0\}\big) = (E \cap C_<) \setminus \{0\},$$

that is, the open disk of radius 1 in E with the origin removed. Thus, there is a continuum of heteroclinic integral curves joining the origin to the periodic integral curve γ.

Next, we turn to non-hyperbolic examples and the discussion of centre manifolds.

Example 3.9.11 Consider the vector field

$$X = y\partial_x - x\partial_y - z\partial_z$$

on $M = \mathbb{R}^3$. This is the harmonic oscillator (3.6.12) with $\omega = 1$ and $\alpha = 0$ in the x–y-plane, combined with a one-dimensional linear vector field in the z-direction. The integral curves are given in cylindrical coordinates r, ϕ, z by

$$\Phi_t(r, \phi, z) = \big(r, \phi + t, z\mathrm{e}^{-t}\big).$$

The set of critical integral curves consists of the origin and of the periodic integral curves γ_r, $r > 0$, given in coordinates by $\gamma_r(t) = (r, \omega t, 0)$. Let C_r denote the surface of the cylinder of radius r centered around the z-axis. The stable and unstable manifolds are

$$S^-\big(\{0\}\big) = z\text{-axis}, \qquad S^+\big(\{0\}\big) = \{0\}, \qquad S^-(\gamma_r) = C_r, \qquad S^+(\gamma_r) = \gamma_r.$$

There are neither homoclinic nor heteroclinic integral curves. The x–y-plane provides a centre manifold for all critical integral curves. Similarly, any open disk centered at the origin in this plane is a centre manifold for the origin and any open annulus containing γ_r is a centre manifold for γ_r.

While Example 3.9.11 shows that a centre manifold cannot be made unique by requiring convergence properties like for $S^{\mp}(\gamma)$, any two centre manifolds still intersect in an open submanifold there, that is, they coincide locally. The next example illustrates that it may happen as well that there exist centre manifolds which do not coincide locally.

Example 3.9.12 Consider the vector field

$$X = x^2\partial_x - y\partial_y$$

on $M = \mathbb{R}^2$. This is Example 3.2.7/2, combined with a linear vector field in one dimension. According to that example, X is not complete and the flow is given by

$$\mathscr{D} = \left\{\big(t, (x, y)\big) \in \mathbb{R} \times M : t < \frac{1}{x}\right\}, \qquad \Phi_t(x, y) = \left(\frac{x}{1 - tx}, y\mathrm{e}^{-t}\right), \qquad (3.9.6)$$

see Fig. 3.3. The only critical integral curve is the origin. The Hessian is

$$\mathrm{Hess}_0(X) = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix},$$

hence $E_0^-(\mathrm{Hess}_0(X))$ is given by the y-axis, $E_0(\mathrm{Hess}_0(X))$ coincides with the x-axis and $E_0^+(\mathrm{Hess}_0(X)) = \{0\}$. It follows that $S^-(\{0\})$ is given by the y-axis, whereas $S^+(\{0\}) = \{0\}$. Since all the points of the closed half-plane $x \leq 0$ converge to the origin under Φ, this illustrates that Proposition 3.9.7 fails to hold in the non-hyperbolic case. Next, let us look for centre manifolds. In the open half-plane $x > 0$, any centre manifold must coincide with the positive x-axis. In the open half-plane $x < 0$, by eliminating t either from the differential equations for the flow or from

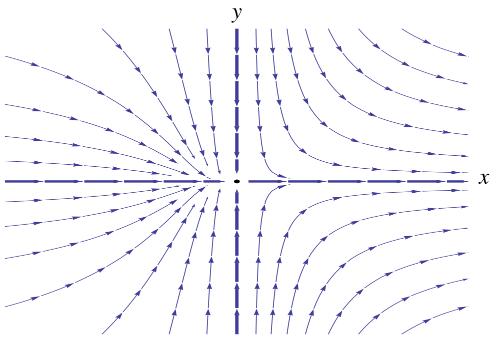


Fig. 3.3 The flow of Example 3.9.12

(3.9.6) one finds that the images of integral curves coincide with the graphs of the functions

$$y = f_\mu(x) = \mu e^{\frac{1}{x}}, \quad \mu \in \mathbb{R}.$$

Since derivatives of f_μ with respect to x to arbitrary order tend to zero as $x \to 0$ from below, f_μ can be extended to a smooth function on $\mathbb{R}$ by letting $f_\mu(x) = 0$ for $x \geq 0$. According to Example 1.6.12/2, the graph of this function is a smooth submanifold. Due to $f'_\mu(0) = 0$, the tangent space of this submanifold at $x = 0$ coincides with the x-axis and hence with $E_0(\mathrm{Hess}_0(X))$. Thus, all of these submanifolds are center manifolds, but neither two of them coincide locally, that is, the equilibrium is not an inner point of the intersection.

While the centre manifolds of Examples 3.9.11 and 3.9.12 are all smooth, this need not be so in general. For counter-examples, see e.g. [296] or [285]. Of course, the above statements about non-uniqueness and non-smoothness of (local) centre manifolds carry over to centre-stable and centre-unstable manifolds.

Let us, furthermore, mention that there exist approximation methods for a systematic construction of centre manifolds, see for example [59] or the second part of Sect. 3.2 in [114].

Remark 3.9.13 (Reduction to the centre manifold) Using centre manifolds, the Grobman-Hartman Theorem can be generalized to arbitrary equilibria as follows. Let $\gamma = \{m\}$ be an equilibrium of X. Let $X^\mp$ denote the linear vector fields induced by $\mathrm{Hess}_m(X)$ on $E_0^\mp(\mathrm{Hess}_m(X))$. For every $k = 1, 2, 3, \ldots$, there exists a centre manifold S^0 for γ of class C^k, as well as open neighbourhoods U of m in M, $V^\mp$ of the origin in $E_0^\mp(\mathrm{Hess}_m(X))$ and V^0 of m in S^0 such that the flow of $X_{\restriction U}$ is topologically conjugate to the flow of the vector field $(X^-_{\restriction V^-}, X^0_{\restriction V^0}, X^+_{\restriction V^+})$ on $V^- \times V^0 \times V^+$, where X^0 denotes the vector field on S^0 induced by restriction of X. See [250, 271] or [17, §32] for a proof and [207, §4.2] for a more detailed discussion. There holds an analogous statement for fixed points of local diffeomorphisms [131, Thm. 7.3] and for periodic integral curves [250].

Thus, for the qualitative analysis of a flow in the vicinity of a critical integral curve it suffices to study the linearization of the flows induced on the minimal stable and unstable manifolds $S^{\mp}(\gamma)$ and the flow induced on an appropriate centre manifold. As already mentioned in Example 3.8.14, the latter flow is in general not topologically conjugate to its linearization.

To conclude this section, let us mention that invariant manifolds for fixed points of local diffeomorphisms can be treated in a completely analogous way. For generalizations of the concept of stable, unstable and centre manifold from critical integral curves to certain classes of submanifolds, see [132]. A well-organized guide to all of this, with hints for further reading, is [257].

Exercises

3.9.1 Complete the proof of the existence of a smooth submanifold structure on the subset generated by a local invariant manifold in Proposition 3.9.3 for the case of a non-complete vector field.

3.9.2 Let γ be a periodic integral curve of period T of the flow Φ. Show that the subspaces

$$E_0^{\mp}\big((\Phi_T)'_m\big), \qquad E_0\big((\Phi_T)'_m\big), \qquad E_0^{\mp}\big((\Phi_T)'_m\big) \oplus E_0\big((\Phi_T)'_m\big), \quad m \in \gamma,$$

combine to smooth vector subbundles of $\mathrm{T}\gamma$ which are invariant under the linearized flow. Reformulate the definition of stable, unstable, centre, centre-stable and centre-unstable manifold for γ in terms of these subbundles.

3.9.3 Determine the flow of the vector field $X = -x\partial_x + (y + x^2)\partial_y$ on $\mathbb{R}^2$ and use this to find the stable and unstable manifolds of the origin.

3.9.4 Consider the vector field $X = y\partial_x + (x - x^3)\partial_y$ on $\mathbb{R}^2$. This models a one-dimensional quartic oscillator (Mexican-hat potential). The associated second order differential equation is a special version of the autonomous Duffing equation.
   (a) Find the equilibrium points and their characteristic exponents.
   (b) Show that $H(x, y) = \frac{1}{2}y^2 - \frac{1}{2}x^2 + \frac{1}{4}x^4$ is a first integral of X, that is, there holds $X(H) = 0$.
   (c) Use H to determine the stable and unstable manifolds of the origin. Compare the result with that for the upper equilibrium of the planar pendulum in Example 3.9.9.

3.9.5 Carry out an analysis similar to Exercise 3.9.4 for

$$X = y\partial_x + x\left(\frac{2}{\sqrt{1+x^2}} - 1\right)\partial_y, \qquad H(x, y) = \frac{1}{2}y^2 + \frac{1}{2}\big(\sqrt{1+x^2} - 2\big)^2.$$

This models a particle which moves without friction along a rod and is attached to a fixed point outside the rod by a spring, whose equilibrium length is twice the distance of the fixed point from the rod.

Chapter 4
Differential Forms

We first present the elementary calculus of differential forms, including the calculus of integration and a discussion of integral invariants. Then, in Sect. 4.3, we give an introduction to de Rham cohomology. Next, in Sects. 4.4 and 4.5, we present some elements of Riemannian geometry, discuss Hodge duality in detail and show how classical vector analysis can be understood in a coordinate-free way using the language of differential forms. In Sect. 4.6, we apply this framework to classical Maxwell electrodynamics. In Sect. 4.7, we give an introduction to the theory of Pfaffian systems and differential ideals. In particular, we derive an equivalent formulation of the classical Frobenius Theorem. Finally, we apply these notions to classical mechanics with constraints.

4.1 Basics

Recall the following notions from Sect. 2.5. A differential k-form on a manifold M is a section in the vector bundle $\Lambda^k \mathrm{T}^* M$. We write $\Omega^k(M)$ for the space of differential k-forms and

$$\Omega^*(M) = \Gamma\Big(\bigwedge \mathrm{T}^* M\Big) \equiv \bigoplus_{k=0}^{\infty} \Omega^k(M)$$

for the exterior algebra. Obviously, $\Omega^0(M) = C^\infty(M)$ and $\Omega^k(M) = 0$ for $k > \dim M$. The exterior product $\wedge\colon \Omega^k(M) \times \Omega^l(M) \to \Omega^{k+l}(M)$ is given by

$$\begin{aligned}
&(\alpha \wedge \beta)(X_1, \ldots, X_{k+l}) \\
&\quad = \frac{1}{k!l!} \sum_{\pi \in \mathrm{S}_{k+l}} \operatorname{sign}(\pi) \alpha(X_{\pi(1)}, \ldots, X_{\pi(k)}) \beta(X_{\pi(k+1)}, \ldots, X_{\pi(k+l)}),
\end{aligned}$$

where $\alpha \in \Omega^k(M)$, $\beta \in \Omega^l(M)$ and $X_i \in \mathfrak{X}(M)$, cf. (2.4.17). The natural pairing of k-forms with k-vectors from multilinear algebra induces a C^∞-valued pairing of

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7_4,

k-differential forms with k-vector fields, given by

$$\langle \alpha_1 \wedge \cdots \wedge \alpha_k, X_1 \wedge \cdots \wedge X_k \rangle = \det\big(\alpha_i(X_j)\big) \tag{4.1.1}$$

for all $\alpha_i \in \Omega^1(M)$ and $X_i \in \mathfrak{X}(M)$. Then, by the definition of the exterior product,

$$\alpha(X_1, \ldots, X_k) = \langle \alpha, X_1 \wedge \cdots \wedge X_k \rangle$$

for all $\alpha \in \Omega^k(M)$ and $X_i \in \mathfrak{X}(M)$. For $k \geq r$, we define the operation of inner multiplication of an r-vector field X with a k-differential form α by

$$\langle X \lrcorner \alpha, Y \rangle := \langle \alpha, X \wedge Y \rangle, \quad Y \in \mathfrak{X}^{k-r}(M). \tag{4.1.2}$$

This operation is $C^\infty(M)$-linear in both arguments.

Let us describe the above structures in a local chart (U, κ) on M. (U, κ) induces local frames $\{\partial_i\}$ and $\{\mathrm{d}\kappa^i\}$ of $\mathrm{T}M$ and T^*M, respectively, and by Example 2.4.5, the induced local frames in $\bigwedge^r \mathrm{T}M$ and $\bigwedge^r \mathrm{T}^*M$ consist, respectively, of the local sections

$$\partial_{i_1} \wedge \cdots \wedge \partial_{i_r}, \qquad \mathrm{d}\kappa^{i_1} \wedge \cdots \wedge \mathrm{d}\kappa^{i_r}, \quad 1 \leq i_1 < \cdots < i_r \leq n.$$

It is common to use the following condensed notation. For a subset $I \subset \{1, \ldots, n\}$ of r elements define

$$\partial_I := \partial_{i_1} \wedge \cdots \wedge \partial_{i_r}, \qquad \mathrm{d}\kappa^I := \mathrm{d}\kappa^{i_1} \wedge \cdots \wedge \mathrm{d}\kappa^{i_r},$$

where $i_1, \ldots, i_r$ denote the elements of I, ordered by magnitude, that is, $I = \{i_1, \ldots, i_r\}$ and $i_1 < \cdots < i_r$. Thus, the local frames under consideration consist, respectively, of the local sections ∂_I and $\mathrm{d}\kappa^I$, where I runs through the subsets of $\{1, \ldots, n\}$ of cardinality r. In particular, by an extension of the summation convention, pairs of capital indices I are summed over those subsets. We have

$$\big\langle \mathrm{d}\kappa^I, \partial^\kappa_J \big\rangle = \delta^I_J, \tag{4.1.3}$$

$$\mathrm{d}\kappa^I \wedge \mathrm{d}\kappa^J = \rho_{I,J} \mathrm{d}\kappa^{I \cup J} \delta_{I \cap J, \varnothing}, \tag{4.1.4}$$

$$\partial^\kappa_J \lrcorner \mathrm{d}\kappa^I = \rho_{J, I \setminus J} \mathrm{d}\kappa^{I \setminus J} \delta_{I \cup J, I}, \tag{4.1.5}$$

where $\delta^I_J = 1$ if $I = J$ and 0 otherwise, and $\rho_{I,J} = (-1)^q$, where q is the number of pairs $(i, j) \in I \times J$ with $i > j$ (Exercise 4.1.1). Then, locally, $\alpha \in \Omega^r(M)$ is represented by

$$\alpha_{\restriction U} = \alpha_I \mathrm{d}\kappa^I \equiv \sum_{i_1 < \cdots < i_r} \alpha_{i_1 \ldots i_r} \mathrm{d}\kappa^{i_1} \wedge \cdots \wedge \mathrm{d}\kappa^{i_r}, \quad \alpha_{i_1 \ldots i_r} = \alpha(\partial_{i_1}, \ldots, \partial_{i_r}), \tag{4.1.6}$$

cf. (2.4.21), and r-vector fields X are represented by

$$X_{\restriction U} = \sum_J X^J \partial^\kappa_J \equiv \sum_{i_1 < \cdots < i_r} X^{i_1 \ldots i_r} \partial_{i_1} \wedge \cdots \wedge \partial_{i_r}. \tag{4.1.7}$$

As a consequence, the operations of exterior product and inner multiplication take the local form

$$(\alpha \wedge \beta)_{\restriction U} = \alpha_{\restriction U} \wedge \beta_{\restriction U} = \sum_I \sum_{J \cap I = \varnothing} \alpha_I \beta_J \rho_{I,J} \mathrm{d}\kappa^{I \cup J}. \tag{4.1.8}$$

$$(X \lrcorner \alpha)_{\restriction U} = X_{\restriction U} \lrcorner \alpha_{\restriction U} = \sum_I \sum_{J \subset I} \alpha_I X^J \rho_{J, I \setminus J} \mathrm{d}\kappa^{I \setminus J}. \tag{4.1.9}$$

Remark 4.1.1

1. Equation (4.1.6) can be used to define functions $\alpha_{i_1 \ldots i_r}$ for all sequences $i_1, \ldots, i_r$, and it is sometimes more convenient to sum over all these sequences rather than just the increasing ones. Since both $\alpha_{i_1 \ldots i_r}$ and $\mathrm{d}\kappa^{i_1} \wedge \cdots \wedge \mathrm{d}\kappa^{i_r}$ are antisymmetric under a permutation of $i_1, \ldots, i_r$, this produces a factor $r!$. Hence,

$$\alpha_{\restriction U} = \frac{1}{r!} \alpha_{i_1 \ldots i_r} \mathrm{d}\kappa^{i_1} \wedge \cdots \wedge \mathrm{d}\kappa^{i_r}, \tag{4.1.10}$$

(summation convention).[1] Let $m \in U$. By an appropriate extension of the functions $\alpha_{i_1 \ldots i_r}$ and κ^i from a smaller neighbourhood $\tilde{U} \subset U$ of m to smooth functions $\tilde{\alpha}_{i_1 \ldots i_r}$ and $\tilde{\kappa}^i$ on M which vanish outside U, we obtain

$$\alpha_{\restriction \tilde{U}} = \frac{1}{r!} \left(\tilde{\alpha}_{i_1 \ldots i_r} \mathrm{d}\tilde{\kappa}^{i_1} \wedge \cdots \wedge \mathrm{d}\tilde{\kappa}^{i_r} \right)_{\restriction \tilde{U}}. \tag{4.1.11}$$

Thus, locally, every k-form is a sum of k-forms of the type $f_0 \, \mathrm{d} f_1 \wedge \cdots \wedge \mathrm{d} f_k$ with $f_i \in C^\infty(M)$.
2. We determine the transformation laws for the local frames and for the corresponding coefficient functions under a change of local chart. Thus, let (V, ρ) be another local chart on M and let us denote

$$A^i_j := \left[\left(\rho \circ \kappa^{-1} \right)' \circ \kappa \right]^i_j, \qquad \tilde{A}^i_j := \left[\left(\kappa \circ \rho^{-1} \right)' \circ \rho \right]^i_j$$

and

$$A^I_J := \sum_{\pi \in \mathrm{S}_r} \mathrm{sign}(\pi) A^{i_1}_{j_{\pi(1)}} \cdots A^{i_r}_{j_{\pi(r)}}, \qquad \tilde{A}^I_J := \sum_{\pi \in \mathrm{S}_r} \mathrm{sign}(\pi) \tilde{A}^{i_{\pi(1)}}_{j_1} \cdots \tilde{A}^{i_{\pi(r)}}_{j_r},$$

for $I = \{i_1, \ldots, i_r\}$ and $J = \{j_1, \ldots, j_r\}$. Then, over $U \cap V$ the following formulae hold (Exercise 4.1.2):[2]

$$\partial^\rho_I = \tilde{A}^J_I \partial^\kappa_J, \qquad \mathrm{d}\rho^I = A^I_J \mathrm{d}\kappa^J, \tag{4.1.12}$$

[1] We caution the reader that for $r \geq 2$ this sum runs over a linearly dependent system of sections. In particular, the left hand side does not determine the coefficient functions on the right hand side uniquely. There is however a unique choice if we limit our attention to functions which are totally antisymmetric in the indices $i_1, \ldots, i_r$, and this choice is given by (4.1.6).

[2] One may also write, for example, $\mathrm{d}\rho^{i_1} \wedge \cdots \wedge \mathrm{d}\rho^{i_r} = A^{i_1}_{j_1} \cdots A^{i_r}_{j_r} \mathrm{d}\kappa^{j_1} \wedge \cdots \wedge \mathrm{d}\kappa^{j_r}$, keeping in mind that on the right hand side the sum runs over a linearly dependent system rather than a basis.

and (4.1.6) and (4.1.7) imply

$$\alpha_I^\rho = \tilde{A}_I^J \alpha_J^\kappa, \qquad X^{\rho,I} = A_J^I X^{\kappa,J}. \tag{4.1.13}$$

Next, let M and N be manifolds and let $\varphi : M \to N$ be a smooth mapping. According to Example 2.4.5, φ induces a linear mapping $\varphi^* : \Omega^*(N) \to \Omega^*(M)$, called the pull-back, by

$$\big(\varphi^*\alpha\big)_m(X_1, \dots, X_r) = \alpha_{\varphi(m)}\big(\varphi'_m(X_1), \dots, \varphi'_m(X_r)\big), \tag{4.1.14}$$

where $\alpha \in \Omega^r(N)$ and $X_i \in \mathrm{T}_m M$. This mapping is an algebra homomorphism, that is,

$$\varphi^*(\alpha \wedge \beta) = \big(\varphi^*\alpha\big) \wedge \big(\varphi^*\beta\big) \tag{4.1.15}$$

for all $\alpha, \beta \in \Omega^*(M)$. If φ is a diffeomorphism, (4.1.14) entails

$$\big(\varphi^*\alpha\big)(X_1, \dots, X_r) = \varphi^*\big(\alpha(\varphi_* X_1, \dots, \varphi_* X_r)\big) \tag{4.1.16}$$

for all $X_i \in \mathfrak{X}(M)$. In local charts (U, κ) on M and (V, ρ) on N, the pull-back is represented by

$$\big(\varphi^*\alpha\big)_{i_1 \dots i_k}(m) = \alpha_{j_1 \dots j_k}\big(\varphi(m)\big) \frac{\partial y^{j_1}}{\partial x^{i_1}}\big(\kappa(m)\big) \dots \frac{\partial y^{j_k}}{\partial x^{i_k}}\big(\kappa(m)\big), \tag{4.1.17}$$

where $y^j = \rho^j \circ \varphi \circ \kappa^{-1}$ (Exercise 4.1.3).

After these introductory remarks, we turn to the discussion of the exterior derivative.

Definition 4.1.2 (Exterior derivative) A system of linear mappings

$$\mathrm{d}\colon \Omega^k(M) \to \Omega^{k+1}(M), \quad k = 0, 1, \dots,$$

is called an exterior derivative on M if

1. $\mathrm{d} \circ \mathrm{d} = 0$,
2. $\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^k \alpha \wedge \mathrm{d}\beta$ for all $\alpha \in \Omega^k(M)$ and $\beta \in \Omega^*(M)$,
3. $\langle \mathrm{d}f, X \rangle = X(f)$ for all $f \in C^\infty(M)$ and $X \in \mathfrak{X}(M)$.

A linear mapping of $\Omega^*(M)$ satisfying condition 2 is called an anti-derivation. Condition 3 means that the exterior derivative of a function f coincides with the ordinary differential of f, defined in Example 2.5.1. This yields existence and uniqueness for the exterior derivative of functions. Moreover, this implies that taking the exterior derivative of a function is a local operation in the following sense:

$$(\mathrm{d}f)_{\restriction U} = \mathrm{d}(f_{\restriction U}) \tag{4.1.18}$$

for all $f \in C^\infty(M)$ and all open subsets U of M, where on the right hand side, d stands for the exterior derivative of functions on the manifold U. Indeed, for every $X \in \mathfrak{X}(M)$ we have $X_{\restriction U} \in \mathfrak{X}(U)$ and $(X(f))_{\restriction U} = X_{\restriction U}(f_{\restriction U})$, and condition 3 implies

$$\big\langle (\mathrm{d}f)_{\restriction U}, X_{\restriction U}\big\rangle = \langle \mathrm{d}f, X\rangle_{\restriction U} = \big(X(f)\big)_{\restriction U} = X_{\restriction U}(f_{\restriction U}) = \big\langle \mathrm{d}(f_{\restriction U}), X_{\restriction U}\big\rangle.$$

Lemma 4.1.3 *An exterior derivative is a local operator, that is, if differential forms α and β coincide on a neighbourhood of a point $m \in M$, then* $\mathrm{d}\alpha(m) = \mathrm{d}\beta(m)$.

Proof Let U be an open neighbourhood of m such that $\alpha_{\restriction U} = \beta_{\restriction U}$. There exists a smaller neighbourhood $\tilde{U} \subset U$ of m and a smooth function f on M such that $f_{\restriction \tilde{U}} = 1$ and $f_{\restriction M\setminus U} = 0$. Then, $f\alpha = f\beta$ on M and hence, by property 2,

$$\mathrm{d}f \wedge \alpha + f\mathrm{d}\alpha = \mathrm{d}f \wedge \beta + f\mathrm{d}\beta. \tag{4.1.19}$$

Since $\mathrm{d}f(m) = 0$ (by property 3) and $f(m) = 1$, (4.1.19) yields the assertion. □

Theorem 4.1.4 (Existence and uniqueness) *On every manifold M there exists a unique exterior derivative. It satisfies*

$$(\mathrm{d}\alpha)_{\restriction U} = \sum_{I,p} \big(\partial_p^\kappa \alpha_I\big)\mathrm{d}\kappa^p \wedge \mathrm{d}\kappa^I \tag{4.1.20}$$

for all $\alpha \in \Omega^k(M)$ and all local charts (U, κ) on M.

Proof First, we show existence. Let $\alpha \in \Omega^k(M)$. We choose a local chart (U,κ) and define $\mathrm{d}\alpha$ pointwise on U by (4.1.20). We must show that this definition does not depend on the choice of the local chart. Thus, let $(\tilde{U}, \tilde{\kappa})$ be another chart. We may assume $U = \tilde{U}$. A simple calculation using the transformation laws (4.1.12) and (4.1.13) yield

$$\sum_{I,p} \big(\partial_p^{\tilde{\kappa}} \tilde{\alpha}_I\big)\mathrm{d}\tilde{\kappa}^p \wedge \mathrm{d}\tilde{\kappa}^I = \sum_{I,p} \big(\partial_p^\kappa \alpha_I\big)\mathrm{d}\kappa^p \wedge \mathrm{d}\kappa^I. \tag{4.1.21}$$

We leave it to the reader to check that the system of linear mappings d so defined is an exterior derivative (Exercise 4.1.5). Next, we prove uniqueness. Let d and d′ be two exterior derivatives on M and let $\alpha \in \Omega^k(M)$. We show that $(\mathrm{d}\alpha)_m = (\mathrm{d}'\alpha)_m$ for all $m \in M$. By Formula (4.1.11), there exists a neighbourhood of m where α coincides with a sum of terms of the form $f_0\mathrm{d}f_1 \wedge \cdots \wedge \mathrm{d}f_k$. Hence, by Lemma 4.1.3, it suffices to show that

$$\mathrm{d}(f_0\mathrm{d}f_1 \wedge \cdots \wedge \mathrm{d}f_k)(m) = \mathrm{d}'(f_0\mathrm{d}f_1 \wedge \cdots \wedge \mathrm{d}f_k)(m).$$

This follows from the uniqueness of the exterior derivative on functions and from the defining properties 1 and 2. □

Remark 4.1.5 By Lemma 4.1.3 and by the uniqueness of d, for every open subset $U \subset M$ and every k-form α, we have

$$(\mathrm{d}\alpha)_{\restriction U} = \mathrm{d}(\alpha_{\restriction U}).$$

The exterior derivative can be characterized in a coordinate-free way as follows.

Proposition 4.1.6 *For $\alpha \in \Omega^k(M)$ and $X_1, \ldots, X_k \in \mathfrak{X}(M)$,*

$$\begin{aligned} \mathrm{d}\alpha(X_0, \ldots, X_k) = {} & \sum_{i=0}^{k} (-1)^i X_i\big(\alpha(X_0, \overset{X_i}{\overset{\curlyvee}{\ldots}}, X_k)\big) \\ & + \sum_{i<j} (-1)^{i+j} \alpha\big([X_i, X_j], X_0, \overset{X_i X_j}{\overset{\curlyvee\ \curlyvee}{\ldots\ldots}}, X_k\big), \end{aligned} \tag{4.1.22}$$

where $\overset{X_i}{\curlyvee}$ means that X_i is omitted.

Proof First, we show that both sides of the above equation are f-linear with respect to every X_j. For the left hand side, this is obvious. Denoting the two terms on the right hand side by $T_i(X_0, \ldots, X_k)$, $i = 1, 2$, and using the derivation property we obtain

$$\begin{aligned} T_1(X_0, \ldots, fX_l, \ldots, X_k) = {} & \sum_{i \neq l} (-1)^i X_i\big(\alpha(X_0, \ldots, fX_l, \overset{X_i}{\overset{\curlyvee}{\ldots}}, X_k)\big) \\ & + (-1)^l fX_l\big(\alpha(X_0, \overset{X_l}{\overset{\curlyvee}{\ldots}}, X_k)\big) \\ = {} & \sum_{i \neq l} \big[(-1)^i X_i(f)\alpha(X_0, \ldots, X_l, \overset{X_i}{\overset{\curlyvee}{\ldots}}, X_k) \\ & + (-1)^i fX_i\big(\alpha(X_0, \ldots, X_l, \overset{X_i}{\overset{\curlyvee}{\ldots}}, X_k)\big)\big] \\ & + (-1)^l fX_l\big(\alpha(X_0, \overset{X_l}{\overset{\curlyvee}{\ldots}}, X_k)\big) \\ = {} & fT_1(X_0, \ldots, X_k) + \sum_{i \neq l} (-1)^i X_i(f)\alpha(X_0, \overset{X_i}{\overset{\curlyvee}{\ldots}}, X_k), \end{aligned}$$

and, using (3.1.1),

$$T_2(X_0, \ldots, fX_l, \ldots, X_k) = \sum_{\substack{i<j \\ i,j \neq l}} (-1)^{i+j} \alpha\big([X_i, X_j], X_0, \ldots, fX_l, \overset{X_i\ X_j}{\overset{\curlyvee\ \curlyvee}{\ldots\ldots}}, X_k\big)$$

$$
\begin{aligned}
&+ \sum_{i<l} (-1)^{i+l} \alpha\big([X_i, f X_l], X_0, \overset{X_i\;X_l}{\ldots\ldots}, X_k\big) \\
&+ \sum_{l<j} (-1)^{l+j} \alpha\big([f X_l, X_j], X_0, \overset{X_l\;X_j}{\ldots\ldots}, X_k\big) \\
= f &\sum_{\substack{i<j \\ i,j\neq l}} (-1)^{i+j} \alpha\big([X_i, X_j], X_0, \ldots, X_l, \overset{X_i\;X_j}{\ldots\ldots}, X_k\big) \\
&+ \sum_{i<l} (-1)^{i+l} f \alpha\big([X_i, X_l], X_0, \overset{X_i\;X_l}{\ldots\ldots}, X_k\big) \\
&+ \sum_{i<l} (-1)^{i+l} X_i(f) \alpha(X_l, X_0, \overset{X_i\;X_l}{\ldots\ldots}, X_k) \\
&+ \sum_{l<j} (-1)^{l+j} f \alpha\big([X_l, X_j], X_0, \overset{X_l\;X_j}{\ldots\ldots}, X_k\big) \\
&- \sum_{l<j} (-1)^{l+j} X_j(f) \alpha(X_l, X_0, \overset{X_l\;X_j}{\ldots\ldots}, X_k).
\end{aligned}
$$

The first, the second and the fourth term on the right hand side give $f T_2(X_0, \ldots, X_k)$. In the remaining two terms we bring the vector fields X_l to their original position. This yields a factor $(-1)^{l-1}$ in the third term and a factor $(-1)^l$ in the fifth term. Thus,

$$
T_2(X_0, \ldots, f X_l, \ldots, X_k) = f T_2(X_0, \ldots, X_k) - \sum_{i \neq l} (-1)^i X_i(f) \alpha(X_0, \overset{X_i}{\ldots}, X_k),
$$

and the right hand side of (4.1.22) is f-linear in every vector field X_j, indeed.

Now, since d, the derivations X_i of $C^\infty(M)$ and the commutator of vector fields are local operations, it is enough to show that the assertion holds for the case $X_i = \partial^\kappa_{j_i}$, where (U, κ) is an arbitrary local chart and $j_0 < \cdots < j_k$. In this case, the right hand side reads

$$
\sum_{i=0}^k (-1)^i \partial^\kappa_{j_i} \Big(\alpha\big(\partial^\kappa_{j_0}, \overset{\partial^\kappa_{j_i}}{\ldots}, \partial^\kappa_{j_k}\big)\Big) = \sum_{i=0}^k (-1)^i \partial^\kappa_{j_i} \alpha_{j_0, \ldots, j_{i-1}, j_{i+1}, \ldots, j_k}
$$

and for the left hand side we obtain, using (4.1.20),

$$
\mathrm{d}\alpha\big(\partial^\kappa_{j_0}, \ldots, \partial^\kappa_{j_k}\big) = \Bigg(\sum_{p,I} \partial^\kappa_p \alpha_I \mathrm{d}\kappa^p \wedge \mathrm{d}\kappa^I\Bigg)\big(\partial^\kappa_{j_0}, \ldots, \partial^\kappa_{j_k}\big)
$$

$$= \sum_{p,I} \partial_p^\kappa \alpha_I \sum_{i=0}^{k} (-1)^i \delta_{j_i}^p \delta_{\{j_0,\dots,j_{i-1},j_{i+1},\dots,j_k\}}^I$$

$$= \sum_{i=0}^{k} (-1)^i \partial_{j_i}^\kappa \alpha_{j_0,\dots,j_{i-1},j_{i+1},\dots,j_k}.$$

This completes the proof. □

The next proposition shows that the exterior derivative commutes with the pull-back operation.

Proposition 4.1.7 *For a smooth mapping* $\varphi\colon M \to N$,

$$\mathrm{d}\big(\varphi^*\alpha\big) = \varphi^*(\mathrm{d}\alpha), \quad \alpha \in \Omega^*(N). \tag{4.1.23}$$

Proof Due to Remarks 4.1.1/1 and 4.1.5, and since $\varphi^*(\alpha_{\restriction U}) = (\varphi^*\alpha)_{\restriction \varphi^{-1}(U)}$ for all open subsets U of N, it suffices to prove the assertion for $\alpha = f_0 \mathrm{d} f_1 \wedge \cdots \wedge \mathrm{d} f_k$, where $f_0, \dots, f_k$ are arbitrary smooth functions on M. Formula (4.1.23) holds for functions, because

$$\big\langle \big(\mathrm{d}\varphi^* f\big)_m, X_m \big\rangle = X_m\big(\varphi^* f\big) = \big(\varphi' X_m\big)(f) = \big\langle (\mathrm{d} f)_{\varphi(m)}, \varphi' X_m \big\rangle = \big\langle \big(\varphi^* \mathrm{d} f\big)_m, X_m \big\rangle$$

for all $m \in M$ and $X_m \in \mathrm{T}_m M$. Using this and (4.1.15), we obtain

$$\mathrm{d}\big(\varphi^*(f_0 \mathrm{d} f_1 \wedge \cdots \wedge \mathrm{d} f_k)\big) = \varphi^*\big(\mathrm{d}(f_0 \mathrm{d} f_1 \wedge \cdots \wedge \mathrm{d} f_k)\big).$$ □

Finally, we derive relations for the Lie derivative of differential forms.

Proposition 4.1.8 *For* $\alpha \in \Omega^k(M)$, $f \in C^\infty(M)$ *and* $X \in \mathfrak{X}(M)$,

$$\mathscr{L}_X \alpha = X \lrcorner \mathrm{d}\alpha + \mathrm{d}(X \lrcorner \alpha), \tag{4.1.24}$$

$$\mathscr{L}_X(\mathrm{d}\alpha) = \mathrm{d}(\mathscr{L}_X \alpha), \tag{4.1.25}$$

$$\mathscr{L}_X(X \lrcorner \alpha) = X \lrcorner (\mathscr{L}_X \alpha), \tag{4.1.26}$$

$$\mathscr{L}_{fX} \alpha = f \mathscr{L}_X \alpha + df \wedge (X \lrcorner \alpha). \tag{4.1.27}$$

Using the notation $X \lrcorner \equiv i_X$, the identity (4.1.24) can be written in the form

$$\mathscr{L}_X = i_X \circ \mathrm{d} + \mathrm{d} \circ i_X.$$

Proof We evaluate both sides of (4.1.24) on a k-vector field $Y_1 \wedge \cdots \wedge Y_k$. For the two terms on the right hand side, Proposition 4.1.6 yields

$$
\begin{aligned}
\langle X \lrcorner \mathrm{d}\alpha, Y_1 \wedge \cdots \wedge Y_k \rangle &= \langle \mathrm{d}\alpha, X \wedge Y_1 \wedge \cdots \wedge Y_k \rangle \\
&= X\big(\langle \alpha, Y_1 \wedge \cdots \wedge Y_k \rangle\big) \\
&\quad + \sum_{i=1}^{k} (-1)^i Y_i \big(\langle \alpha, X \wedge \overset{Y_i}{\overset{\curlyvee}{\dots}} \wedge Y_k \rangle\big) \\
&\quad + \sum_{l=1}^{k} (-1)^l \big\langle \alpha, [X, Y_l] \wedge Y_1 \wedge \overset{Y_l}{\overset{\curlyvee}{\dots}} \wedge Y_k \big\rangle \\
&\quad + \sum_{i<l} (-1)^{i+l} \big\langle \alpha, [Y_i, Y_l] \wedge X \wedge Y_1 \wedge \overset{Y_i Y_l}{\overset{\curlyvee \, \curlyvee}{\dots}} \wedge Y_k \big\rangle, \\
\big\langle \mathrm{d}(X \lrcorner \alpha), Y_1 \wedge \cdots \wedge Y_k \big\rangle &= \sum_{i=1}^{k} (-1)^{i+1} Y_i \big(\langle X \lrcorner \alpha, Y_1 \wedge \overset{Y_i}{\overset{\curlyvee}{\dots}} \wedge Y_k \rangle\big) \\
&\quad + \sum_{i<l} (-1)^{i+l} \big\langle X \lrcorner \alpha, [Y_i, Y_l] \wedge Y_1 \wedge \overset{Y_i Y_l}{\overset{\curlyvee \, \curlyvee}{\dots}} \wedge Y_k \big\rangle \\
&= \sum_{i=1}^{k} (-1)^{i+1} Y_i \big(\langle \alpha, X \wedge Y_1 \wedge \overset{Y_i}{\overset{\curlyvee}{\dots}} \wedge Y_k \rangle\big) \\
&\quad + \sum_{i<l} (-1)^{i+l} \big\langle \alpha, X \wedge [Y_i, Y_l] \wedge Y_1 \wedge \overset{Y_i Y_l}{\overset{\curlyvee \, \curlyvee}{\dots}} \wedge Y_k \big\rangle.
\end{aligned}
$$

According to Propositions 3.3.2 and 3.3.3, the sum of these terms gives

$$
X\big(\langle \alpha, Y_1 \wedge \cdots \wedge Y_k \rangle\big) - \sum_{l=1}^{k} \big\langle \alpha, Y_1 \wedge \cdots \wedge [X, Y_l] \wedge \cdots \wedge Y_k \big\rangle = \langle \mathscr{L}_X \alpha, Y_1 \wedge \cdots \wedge Y_k \rangle .
$$

The proofs of (4.1.25)–(4.1.27) are left to the reader (Exercise 4.1.6). □

For later use, we note the following consequence of (4.1.25). Let X be a time-dependent vector field on M with flow $\Phi : \mathscr{D} \to M$. Then, for all $\alpha \in \Omega^*(M)$ and all $(t, t_0, m) \in \mathscr{D}$,

$$
\frac{\mathrm{d}}{\mathrm{d}t} \big(\Phi^*_{t,t_0} \alpha\big)_m = \big(\Phi^*_{t,t_0} (\mathscr{L}_{X_t} \alpha)\big)_m , \tag{4.1.28}
$$

where $\mathscr{L}_{X_t}$ is the Lie derivative with respect to the ordinary vector field X_t with fixed t. Indeed, in the case $\alpha = f \in C^\infty(M)$, we find

$$
\frac{\mathrm{d}}{\mathrm{d}t} \big(\Phi^*_{t,t_0} f\big)_m = \frac{\mathrm{d}}{\mathrm{d}t} f\big(\Phi_{t,t_0}(m)\big) = \big(X_t(f)\big)\big(\Phi_{t,t_0}(m)\big) = \big(\Phi^*_{t,t_0} \circ \mathscr{L}_{X_t} f\big)(m).
$$

Since both Φ^*_{t,t_0} and $\mathscr{L}_{X_t}$ commute with the exterior derivative on M and since both sides of (4.1.28) are derivations with respect to the wedge product, the assertion follows. Note that (4.1.28) also holds for time-independent vector fields. In this form, it is a special case of (3.3.3).

We conclude this section by two remarks.

Remark 4.1.9 (Structure of $\Omega^*(M)$)

1. By the defining properties 1 and 2 of the exterior derivative, $(\Omega^*(M), +, \wedge, \mathrm{d})$ is an associative graded commutative differential algebra, called the Cartan algebra.
2. Since $\mathrm{d}^2 = 0$, the sequence of linear mappings

$$0 \to \Omega^0(M) \xrightarrow{\mathrm{d}} \Omega^1(M) \xrightarrow{\mathrm{d}} \cdots \xrightarrow{\mathrm{d}} \Omega^n(M) \xrightarrow{\mathrm{d}} 0 \tag{4.1.29}$$

   is a chain complex, called the de Rham complex of M. It contains important information about the topological structure of M, see Sect. 4.3.
3. The following generalization of the de Rham complex is an important element of noncommutative geometry. Let $\mathfrak{A}$ be a unital associative $*$-algebra over $\mathbb{C}$. An involutive graded differential algebra over $\mathfrak{A}$ is a tuple

$$\left\{\Lambda^*_{\mathfrak{A}}, \cdot, *, \mathrm{d}\right\},$$

   where $\Lambda^*_{\mathfrak{A}}$ is an $\mathfrak{A}$-bimodule of the form $\Lambda^*_{\mathfrak{A}} = \bigoplus_{k=0}^{\infty} \Lambda^k_{\mathfrak{A}}$, with $\Lambda^0_{\mathfrak{A}} \equiv \mathfrak{A}$, $\cdot$ is an $\mathfrak{A}$-bilinear multiplication in $\Lambda^*_{\mathfrak{A}}$ fulfilling $\Lambda^k_{\mathfrak{A}} \cdot \Lambda^l_{\mathfrak{A}} \subset \Lambda^{k+l}_{\mathfrak{A}}$ and $\mathrm{d} \colon \Lambda^k_{\mathfrak{A}} \to \Lambda^{k+1}_{\mathfrak{A}}$ is a derivation fulfilling $\mathrm{d}(\lambda^*) = (-1)^k(\mathrm{d}\lambda)^*$ for any $\lambda \in \Lambda^k_{\mathfrak{A}}$. Moreover, for given $e \in \mathrm{End}_{\mathfrak{A}}(\mathfrak{A}^p)$ with $e^2 = e$, $\mathscr{E} := e\mathfrak{A}^p$ is a right projective module over $\mathfrak{A}$. In view of the Serre-Swan Theorem [156], finitely generated projective modules naturally generalize vector bundles.

The second remark introduces smooth families of differential forms and differential forms with values in a vector space. Both concepts will occasionally be used later on.

Remark 4.1.10

1. A one-parameter family of r-forms $\{\alpha_t : t \in [a, b]\}$ on M is said to be smooth if the mapping

$$[a, b] \times M \to \bigwedge^r \mathrm{T}^* M, \quad (t, m) \mapsto \alpha_t(m),$$

   is smooth. Given such a family, by pointwise integration $\int_a^b \alpha_t(m)\,\mathrm{d}t$ in $\mathrm{T}^*_m M$, $m \in M$, we obtain an r-form on M, denoted by $\int \alpha_t \mathrm{d}t$. Similarly, by pointwise differentiation $\frac{\mathrm{d}}{\mathrm{d}s}\restriction_t \alpha_s(m)$ in $\mathrm{T}^*_m M$, we obtain a smooth family of r-forms on M, denoted by $\frac{\mathrm{d}}{\mathrm{d}t}\alpha_t$. In this sense, one may omit the point m in the Formulae (3.3.3) and (4.1.28). Application of the operations of exterior derivative, pull-back, inner product or Lie derivative to a smooth family of differential forms yields a smooth family again. Using Proposition 4.1.6, one can show that these operations commute with the operations of integration or differentiation with respect to the parameter. That is,

$$\mathrm{d}\int \alpha_t \mathrm{d}t = \int \mathrm{d}\alpha_t \mathrm{d}t, \qquad \mathrm{d}\frac{\mathrm{d}}{\mathrm{d}t}\alpha_t = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{d}\alpha_t \tag{4.1.30}$$

and similar formulae with d replaced by φ^* for any smooth mapping $\varphi : N \to M$ or by $X \lrcorner$ or $\mathscr{L}_X$ for any vector field X on M (Exercise 4.1.7).

2. Let V be a finite-dimensional $\mathbb{K}$-vector space. A differential r-form on M with values in V is a section in the vector bundle whose fibres are the r-forms on $\mathrm{T}_m M$ with values in V, that is, the antisymmetric r-linear mappings

$$\mathrm{T}_m M \times \cdots \times \mathrm{T}_m M \to V.$$

We denote the vector space of these forms by $\Omega^r(M, V)$ and we define $\Omega^*(M, V) := \bigoplus_{r=0}^n \Omega^r(M, V)$. The bundle of r-forms on $\mathrm{T}_m M$ with values in V can be identified with the tensor product $\bigwedge^r \mathrm{T}^*M \otimes (M \times V)$. Hence, $\Omega^r(M, V)$ can be identified with the space of sections of this bundle, which in turn can be identified with $\Omega^r(M) \otimes V$. Accordingly, if $\{e_a\}$ is a basis in V, every differential r-form α with values in V can be written in the form

$$\alpha = \alpha^a \otimes e_a$$

(summation convention), where α^a are uniquely determined ordinary differential r-forms and the e_a can be interpreted either as elements of V, in which case α is viewed as an element of $\Omega^r(M) \otimes V$, or as global sections in $M \times V$, in which case α is viewed as a section of $\bigwedge^r \mathrm{T}^*M \otimes (M \times V)$. As a consequence, with respect to a local chart (U, κ) on M, α has the local representation

$$\alpha_{\restriction U} = \alpha^a_{i_1,\ldots,i_k} \mathrm{d}\kappa^{i_1} \wedge \cdots \wedge \mathrm{d}\kappa^{i_k} \otimes e_a \tag{4.1.31}$$

with smooth functions $\alpha^a_{i_1,\ldots,i_k} : U \to \mathbb{R}$. The statements of this section about ordinary differential forms carry over in an obvious way to differential forms with values in V, except for those which have to do with the exterior product. An exterior product exists on $\Omega^*(M, V)$ only if V carries in addition the structure of an algebra. If so, the exterior product is defined by (2.4.17).

Exercises

4.1.1 Prove Formulae (4.1.4) and (4.1.5).

4.1.2 Prove the transformation laws given in (4.1.12).

4.1.3 Prove the local Formula (4.1.17) for the pull-back of differential forms.

4.1.4 Prove that the operation of inner multiplication is f-linear.

4.1.5 Complete the proof of Theorem 4.1.4 by verifying that the system of linear mappings d defined by (4.1.21) is an exterior derivative.

4.1.6 Prove Eqs. (4.1.25)–(4.1.27).

4.1.7 Use Proposition 4.1.6 to prove Formula (4.1.30) as well as analogous formulae for the operations φ^*, $X \lrcorner$ and $\mathscr{L}_X$.

4.2 Integration and Integral Invariants

The idea of integration of a differential n-form α on an n-dimensional manifold M is as follows. We choose an atlas and take the pull-backs of α, weighted with the elements of a subordinate partition of unity, by the local chart mappings. This yields a family of differential n-forms on open subsets of $\mathbb{R}^n$, which we integrate and sum up.

Definition 4.2.1 (Orientation) Let M be a manifold. An atlas $\mathscr{A}_M$ of M is called oriented, or an orientation of M, if for every pair of charts (U_1, κ_1) and (U_2, κ_2) belonging to $\mathscr{A}_M$ and such that $U_1 \cap U_2 \neq \varnothing$ one has

$$\det\big(\kappa_1 \circ \kappa_2^{-1}\big)' > 0. \tag{4.2.1}$$

A manifold is called orientable if it admits an orientation. It is called oriented if an orientation has been chosen.

Let M be oriented by the atlas $\mathscr{A}$. A local chart on M is said to be oriented if $\mathscr{A}$ remains oriented when this chart is added to it. A local frame in $\mathrm{T}M$ is said to be oriented if the matrix transforming this frame at some point to the frame $\{\partial_i\}$ of some chart in $\mathscr{A}$ has positive determinant.

Remark 4.2.2 The set of ordered bases of a vector space V decomposes into the following two equivalence classes: $(e_1, \dots, e_n)$ and $(e'_1, \dots, e'_n)$ are equivalent iff the transition matrix A, defined by $e'_i = A_i{}^j e_j$, has positive determinant. Each of these two equivalence classes defines an orientation of V. This definition immediately carries over to manifolds: choosing an orientation means choosing orientations in each tangent space in a smooth way. Indeed, the local frame $\{\partial_1^\kappa, \dots, \partial_n^\kappa\}$ induced by the local chart (U, κ) provides an orientation of the open subset U and (4.2.1) yields the consistency condition for the orientations of all charts belonging to the chosen atlas.

Definition 4.2.3 (Volume form) A nowhere-vanishing n-form on an n-dimensional manifold is called a volume form.

Proposition 4.2.4 *A manifold is orientable iff it admits a volume form.*

Proof First, assume that M is orientable. Let $\{(U_i, \kappa_i)\}$ be a countable, locally finite, oriented atlas and let $\{f_i\}$ be a subordinate partition of unity. We define the following family of n-forms on M:

$$(\mathrm{v}_i)_m := \begin{cases} 0 & m \notin U_i \\ f_i(m)\mathrm{d}\kappa_i^1 \wedge \cdots \wedge \mathrm{d}\kappa_i^n & m \in U_i . \end{cases}$$

Since $\{\mathrm{supp}(f_i)\}$ is locally finite, $\mathsf{v} := \sum_i \mathsf{v}_i$ is a well-defined smooth n-form on M. It remains to show that $\mathsf{v}_m \neq 0$ for all $m \in M$. Choose i_0 such that $m \in U_{i_0}$. By the transformation law (4.1.12) for local frames, in the chart (U_{i_0}, κ_{i_0}) we have

$$\mathsf{v}_m = \Big(f_{i_0}(m) + \sum f_i(m) \det\big(\kappa_i \circ \kappa_{i_0}^{-1}\big)'\big(\kappa_{i_0}(m)\big)\Big) \mathrm{d}\kappa_{i_0}^1 \wedge \cdots \wedge \mathrm{d}\kappa_{i_0}^n,$$

where the sum runs over all i such that $m \in U_i$. Since $f_i \geq 0$ and $f_{i_0}(m) + \sum f_i(m) = 1$, and since the determinants are positive, we conclude $\mathsf{v}_m \neq 0$. Conversely, assume that there exists a volume form v. In a local chart (U, κ) the n-form v is given by

$$\mathsf{v}_{\restriction U} = h \mathrm{d}\kappa^1 \wedge \cdots \wedge \mathrm{d}\kappa^n.$$

Since $\mathsf{v}_m \neq 0$ for all $m \in U$, we have either $h > 0$ or $h < 0$. If $h > 0$, we leave κ as it is. If $h < 0$, we define a new chart by $\overline{\kappa} = (-\kappa^1, \kappa^2, \ldots, \kappa^n)$. By this procedure, we obtain an atlas with $h > 0$ in every chart. Then, for any two charts (U, κ) and $(\tilde{U}, \tilde{\kappa})$ of this atlas, we have $\tilde{h} = \det(\tilde{\kappa} \circ \kappa^{-1})' h$ and thus $\det(\tilde{\kappa} \circ \kappa^{-1})' > 0$. □

Example 4.2.5

1. The vector space $\mathbb{R}^n$ is orientable. One possible orientation, called the standard orientation, is given by the identity chart mapping. A volume form corresponding to this orientation is $\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$ with $x^1, \ldots, x^n$ being the standard coordinates.
2. The Möbius strip is not orientable (Exercise 4.2.1).
3. The odd-dimensional real projective spaces are orientable, the even-dimensional real projective spaces are not orientable (Exercise 4.2.1).
4. The level set of a regular value of a differentiable mapping $f : \mathbb{R}^n \to \mathbb{R}^m$ is orientable (Exercise 4.2.1).
5. Every parallelizable manifold is orientable, because if $\{X_1, \ldots, X_n\}$ is a global frame in $\mathrm{T}M$ and $\{\xi_1, \ldots, \xi_n\}$ is the dual frame in T^*M, then $\xi_1 \wedge \cdots \wedge \xi_n$ is a volume form. In particular, Lie groups are orientable, see Chap. 5.

Let $\{x^i\}$ denote the standard coordinates on $\mathbb{R}^n$ and let there be chosen the volume form $\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$. Then, for every n-form $\alpha \in \Omega^n(\mathbb{R}^n)$ there exists a unique function $f : \mathbb{R}^n \to \mathbb{R}$ such that $\alpha = f \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$. We define

$$\int_{\mathbb{R}^n} \alpha := \int_{\mathbb{R}^n} f \mathrm{d}^n x, \tag{4.2.2}$$

provided the integral on the right hand side exists. Here, $\mathrm{d}^n x$ denotes the Lebesgue measure[3] on $\mathbb{R}^n$.

[3]For our purposes, the Riemann integral would do as well.

Definition 4.2.6 (Integral) Let M be an oriented manifold of dimension n. Let $\{(U_i, \kappa_i)\}$ be a countable, locally finite and oriented atlas and let $\{f_i\}$ be a subordinate partition of unity. A differential n-form $\alpha \in \Omega^n(M)$ is called integrable if the integrals $\int_{\kappa_i(U_i)} (\kappa_i^{-1})^*(f_i \alpha)$ exist and the family of their absolute values is summable. If α is integrable, we define

$$\int_M \alpha := \sum_i \int_{\kappa_i(U_i)} \big(\kappa_i^{-1}\big)^* (f_i \alpha). \tag{4.2.3}$$

One easily shows that the right hand side of Eq. (4.2.3) neither depends on the choice of the atlas nor on the choice of the partition of unity (Exercise 4.2.2). Obviously, the integral (4.2.3) always exists if M is compact or if α has compact support. Sometimes, one wishes to integrate a differential k-form over a k-dimensional submanifold:

Definition 4.2.7 (Integral over a submanifold) Let (N, φ) be a k-dimensional oriented submanifold of M and let $\beta \in \Omega^k(M)$. The integral of β over N is defined by

$$\int_N \beta := \int_N \varphi^* \beta. \tag{4.2.4}$$

Next, we extend the famous Stokes Theorem of classical calculus to the case of manifolds. For that purpose, we need the notion of manifold with boundary. Consider the closed half-space $\mathbb{R}^n_- := \{\mathbf{x} \in \mathbb{R}^n : x_1 \leq 0\}$, equipped with the relative topology induced from $\mathbb{R}^n$. The boundary of $\mathbb{R}^n_-$ is defined to be the closed subset $\partial \mathbb{R}^n_- = \{0\} \times \mathbb{R}^{n-1}$. The boundary of an open subset $U \subset \mathbb{R}^n_-$ is defined to be $\partial U := U \cap \partial \mathbb{R}^n_-$. This is an open subset of $\partial \mathbb{R}^n_- \cong \mathbb{R}^{n-1}$. Note that, in general, ∂U does not coincide with the topological boundary of U as a subset of $\mathbb{R}^n_-$.

Definition 4.2.8 Let $U \subset \mathbb{R}^n_-$ be open and let $\mathbf{x} \in U$. A mapping $\Phi : U \to \mathbb{R}^m$ is smooth at $\mathbf{x}$ if there exists an open neighbourhood $\tilde{U}$ of $\mathbf{x}$ in $\mathbb{R}^n$ and a smooth mapping $\tilde{\Phi} : \tilde{U} \to \mathbb{R}^m$ such that $\Phi_{\restriction U \cap \tilde{U}} = \tilde{\Phi}_{\restriction U \cap \tilde{U}}$. A diffeomorphism between two open subsets of $\mathbb{R}^n_-$ is a smooth bijective mapping whose inverse is also smooth.

Remark 4.2.9

1. If $\Phi : U \to V$ is a diffeomorphism of open subsets of $\mathbb{R}^n_-$, then $\Phi(\partial U) = \partial V$ and the restriction $\Phi_{\restriction \partial U} : \partial U \to \partial V$ is a diffeomorphism of open subsets of $\mathbb{R}^{n-1}$. For $\mathbf{x} \in \partial U$, the first derivative $\Phi'(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^n$ is of the form

$$\Phi'(\mathbf{x}) = \left(\begin{array}{c|ccc} \frac{\partial \Phi_1}{\partial x_1} & 0 & \cdots & 0 \\ \hline \frac{\partial \Phi_2}{\partial x_1} & \frac{\partial \Phi_2}{\partial x_2} & \cdots & \frac{\partial \Phi_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial \Phi_n}{\partial x_1} & \frac{\partial \Phi_n}{\partial x_2} & \cdots & \frac{\partial \Phi_n}{\partial x_n} \end{array} \right), \qquad \frac{\partial \Phi_1}{\partial x_1} > 0. \tag{4.2.5}$$

Hence, $\Phi'(\mathbf{x})$ maps the subspace $\{0\} \times \mathbb{R}^{n-1}$ and the half-spaces $\mathbb{R}^n_\pm$ to themselves, respectively.

2. While a differentiable mapping $\Phi : U \to \mathbb{R}^m$ has many distinct extensions $\tilde{\Phi} : \tilde{U} \to \mathbb{R}^m$ at a boundary point $\mathbf{x} \in U$, its derivative $\Phi'(\mathbf{x})$ at $\mathbf{x}$ is unique: since the derivative exists, it suffices to check uniqueness of the directional derivatives. The latter follows from the fact that for arbitrary $\mathbf{h} \in \mathbb{R}^n$ either $\mathbf{h}$ or $-\mathbf{h}$ (or both) are in $\mathbb{R}^n_-$.

Now, we carry over the notions of Sect. 1.1 to the present situation by replacing the linear space $\mathbb{R}^n$ in Definitions 1.1.1–1.1.4 by the half-space $\mathbb{R}^n_-$. This way, we obtain the notions of

– topological manifold with boundary, including its dimension,
– local chart, compatibility of local charts and atlas for a topological manifold with boundary,
– manifold with boundary and differentiable mapping,
– orientation for a manifold with boundary.

In addition, by replacing the term manifold by the term manifold with boundary in the respective definition, we obtain the notion of partition of unity for a manifold with boundary and the notion of vector bundle with boundary, including the notions of morphism and section. Next, we construct the tangent space at a point m of a manifold M with boundary in the same way as for ordinary manifolds, cf. Sect. 1.4, with the following changes.

– Instead of curves through m and ordinary derivatives in (1.4.2), we take curves $\gamma : [0, 1] \to M$ starting at m and curves $\gamma : [-1, 0] \to M$ ending at m and use one-sided derivatives.
– For $\mathbf{x} \in \mathbb{R}^n$, we define the curve $\gamma^{\mathbf{x}}$ in (1.4.3) as a curve starting at $\mathbf{x}$ (that is, $t \in [0, 1]$) if $\mathbf{x} \in \mathbb{R}^n_-$ and as a curve ending at $\mathbf{x}$ (that is, $t \in [-1, 0]$) otherwise.

This way, the tangent spaces become linear spaces. The construction of the tangent bundle and the tangent mapping carries over without change, rendering vector bundles with boundary and morphisms thereof. The same is true for the various tensor bundles. Based on this, the calculus on manifolds carries over to manifolds with boundary in an obvious way, including the notions of vector field and flow, tensor field, transport and Lie derivative, differential form, pull-back, exterior derivative and integral, as well as immersion and submersion, submanifold and vector subbundle.

Next, we define the interior and the boundary of a manifold with boundary. As a consequence of Remark 4.2.9/1, the transition mappings of local charts of a topological manifold with boundary map points in $\partial\mathbb{R}^n_-$ to points in $\partial\mathbb{R}^n_-$. That is, if for $m \in M$ there holds $\kappa(m) \in \partial\mathbb{R}^n_-$ for some local chart κ at m, then this holds for any local chart at m.

Definition 4.2.10 (Boundary and interior) Let M be a manifold with boundary. A point $m \in M$ is called a boundary point if there exists a local chart (U, κ) at m

with $\kappa(m) \in \partial\mathbb{R}^n_-$. Otherwise, m is called an interior or inner point. The subset of boundary points is denoted by ∂M and is called the boundary of M. The subset of interior points is denoted by $\mathrm{Int}(M)$ and is called the interior of M.

Example 4.2.11 The closed unit disk D^n of $\mathbb{R}^n$ is a smooth manifold with boundary. The interior is the open unit ball and the boundary coincides with the smooth manifold S^{n-1} (Exercise 4.2.4).[4]

Example 4.2.12 An ordinary manifold is a manifold with boundary. To see this, replace the coordinate x_1 by $-\mathrm{e}^{x_1}$ in each local chart. For an ordinary manifold, $\partial M = \varnothing$ and $\mathrm{Int}(M) = M$. Hence any statement for manifolds with boundary holds, in particular, for ordinary manifolds.

Proposition 4.2.13 *Let M be a manifold with boundary and let $n = \dim M$.*

1. $\mathrm{Int}(M)$ *and* ∂M *are embedded submanifolds of* M *of dimension* n *and* $n-1$*, respectively.*
2. $\mathrm{Int}(M)$ *and* ∂M *are ordinary manifolds, that is,* $\partial\,\mathrm{Int}(M) = \varnothing$ *and* $\partial(\partial M) = \varnothing$.
3. *Every orientation on* M *induces orientations on* $\mathrm{Int}(M)$ *and on* ∂M.

Proof For the first assertion, we construct atlases on $\mathrm{Int}(M)$ and ∂M by restriction of appropriate local charts of an atlas for M (Exercise 4.2.6). Then, the second assertion is obvious. For $\mathrm{Int}(M)$, the third assertion is obvious as well. For ∂M, this assertion follows from the fact that the derivative of the transition mapping of two induced charts is given by the lower right block of the restriction of the original transition mapping to boundary points, given in (4.2.5). □

Now we can extend Stokes' Theorem of classical integral calculus to manifolds.

Theorem 4.2.14 (Stokes) *Let M be an n-dimensional compact and oriented manifold with boundary and let ∂M be endowed with the induced orientation. Then, for every $(n-1)$-form β on M,*

$$\int_M \mathrm{d}\beta = \int_{\partial M} \beta. \tag{4.2.6}$$

We note that all the integral theorems of classical vector analysis follow from Stokes' Theorem, see Sect. 4.5.

Proof We choose a finite atlas $\{(U_i, \kappa_i)\}$ and a subordinate partition of unity $\{f_i\}$. By finiteness, it is enough to show that $\int_M \mathrm{d}(f_i\beta) = \int_{\partial M} f_i\beta$ for all i. Thus, we may assume that β has support in one chart domain U_k and that, therefore, the partition

[4] Beware that the interior and the boundary of a manifold with boundary are defined by the manifold alone, whereas the interior and the boundary of a subset of a topological space are defined with respect to the ambient space. By chance, in this example, these two notions coincide.

of unity $\{f_i\}$ can be chosen so that $f_{k \restriction \mathrm{supp}(\beta)} = 1$ and $f_{i \restriction \mathrm{supp}(\beta)} = 0$ for all $i \neq k$. Under these assumptions, the calculation of $\int_M \mathrm{d}\beta$ reduces to

$$\int_M \mathrm{d}\beta = \int_{\kappa_k(U_k)} \big(\kappa_k^{-1}\big)^* \mathrm{d}\beta = \int_{\kappa_k(U_k)} \mathrm{d}\big(\big(\kappa_k^{-1}\big)^* \beta\big).$$

We expand

$$\big(\kappa_k^{-1}\big)^* \beta = \sum_{l=1}^{n} b_l \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^{l-1} \wedge \mathrm{d}x^{l+1} \wedge \cdots \wedge \mathrm{d}x^n \tag{4.2.7}$$

with smooth[5] functions b_l on $\kappa_k(U_k)$. Then,

$$\mathrm{d}\big(\big(\kappa_k^{-1}\big)^* \beta\big) = \sum_{l=1}^{n} (-1)^{l-1} \frac{\partial b_l}{\partial x^l} \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$$

and hence by (4.2.2)

$$\int_M \mathrm{d}\beta = \sum_{l=1}^{n} (-1)^{l-1} \int_{\kappa_k(U_k)} \frac{\partial b_l}{\partial x^l} \mathrm{d}^n x. \tag{4.2.8}$$

Since $\mathrm{supp}(b_l) \subset \kappa_k(\mathrm{supp}(\beta))$, the functions b_l have compact support in $\kappa_k(U_k)$. First, this implies that we can extend them by zero to $\mathbb{R}^n_-$. Second, this implies that there exists $R > 0$ such that $\mathrm{supp}(b_l)$ is contained in the interior of the half-cube $[-R, 0] \times [-R, R]^{n-1}$, so that we can replace the range of integration on the right hand side of (4.2.8) by this half-cube. Integration over x^l in the l-th term yields

$$\begin{aligned}
\int_M \mathrm{d}\beta = & \int_{[-R,R]^{n-1}} \big(b_1\big(0, x^2, \ldots, x^n\big) - b_1\big(-R, x^2, \ldots, x^n\big)\big) \mathrm{d}x^2 \ldots \mathrm{d}x^n \\
& + \sum_{l>1} (-1)^{l-1} \int_{[-R,0]\times[-R,R]^{n-2}} \big(b_l\big(x^1, \ldots, x^{l-1}, R, x^{l+1}, \ldots, x^n\big) \\
& - b_l\big(x^1, \ldots, x^{l-1}, -R, x^{l+1}, \ldots, x^n\big)\big) \mathrm{d}x^1 \ldots \mathrm{d}x^{l-1} \mathrm{d}x^{l+1} \ldots \mathrm{d}x^n \\
= & \int_{[-R,R]^{n-1}} b_1\big(0, x^2, \ldots, x^n\big) \mathrm{d}x^2 \ldots \mathrm{d}x^n.
\end{aligned}$$

By extending the range of integration to $\mathbb{R}^{n-1}$ we finally arrive at

$$\int_M \mathrm{d}\beta = \int_{\mathbb{R}^{n-1}} b_1\big(0, x^2, \ldots, x^n\big) \mathrm{d}x^2 \ldots \mathrm{d}x^n. \tag{4.2.9}$$

[5] In the sense of Definition 4.2.8.

Next, we calculate $\int_{\partial M} \beta$. For that purpose, we need the natural inclusion mapping $\iota : \partial M \to M$ and the mappings

$$p : \mathbb{R}^n \to \mathbb{R}^{n-1}, \qquad p(x_1, \dots, x_n) := (x_2, \dots, x_n),$$
$$j : \mathbb{R}^{n-1} \to \mathbb{R}^n, \qquad j(x_2, \dots, x_n) := (0, x_2, \dots, x_n).$$

To evaluate the integral we use the induced atlas $\{(\tilde{U}_i, \tilde{\kappa}_i)\}$ on ∂M, defined by $\tilde{U}_i := U_i \cap \partial M$ and $\tilde{\kappa}_i := p \circ \kappa_i \circ \iota$, and the induced subordinate partition of unity $\{\tilde{f}_i\}$ of ∂M, defined by $\tilde{f}_i := f_i \circ \iota$. Since $j \circ \tilde{\kappa}_i = \kappa_i \circ \iota$, Formula (4.2.4) yields

$$\int_{\partial M} \beta \equiv \int_{\partial M} \iota^* \beta = \int_{\tilde{\kappa}_k(\tilde{U}_k)} \left(\tilde{\kappa}_k^{-1}\right)^* \iota^* \beta = \int_{\tilde{\kappa}_k(\tilde{U}_k)} j^* \left(\kappa_k^{-1}\right)^* \beta.$$

Plugging in the expansion (4.2.7) for $(\kappa_k^{-1})^* \beta$ and using that, due to $j^* \mathrm{d}x^1 = 0$, only the contribution of $l = 1$ survives, we obtain

$$\int_{\partial M} \beta = \int_{\tilde{\kappa}_k(\tilde{U}_k)} b_1\left(0, x^2, \dots, x^n\right) \mathrm{d}x^2 \dots \mathrm{d}x^n.$$

As before, b_1 can be extended by zero to $\mathbb{R}^{n-1}$ and the range of integration can be replaced by $\mathbb{R}^{n-1}$. Then, comparison with (4.2.9) yields the assertion. □

In the remaining part of this section we discuss integral invariants. They play a role in symplectic geometry and in the theory of Hamiltonian systems, see Sect. 9.3. For an exhaustive presentation we refer to [181], see also [60].

Definition 4.2.15 Let $X \in \mathfrak{X}(M)$. A differential k-form α on M is called

1. invariant with respect to X if $\mathscr{L}_X \alpha = 0$,
2. relatively invariant with respect to X if $X \lrcorner \mathrm{d}\alpha = 0$,
3. absolutely invariant with respect to X if $X \lrcorner \alpha = 0$ and $X \lrcorner \mathrm{d}\alpha = 0$.

Equivalently, α is invariant with respect to X iff the flow Φ of X fulfils $\Phi_t^* \alpha = \alpha$ for all t. Due to

$$\mathscr{L}_{fX} \alpha = f \mathscr{L}_X \alpha + \mathrm{d}f \wedge (X \lrcorner \alpha),$$

α is absolutely invariant with respect to X iff it is invariant with respect to fX for all $f \in C^\infty(M)$. If α is relatively invariant with respect to X, then $\mathscr{L}_X \alpha$ is closed.

Proposition 4.2.16 (Poincaré-Cartan) *Let X be a complete vector field on M with flow Φ and let $\alpha \in \Omega^k(M)$. Then,*

1. *α is invariant with respect to X iff*

$$\int_N (\Phi_t \circ \varphi)^* \alpha = \int_N \varphi^* \alpha \tag{4.2.10}$$

for all oriented compact k-dimensional manifolds N with boundary, for all smooth mappings $\varphi\colon N \to M$ and for all $t \in \mathbb{R}$,

2. *α is relatively invariant with respect to X iff*

$$\int_{\partial N} (\Phi_t \circ \varphi \circ \iota)^* \alpha = \int_{\partial N} (\varphi \circ \iota)^* \alpha \tag{4.2.11}$$

for all oriented compact $(k+1)$-dimensional manifolds N with boundary ∂N, for all smooth mappings $\varphi\colon N \to M$ and for all $t \in \mathbb{R}$. Here, $\iota\colon \partial N \to N$ denotes the natural inclusion mapping,

3. *α is absolutely invariant with respect to X iff (4.2.10) holds for the flows of fX for all $f \in C^\infty(M)$.*

Accordingly, invariant forms are often called integral invariants.

Proof 1. If α is invariant, then $\Phi_t^*\alpha = \alpha$ and the statement follows trivially. Conversely, assume that (4.2.10) holds. Then, by assumption, for every embedding $\iota : \mathrm{D}^k \to N$ of the closed unit ball $\mathrm{D}^k \subset \mathbb{R}^k$ we have

$$\int_{\mathrm{D}^k} (\Phi_t \circ \varphi \circ \iota)^* \alpha = \int_{\mathrm{D}^k} (\varphi \circ \iota)^* \alpha,$$

because D^k is compact. But the Lebesgue measurable sets $A \subset N$ are generated by balls, that is, for every such set we obtain $\int_A (\Phi_t \circ \varphi)^*\alpha = \int_A \varphi^*\alpha$ and, therefore, $(\Phi_t \circ \varphi)^*\alpha = \varphi^*\alpha$. If now N runs through all k-dimensional manifolds, we can conclude $\Phi_t^*\alpha = \alpha$.

2. Let α be relatively invariant. Then, $\mathscr{L}_X \mathrm{d}\alpha = \mathrm{d}\mathscr{L}_X\alpha = 0$, that is, $\mathrm{d}\alpha$ is invariant. Using Stokes' Theorem and point 1, we obtain

$$\int_{\partial N} (\Phi_t \circ \varphi \circ \iota)^* \alpha = \int_N (\Phi_t \circ \varphi)^* \mathrm{d}\alpha = \int_N \varphi^* \mathrm{d}\alpha = \int_{\partial N} \iota^* \circ \varphi^* \alpha.$$

The proof of the converse direction is analogous to point 1.

3. This follows from the remark after Definition 4.2.15. □

Remark 4.2.17 Let $\alpha, \beta \in \Omega^*(M)$ be invariant under $X \in \mathfrak{X}(M)$. Since the operations i_X and d commute with $\mathscr{L}_X$ and since $\mathscr{L}_X$ is a derivation of the tensor algebra, $i_X\alpha$, $\mathrm{d}\alpha$ and $\alpha \wedge \beta$ are invariant with respect to X, too. Thus, the set of X-invariant differential forms on M forms a subalgebra of $\Omega^*(M)$ which is closed under i_X and d.

Besides the forms which are invariant with respect to a given vector field, it is also interesting to study the vector fields leaving invariant a given form. The kernel of an antisymmetric multilinear form η on a vector space V is defined as

$$\ker \eta := \{v \in V : v \lrcorner \eta = 0\}. \tag{4.2.12}$$

Definition 4.2.18 Let $\alpha \in \Omega^k(M)$.

1. The subspace $F^\alpha_m := \ker \alpha_m \cap \ker(\mathrm{d}\alpha)_m$ of $\mathrm{T}_m M$ is called the characteristic subspace of α at m.
2. A local vector field X on M with domain U is called characteristic for α if $X_m \in F^\alpha_m$ for all $m \in U$.
3. The distribution spanned by the characteristic local vector fields for α is called the characteristic distribution of α and is denoted by D^α.

While the dimension of F^α_m is locally non-increasing, it may suddenly decrease. Accordingly, the subset $F^\alpha := \bigcup_{m\in M} F^\alpha_m$ of $\mathrm{T}M$ need not establish a distribution in the sense of Definition 3.5.1.[6] Rather, the characteristic distribution D^α is the maximal distribution on M contained in F^α. If $\dim F^\alpha_m$ is constant on M, then D^α and F^α coincide, and both are regular distributions. We also note that, by the above definition, a vector field X is characteristic for α iff

$$X \lrcorner \alpha = 0, \qquad X \lrcorner \mathrm{d}\alpha = 0, \tag{4.2.13}$$

that is, iff α is absolutely invariant with respect to X.

Example 4.2.19 Let $M = \mathbb{R}^2$ with standard coordinates x, y. For $\alpha = \mathrm{d}x$, all characteristic subspaces $F^\alpha_{(x,y)}$ coincide with the y-axis. Therefore, $F^\alpha = \mathrm{D}^\alpha$ and both are regular distributions of rank 1. For $\alpha = x\mathrm{d}x$, $F^\alpha_{(x,y)}$ is given by the y-axis if $x \neq 0$ and by $\mathbb{R}^2$ otherwise. In contrast, since the characteristic vector fields are continuous, $D^\alpha_{(x,y)}$ coincides with the y-axis everywhere.

Proposition 4.2.20 *For every $\alpha \in \Omega^k(M)$, the characteristic distribution D^α is integrable.*

The corresponding foliation of M by the maximal integral manifolds of D^α is called the characteristic foliation of α.

Proof We show that D^α is homogeneous and apply Theorem 3.5.10. Let $X \in \mathfrak{X}^{D^\alpha}_{\mathrm{loc}}(M)$. Since X is characteristic for α, α is invariant with respect to X. Hence,

$$\Phi^{X*}_t \alpha = \alpha. \tag{4.2.14}$$

Now, let $(t, m) \in \mathscr{D}^X$. For $Y_m \in D^\alpha_m$, choose a characteristic vector field Y taking this value at m. Due to (4.2.14), $\Phi^X_{t*} Y$ is characteristic for α. Hence,

$$\left(\Phi^X_t\right)'_m Y_m = \left(\Phi^X_{t*} Y\right)_{\Phi^X_t(m)} \in D^\alpha_{\Phi^X_t(m)}.$$

[6]Nevertheless, in [181], F^α is referred to as the characteristic distribution of α, which is consistent with the notion of distribution used there.

It follows that $(\Phi_t^X)'_m D_m^\alpha \subset D^\alpha_{\Phi_t^X(m)}$. The same argument yields

$$\left(\Phi_{-t}^X\right)'_{\Phi_t^X(m)} D^\alpha_{\Phi_t^X(m)} \subset D_m^\alpha,$$

and hence equality. Thus, D^α is homogeneous and hence integrable. □

Remark 4.2.21 If $\dim F_m^\alpha$ is constant on M and hence $D^\alpha = F^\alpha$ is regular, it is enough to prove involutivity:

$$[X,Y]\lrcorner\alpha = (\mathscr{L}_X Y)\lrcorner\alpha = \mathscr{L}_X(Y\lrcorner\alpha) - Y\lrcorner\mathscr{L}_X\alpha = 0,$$
$$[X,Y]\lrcorner \mathrm{d}\alpha = \mathscr{L}_{[X,Y]}\alpha = \mathscr{L}_X\mathscr{L}_Y\alpha - \mathscr{L}_Y\mathscr{L}_X\alpha = 0$$

for arbitrary characteristic vector fields X, Y of α.

Exercises

4.2.1 Show the following.
 (a) The Möbius strip is not orientable.
 (b) The real projective space $\mathbb{R}P^n$ is orientable iff n is odd.
 (c) Level sets of smooth mappings $f : \mathbb{R}^n \to \mathbb{R}^m$ at regular values are orientable.

4.2.2 Show that the right hand side of Eq. (4.2.3) neither depends on the choice of the atlas nor on the choice of the partition of unity.

4.2.3 Prove the statements of Remark 4.2.9/1.

4.2.4 Prove the statements in Example 4.2.11.

4.2.5 Show that the closed unit disk $\mathrm{D}^n = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \leq 1\}$ is a smooth manifold with boundary and find the interior and the boundary.

4.2.6 Prove Proposition 4.2.13/1.

4.3 De Rham Cohomology

In this section we show that the Cartan algebra of differential forms on a manifold contains important information about the topology of this manifold.

Definition 4.3.1 Let M be a manifold.

1. A k-form α is called closed if $\mathrm{d}\alpha = 0$. The set of closed k-forms is denoted by $Z^k(M)$.
2. A k-form α is called exact if there exists a $(k-1)$-form β such that $\alpha = \mathrm{d}\beta$. In this case, β is called a potential for α. The set of exact k-forms is denoted by $B^k(M)$.

Both $Z^k(M)$ and $B^k(M)$ are vector spaces. Since $\mathrm{d}^2 = 0$, we have $B^k(M) \subset Z^k(M)$, that is, every exact form is closed.

Definition 4.3.2 Let M be a manifold.

1. The additive Abelian group underlying the vector space $H^k(M) := Z^k(M)/B^k(M)$ is called the k-th de Rham cohomology group of M.
2. The numbers $\dim H^i(M)$ are called the Betti numbers of M and the sum

$$\chi(M) := \sum_{i=0}^{n} (-1)^i \dim H^i(M) \tag{4.3.1}$$

is called the Euler characteristic of M.

Remark 4.3.3

1. For $k > \dim M$ we have $Z^k(M) = 0$ and thus $H^k(M) = 0$.
2. Consider the direct sums

$$Z^*(M) := \bigoplus_{k=0}^{n} Z^k(M), \qquad B^*(M) := \bigoplus_{k=0}^{n} B^k(M).$$

For $\alpha \in Z^k(M)$ and $\beta \in Z^r(M)$ we have

$$\mathrm{d}(\alpha \wedge \beta) = (\mathrm{d}\alpha) \wedge \beta + (-1)^k \alpha \wedge (\mathrm{d}\beta) = 0,$$

that is, $\alpha \wedge \beta \in Z^{k+r}(M)$. Thus, $Z^*(M)$ is a subalgebra of $\Omega^*(M)$. For $\alpha \in B^k(M)$ with potential $\tilde{\alpha}$ and for $\beta \in Z^r(M)$ we get

$$\mathrm{d}(\tilde{\alpha} \wedge \beta) = (\mathrm{d}\tilde{\alpha}) \wedge \beta = \alpha \wedge \beta,$$

that is, $\alpha \wedge \beta \in B^{k+r}(M)$. The same is true for $\alpha \in Z^k(M)$ and $\beta \in B^r(M)$. Thus, $B^*(M)$ is a two-sided ideal in $Z^*(M)$. We conclude that the exterior product induces a multiplication in $H^*(M) := Z^*(M)/B^*(M)$,

$$[\alpha] \cup [\beta] := [\alpha \wedge \beta],$$

called the cup product. This way, $H^*(M)$ becomes an associative, graded commutative algebra with grading

$$H^*(M) = \bigoplus_{k=0}^{n} H^k(M).$$

$H^*(M)$ is called the cohomology algebra of M and the underlying ring is called the cohomology ring of M.
3. Let $\varphi\colon M_1 \to M_2$ be a smooth mapping. Since the pull-back by φ commutes with the operation of taking the exterior differential, it maps closed forms to closed forms and exact forms to exact forms. Since it is linear, it induces a homomorphism of the cohomology groups, denoted by the same symbol:

$$\varphi^*\colon H^k(M_2) \to H^k(M_1), \qquad \varphi^*\big([\alpha]\big) := \big[\varphi^*(\alpha)\big]. \tag{4.3.2}$$

Equation (4.1.15) implies

$$\varphi^*\big([\alpha]\cup[\beta]\big) = \big(\varphi^*[\alpha]\big)\cup\big(\varphi^*[\beta]\big),$$

that is, φ^* is a homomorphism of algebras. Moreover,

$$(\chi\circ\varphi)^* = \varphi^*\circ\chi^*, \qquad \mathrm{id}_M^* = \mathrm{id}_{H^*(M)}. \tag{4.3.3}$$

Example 4.3.4

1. We compute $H^0(M)$. From $\mathrm{d}f = 0$ we conclude that the function f must be constant on each connected component of M. Thus, f is given by q real numbers, where q denotes the number of connected components, and we obtain $Z^0(M) = \mathbb{R}^q$. Since exact 0-forms do not exist, we have $B^0(M) = 0$ and thus
$$H^0(M) = \mathbb{R}^q.$$
2. Let $M = \mathbb{R}$. By point 1, we have $H^0(\mathbb{R}) = \mathbb{R}$. Let us calculate $H^1(\mathbb{R})$. Let $\alpha \in \Omega^1(\mathbb{R})$. Then, for dimensional reasons α is closed. Moreover, it is also exact. Indeed, writing $\alpha(x) = a(x)\mathrm{d}x$ and defining the function $f(x) := \int_0^x a(y)\mathrm{d}y$, we get
$$\mathrm{d}f(x) = \frac{\mathrm{d}f}{\mathrm{d}x}(x)\mathrm{d}x = a(x)\mathrm{d}x = \alpha(x).$$
Thus, $B^1(\mathbb{R}) = Z^1(\mathbb{R}) = \Omega^1(\mathbb{R})$ and we conclude $H^1(\mathbb{R}) = 0$.
3. Let $M = \mathrm{S}^1$. By point 1, we have $H^0(\mathrm{S}^1) = \mathbb{R}$. We show that $H^1(\mathrm{S}^1) = \mathbb{R}$. In the standard parameterization of S^1 by $x \in \mathbb{R}$, a 1-form α on S^1 is given by $\alpha = a(x)\mathrm{d}x$, with $a(x) = a(x + 2\pi n)$. Since $\dim \mathrm{S}^1 = 1$, α is closed. It is exact iff there exists a function f, fulfilling $f(x + 2\pi n) = f(x)$, such that $\alpha = \mathrm{d}f$, that is, $a(x) = \frac{\mathrm{d}f}{\mathrm{d}x}(x)$. Then, using $x \in (0, 2\pi)$ as a coordinate, we obtain
$$\int_{\mathrm{S}^1}\alpha = \int_0^{2\pi} a(x)\mathrm{d}x = \int_0^{2\pi}\frac{\mathrm{d}f}{\mathrm{d}x}(x)\mathrm{d}x = f(2\pi) - f(0) = 0.$$
This means that two forms α_1 and α_2 belong to the same cohomology class iff
$$\int_{\mathrm{S}^1}\alpha_1 = \int_{\mathrm{S}^1}\alpha_2.$$
Thus, $H^1(\mathrm{S}^1)$ is in one-to-one correspondence with the values of such integrals. If we put $\alpha = (2\pi)^{-1}r\mathrm{d}x$, $r \in \mathbb{R}$, then we get $\int_{\mathrm{S}^1}\alpha = r$. Thus, $H^1(\mathrm{S}^1) = \mathbb{R}$.
4. For the Euler characteristics we obtain $\chi(\mathbb{R}) = 1$ and $\chi(\mathrm{S}^1) = 0$. This shows, in particular, that $\mathbb{R}$ and S^1 are topologically distinct, cf. Corollary 4.3.10.
5. For the one-point space $\{*\}$ we get
$$H^k\big(\{*\}\big) = \begin{cases}\mathbb{R} & k = 0\\ 0 & k > 0.\end{cases}$$

Definition 4.3.5 (Smooth homotopy) Let M and N be manifolds.

1. A smooth homotopy of a smooth mapping $\varphi\colon M \to N$ is a smooth mapping $F\colon M \times [0,1] \to N$ such that $F(\cdot,0) = \varphi$. A smooth homotopy of two smooth mapping $\varphi_0, \varphi_1\colon M \to N$ is a smooth mapping $F\colon M \times [0,1] \to N$ such that $F(\cdot,i) = \varphi_i$, $i = 1,2$. If a smooth homotopy exists, φ_0 and φ_1 are said to be smoothly homotopic.
2. M and N are called smoothly homotopy-equivalent if there exist smooth mappings $\varphi\colon M \to N$ and $\chi : N \to M$ such that the compositions $\chi \circ \varphi\colon M \to M$ and $\varphi \circ \chi : N \to N$ are smoothly homotopic to the respective identical mapping.
3. M is called contractible if for some $m_0 \in M$ the constant mapping $M \ni m \mapsto m_0 \in M$ and the identical mapping id_M are smoothly homotopic.

Remark 4.3.6

1. There is an analogous notion of homotopy in the category of topological spaces and continuous mappings. In the case of smooth manifolds and smooth mappings, the two notions of homotopy are related as follows. If two smooth mappings are continuously homotopic, they are smoothly homotopic. Every continuous mapping between manifolds is continuously homotopic to a smooth mapping, see [130]. After having clarified this, in the sequel we may drop the adjective smooth and just speak of homotopies.
2. To be homotopic is an equivalence relation in the set of smooth mappings from N to M. The set of homotopy classes of mappings $\mathrm{S}^i \to M$, such that some chosen point $s_0 \in \mathrm{S}^i$ is mapped to some chosen point $x_0 \in M$, is called the i-th homotopy group[7] of M and is denoted by $\pi_i(M, x_0)$. If M is connected, $\pi_i(M, x_0)$ is naturally isomorphic to $\pi_i(M, x_1)$ for all $x_0, x_1 \in M$. Hence, in this case, it is common to suppress the base point x_0 in the notation and to simply write $\pi_i(M)$.
3. To be homotopy-equivalent is an equivalence relation in the totality of smooth manifolds. The equivalence class of a manifold M is called the homotopy type of M. This terminology extends in an obvious way to topological spaces.
4. Let M and N be compact oriented manifolds with volume forms v_M and v_N, respectively, and let $\varphi : M \to N$ be a smooth mapping. The mapping degree $\deg(\varphi)$ of φ is defined by
$$\deg(\varphi) \int_M \mathrm{v}_M = \int_M \varphi^* \mathrm{v}_N .$$
One can show that $\deg(\varphi)$ is an integer, see for example [76]. There, the reader can find an exhaustive discussion of this important topological concept. In the special case $M = N = \mathrm{S}^1$, the mapping degree can be interpreted as the winding number. It defines an isomorphism from $\pi_1(\mathrm{S}^1)$ to $\mathbb{Z}$ (Exercise 4.3.4).

[7] For $i = 1$, the group multiplication is induced from an appropriate composition of closed paths. For $i > 1$ there is an analogous construction, see [76] for a detailed presentation. In contrast, $\pi_0(M, x_0)$, which is in bijective correspondence to the set of connected components of M, in general does not carry a group structure.

5. A manifold is contractible iff it is homotopy-equivalent to the one-point space $\{*\}$: for $m_0 \in M$, let $\varrho_{m_0} : M \to M$ denote the constant mapping $m \mapsto m_0$. Define

$$\varphi\colon M \to \{*\}, \qquad \varphi(m) := *, \qquad \chi : \{*\} \to M, \qquad \chi(*) := m_0.$$

By construction, we have $\varphi \circ \chi = \mathrm{id}_{\{*\}}$ and $\chi \circ \varphi = \varrho_{m_0}$. Hence, $\chi \circ \varphi$ is homotopic to id_M iff so is ρ_{m_0}.

Now, let M be a manifold and let $I = [0, 1]$. Every $\Omega \in \Omega^k(M \times I)$ can be uniquely decomposed into $\Omega = \alpha_1 + \alpha_2 \wedge \mathrm{d}t$, where $\alpha_1(\cdot, t) \in \Omega^k(M)$ and $\alpha_2(\cdot, t) \in \Omega^{k-1}(M)$ for every $t \in I$. Consider the family of mappings

$$D_k\colon \Omega^k(M \times I) \to \Omega^{k-1}(M), \qquad (D_k\Omega)(m) := (-1)^{k-1} \int_0^1 \alpha_2(m, t)\mathrm{d}t, \tag{4.3.4}$$

where the integral is to be understood as an integral of a function on the interval I with values in the vector space $\bigwedge^{k-1} \mathrm{T}^*_m M$.

Lemma 4.3.7 *We have*

$$\mathrm{d}(D_k\Omega) + D_{k+1}(\mathrm{d}\Omega) = \Omega_{\restriction t=1} - \Omega_{\restriction t=0}. \tag{4.3.5}$$

Proof Let $m \in M$ and let (U, κ) be a local chart at m. Let

$$\alpha_2(m, t) = a_I(m, t)\mathrm{d}\kappa^I, \qquad \alpha_1(m, t) = b_J(m, t)\mathrm{d}\kappa^J$$

(summation convention) be the corresponding local formulae for the components of Ω. Then,

$$(D_k\Omega)(m) = (-1)^{k-1}\left(\int_0^1 a_I(m, t)\mathrm{d}t\right)\mathrm{d}\kappa^I$$

and thus

$$\big(\mathrm{d}(D_k\Omega)\big)(m) = (-1)^{k-1}\left(\int_0^1 \partial^\kappa_p a_I(m, t)\mathrm{d}t\right)\mathrm{d}\kappa^p \wedge \mathrm{d}\kappa^I.$$

On the other hand,

$$\begin{aligned}
&\big(D_{k+1}(\mathrm{d}\Omega)\big)(m)\\
&\quad= \big(D_{k+1}(\mathrm{d}\alpha_1)\big)(m) + \big(D_{k+1}(\mathrm{d}\alpha_2 \wedge \mathrm{d}t)\big)(m)\\
&\quad= \big(D_{k+1}\big(\partial^\kappa_q b_J \mathrm{d}\kappa^q \wedge \mathrm{d}\kappa^J + \partial_t b_J \mathrm{d}t \wedge \mathrm{d}\kappa^J\big)\big)(m)\\
&\qquad+ \big(D_{k+1}\big(\partial^\kappa_p a_I \mathrm{d}\kappa^p \wedge \mathrm{d}\kappa^I \wedge \mathrm{d}t\big)\big)(m)\\
&\quad= (-1)^k(-1)^k\left(\int_0^1 \partial_t b_J(m, t)\mathrm{d}t\right)\mathrm{d}\kappa^J
\end{aligned}$$

$$
\begin{aligned}
&+(-1)^k\left(\int_0^1 \partial_p^\kappa a_I(m,t)\mathrm{d}t\right)\mathrm{d}\kappa^p\wedge \mathrm{d}\kappa^I\\
&=\big(b_J(m,1)-b_J(m,0)\big)\mathrm{d}\kappa^J-\big(\mathrm{d}(D_k\Omega)\big)(m)\\
&=\Omega_{\restriction t=1}(m)-\Omega_{\restriction t=0}(m)-\big(\mathrm{d}(D_k\Omega)\big)(m).
\end{aligned}
$$

This completes the proof. □

Proposition 4.3.8 *Homotopic mappings induce the same homomorphism of the de Rham cohomology groups.*

Proof Let $\varphi_0,\varphi_1\colon M\to N$ be smooth mappings and let $F\colon M\times I\to N$ be a homotopy fulfilling $F(m,0)=\varphi_0(m)$ and $F(m,1)=\varphi_1(m)$. Let $\alpha\in\Omega^k(N)$. If we set $\Omega=F^*\alpha$ in (4.3.5), we get

$$\mathrm{d}D_kF^*\alpha+D_{k+1}\mathrm{d}F^*\alpha=\varphi_1^*\alpha-\varphi_0^*\alpha.$$

If α is a representative of a cohomology class in $H^k(N)$, then $\mathrm{d}D_kF^*\alpha=\varphi_1^*\alpha-\varphi_0^*\alpha$. Thus, $\varphi_1^*\alpha$ and $\varphi_0^*\alpha$ differ by an exact form, that is, they define the same cohomology class. □

Remark 4.3.9 Given smooth mappings $\varphi_0,\varphi_1:M\to N$, a family of mappings $H_k:\Omega^k(N)\to\Omega^{k-1}(M)$ fulfilling

$$H_{k+1}\circ\mathrm{d}+\mathrm{d}\circ H_k=\varphi_1^*-\varphi_0^*$$

is called a homotopy operator for φ_0 and φ_1. If F is a homotopy of φ_0 and φ_1, then $H_k=D_k\circ F^*$ is a homotopy operator for φ_0 and φ_1.

Corollary 4.3.10 *Homotopy-equivalent manifolds possess isomorphic de Rham cohomology groups.*

Proof Let M and N be homotopy-equivalent and let $\varphi\colon M\to N$ and $\chi:N\to M$ be smooth mappings yielding this equivalence. We consider the induced homomorphisms $\varphi^*\colon H^k(N)\to H^k(M)$ and $\chi^*\colon H^k(M)\to H^k(N)$. Since $\chi\circ\varphi\sim \mathrm{id}_M$ and $\varphi\circ\chi\sim\mathrm{id}_N$, respectively, Proposition 4.3.8 and (4.3.3) imply

$$\varphi^*\circ\chi^*=(\chi\circ\varphi)^*=\mathrm{id}_{M_1}^*=\mathrm{id}_{H^k(M_1)},$$

and, in the same way, $\chi^*\circ\varphi^*=\mathrm{id}_{H^k(M_2)}$. Thus, φ^* and χ^* are isomorphisms, which are inverse to each other. □

Corollary 4.3.11 (Lemma of Poincaré) *If M is contractible, then*

$$H^k(M)=\begin{cases}\mathbb{R} & \mid k=0\\ 0 & \mid k>0.\end{cases}$$

Proof By Remark 4.3.6/5, M is homotopy-equivalent to the one-point space. Thus, Corollary 4.3.10 implies $H^k(M) \cong H^k(\{*\})$. □

Remark 4.3.12

1. The Lemma of Poincaré tells us, in particular, that on a contractible manifold every closed form is exact.
2. The operator D_k, defined by (4.3.4), yields an explicit construction for potentials of a closed form on a contractible manifold: let $F: M \times I \to M$ be a homotopy of M to m_0, that is, $F(m, 0) = m$ and $F(m, 1) = m_0$. By (4.3.5), we have $\mathrm{d}D_k F^*\alpha + D_{k+1}\mathrm{d}F^*\alpha = -\alpha$ for all $\alpha \in \Omega^k(M)$. If $\mathrm{d}\alpha = 0$, we read off the following potential for α:

$$\beta = -D_k F^*\alpha. \tag{4.3.6}$$

Example 4.3.13 Let $\alpha \in \Omega^2(\mathbb{R}^3)$ be given by

$$\alpha = xy\mathrm{d}x \wedge \mathrm{d}y + 2x\mathrm{d}y \wedge \mathrm{d}z + 2y\mathrm{d}x \wedge \mathrm{d}z,$$

in standard coordinates x, y and z on $\mathbb{R}^3$. Obviously, $\mathrm{d}\alpha = 0$. To construct a potential we choose a homotopy to the origin:

$$F\big((x, y, z), t\big) = (tx, ty, tz), \quad t \in [0, 1].$$

Then,

$$\begin{aligned} F^*\alpha =& \big(xy^2t^3 + 2yzt^2\big)\mathrm{d}x \wedge \mathrm{d}t + \big(2xzt^2 - x^2yt^3\big)\mathrm{d}y \wedge \mathrm{d}t - 4xyt^2\mathrm{d}z \wedge \mathrm{d}t \\ &+ xyt^4\mathrm{d}x \wedge \mathrm{d}y + 2xt^3\mathrm{d}y \wedge \mathrm{d}z + 2yt^3\mathrm{d}x \wedge \mathrm{d}z. \end{aligned}$$

Thus, we obtain the potential

$$\begin{aligned} \beta =& - D_2F^*\alpha \\ =& (-1)^{2-1}\bigg[\bigg(\int_0^1 \big(xy^2t^3 + 2yzt^2\big)\mathrm{d}t\bigg)\mathrm{d}x + \bigg(\int_0^1 \big(2xzt^2 - x^2yt^3\big)\mathrm{d}t\bigg)\mathrm{d}y \\ &- \bigg(\int_0^1 4xyt^2\mathrm{d}t\bigg)\mathrm{d}z\bigg] \\ =& -\bigg(\frac{1}{4}xy^2 + \frac{2}{3}yz\bigg)\mathrm{d}x + \bigg(\frac{1}{4}x^2y - \frac{2}{3}xz\bigg)\mathrm{d}y + \frac{4}{3}xy\mathrm{d}z. \end{aligned}$$

The following generalization of the Poincaré Lemma will be useful later on. The proof uses the Tubular Neighbourhood Theorem for embedded submanifolds, cf. Remark 6.4.7, and the fact that the normal bundle of a submanifold N is homotopy equivalent to N. We leave the latter as an advanced exercise to the reader (Exercise 4.3.5).

Proposition 4.3.14 (Generalized Poincaré Lemma) *Let (N, φ) be an embedded submanifold of M and let α be a closed k-form fulfilling $\varphi^*\alpha = 0$. Then, there exists a $(k-1)$-form β on a neighbourhood of N in M fulfilling*

$$\alpha = \mathrm{d}\beta, \qquad \beta_{\restriction N} = 0.$$

If, in addition, α vanishes[8] *on N, then β can be chosen such that, in any local chart, the first derivatives of the coefficients of β vanish on N.*

Algebraic topology provides a lot of tools for calculating the de Rham cohomology groups of a manifold, see [52], [96] and also [55]. Here, we only make some elementary remarks. First, the cohomology groups of a direct product $M \times N$ can be computed by means of the Künneth Formula:

$$H^r(M \times N) = \bigoplus_{p+q=r} H^p(M) \otimes H^q(N). \tag{4.3.7}$$

Thus,

$$H^*(M \times N) = H^*(M) \otimes H^*(N), \tag{4.3.8}$$

which is to be understood as a tensor product of associative algebras.

Example 4.3.15 Consider the 2-torus $\mathrm{T}^2 = \mathrm{S}^1 \times \mathrm{S}^1$. We calculate its de Rham cohomology groups:

$$\begin{aligned} H^0(\mathrm{T}^2) &= H^0(\mathrm{S}^1) \otimes H^0(\mathrm{S}^1) = \mathbb{R} \otimes \mathbb{R} = \mathbb{R} \\ H^1(\mathrm{T}^2) &= \big(H^1(\mathrm{S}^1) \otimes H^0(\mathrm{S}^1)\big) \oplus \big(H^0(\mathrm{S}^1) \otimes H^1(\mathrm{S}^1)\big) \\ &= (\mathbb{R} \otimes \mathbb{R}) \oplus (\mathbb{R} \otimes \mathbb{R}) = \mathbb{R}^2 \\ H^2(\mathrm{T}^2) &= H^1(\mathrm{S}^1) \otimes H^1(\mathrm{S}^1) = \mathbb{R} \otimes \mathbb{R} = \mathbb{R}. \end{aligned}$$

As a second tool, one has the following integral criterion for exact forms.

Proposition 4.3.16 *A closed differential k-form α on a manifold M is exact iff $\int_N \alpha = 0$, for every compact k-dimensional submanifold N of M.*

We give a sketch of the proof. Assume that $\alpha \in \Omega^k(M)$ is exact, $\alpha = \mathrm{d}\beta$. Then, using $\partial N = \varnothing$ and Stokes' Theorem, we get $\int_N \alpha = \int_{\partial N} \beta = 0$. The converse follows from the de Rham Theorem, which states that there exists a natural isomorphism between the de Rham cohomology algebra and the real-valued differentiable singular cohomology algebra.[9] The latter is dual to the differentiable singular homology. We provide the reader with a rough idea of these notions. Singular

[8] If α vanishes on N, then $\varphi^*\alpha$ vanishes, too. Note that the converse is, of course, not true in general.

[9] For a proof see [302], Sects. 4.17, 5.36 and 5.45.

homology theory is built from formal real linear combinations (called k-chains) $\sigma = \sum_i \lambda_i \sigma_i$ of differentiable mappings (called differentiable singular simplices) $\sigma_i : \Delta^k \to M$, where

$$\Delta^k = \left\{ \mathbf{x} \in \mathbb{R}^k : x^i \geq 0, x^1 + \cdots + x^k \leq 1 \right\}$$

denotes the standard k-dimensional simplex. There is a natural boundary operator defined by

$$\partial \sigma_i = \sum_{j=0}^{k} (-1)^j \sigma_{i \restriction \Delta_j^k},$$

where $\Delta_1^k, \ldots, \Delta_k^k$ denote the faces of Δ^k, defined by $x^k = 0$, and Δ_0^k denotes the remaining face. This operator is extended to arbitrary chains by linearity. One easily shows that $\partial \circ \partial = 0$. Now, the k-th differential singular homology group of M with real coefficients is defined by $\ker \partial_k / \operatorname{im} \partial_{k+1}$. Elements of $\ker \partial_k$ and $\operatorname{im} \partial_{k+1}$ are called differentiable k-cycles and differentiable k-boundaries, respectively. The corresponding differentiable singular cohomology is then built from linear functionals (called singular cochains), which assign to each singular simplex a number. Now, the link to de Rham cohomology is provided by defining how to evaluate a closed k-form α on a k-cycle $\sigma = \sum_i \lambda_i \sigma_i$:

$$\langle \alpha, \sigma \rangle := \sum_i \lambda_i \int_{\Delta^k} \sigma_i^* \alpha. \tag{4.3.9}$$

This induces a linear mapping from the k-th de Rham cohomology into the real differentiable singular cohomology, which turns out to be an isomorphism. It can be easily seen that the integrals over all k-dimensional compact submanifolds of a closed k-form α vanish iff α vanishes on every k-cycle. Then, injectivity of the isomorphism defined by (4.3.9) implies that a closed form vanishing on all k-cycles must be exact.

Example 4.3.17 As an application of Proposition 4.3.16, let us calculate the de Rham cohomology groups of the spheres S^n, $n \geq 1$:

$$H^k\left(S^n\right) = \begin{cases} \mathbb{R} & \mid k = 0, n, \\ 0 & \mid \text{otherwise}. \end{cases} \tag{4.3.10}$$

The case $k = 0$ was dealt with in Example 4.3.4/1. Let $0 < k < n$ and let $\alpha \in \Omega^k(S^n)$ be closed. Let $N \subset S^n$ be an arbitrary compact k-dimensional submanifold of S^n. Then, there exists a point $\mathbf{x} \in S^n$ such that N is contained in the open submanifold $S^n_{\mathbf{x}} := S^n \setminus \{\mathbf{x}\} \subset S^n$. Since $S^n_{\mathbf{x}}$ is contractible, according to the Poincaré Lemma, $\alpha_{\restriction S^n_{\mathbf{x}}}$ is exact. Applying Proposition 4.3.16 to $M = S^n_{\mathbf{x}}$ we get $\int_N \alpha = 0$. Since N was arbitrary, the same proposition, applied to $M = S^n$, yields that α is exact on S^n. We conclude $H^k(S^n) = 0$ for $0 < k < n$. Consider the case $k = n$. By the theorem on invariance of domain, see Footnote 38 on page 159, the only compact n-dimensional submanifold of S^n is S^n itself. Hence, Proposition 4.3.16 entails that $H^n(\mathrm{S}^n)$ is in one-to-one correspondence with values of the integrals $\int_{\mathrm{S}^n} \alpha$. Since S^n is orientable, it admits a volume form v. By multiplying v by a real number, one can obtain any real value for the above integral. Thus, $H^n(\mathrm{S}^n) = \mathbb{R}$. It also follows that $\chi(\mathrm{S}^n) = 0$ if n is odd and $\chi(\mathrm{S}^n) = 2$ if n is even.

Exercises

4.3.1 Check that the 1-form β constructed in Example 4.3.13 is a potential for the 2-form α, indeed.

4.3.2 Calculate the de Rham cohomology groups for the annulus

$$A = \left\{ \mathbf{x} \in \mathbb{R}^2 : 1 < \left(x^2 + y^2\right)^{\frac{1}{2}} < 2 \right\}.$$

4.3.3 Consider the following 1-form on $\mathbb{R}^2 \setminus \{0\}$:

$$\alpha = \frac{x dy - y dx}{x^2 + y^2}.$$

(a) Prove that α is closed. Calculate the integral of the restriction of α to the unit circle $\mathrm{S}^1 \subset \mathbb{R}^2 \setminus \{0\}$ and explain why α is not exact.

(b) Show that the curves $\gamma_n : [0, 1] \to \mathbb{R}^2 \setminus \{0\}$, $\gamma_n(t) = (\cos(2\pi n t), \sin(2\pi n t))$ are not homotopic for different values of $n \in \mathbb{Z}$.
Hint. Calculate the integrals of α over γ_n.

(c) Let β be another closed 1-form on $\mathbb{R}^2 \setminus \{0\}$. Show that there exists a number c and a function $f : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}$ such that $\beta = c\alpha + \mathrm{d}f$. Conclude that α generates the first de Rham cohomology of $\mathbb{R}^2 \setminus \{0\}$.

4.3.4 Prove that the mapping degree defines an isomorphism from $\pi_1(\mathrm{S}^1)$ to $\mathbb{Z}$, cf. Remark 4.3.6/4.

4.3.5 Show that every vector bundle is homotopy equivalent to its base manifold. Use this and the Tubular Neighbourhood Theorem for embedded submanifolds, stated in Remark 6.4.7, to prove the Generalized Poincaré Lemma 4.3.14.

4.4 Riemannian Manifolds

In this section, we discuss some elements of Riemannian geometry needed in the sequel. A systematic treatment will be contained in Part II of this book.

Definition 4.4.1 (Riemannian manifold) Let M be an n-dimensional manifold. A metric on M is a covariant tensor field $\mathrm{g} \in \Gamma(\mathbb{T}^0_2(M))$ such that for every $m \in M$ the bilinear form

$$\mathrm{g}_m : \mathrm{T}_m M \times \mathrm{T}_m M \to \mathbb{R}$$

is symmetric and non-degenerate. If g_m is positive-definite for all $m \in M$, then g is called a Riemannian metric and the pair (M, g) is called a Riemannian manifold. Otherwise, g is called a pseudo-Riemannian metric and the pair (M, g) is called a pseudo-Riemannian manifold.

Let $m \in M$. It is well-known from linear algebra that there exists an orthonormal basis $\{e_i\}$ in $\mathrm{T}_m M$, that is,

$$\mathrm{g}_m(e_i, e_j) = \begin{pmatrix} +\mathbb{1}_r & 0 \\ 0 & -\mathbb{1}_s \end{pmatrix}.$$

We denote $\eta_{ij}(m) = \mathrm{g}_m(e_i, e_j)$. According to the Theorem of Sylvester, the integers r and s do not depend on the choice of the basis. Using Proposition 2.3.15 and an inductive orthonormalization procedure one can show that $\{e_i\}$ can be extended to a local orthonormal frame[10] at m_0 (Exercise 4.4.1). Thus, $\mathrm{g}(e_i, e_j) = \eta_{ij}$ in a neighbourhood of m_0, and the numbers r and s are constant on each connected component of M. The pair (r, s) is called the signature[11] and s is called the index of g.

Proposition 4.4.2 *Every manifold admits a Riemannian metric.*

Proof Choose a countable atlas $\{(U_i, \kappa_i) : i \in I\}$ on M and a subordinate partition of unity $\{f_i\}$. Via κ_i, the standard scalar product on $\mathbb{R}^n$ induces a Riemannian metric g_i on U_i. For every i, the tensor field $f_i \mathrm{g}_i$ extends to a smooth tensor field $\tilde{\mathrm{g}}_i$ on M with $\mathrm{supp}(\tilde{\mathrm{g}}_i) = \mathrm{supp}(f_i)$. Since the covering $\{\mathrm{supp}(f_i)\}$ is locally finite, we can define $\mathrm{g} := \sum_i \tilde{\mathrm{g}}_i$ and this is a smooth covariant tensor field of second order on M. Since on the interiors of their supports, the $\tilde{\mathrm{g}}_i$ are Riemannian metrics, and since the sum of Riemannian metrics is a Riemannian metric, g is a Riemannian metric. □

Since g_m is a non-degenerate bilinear form, it can be viewed as an isomorphism of vector spaces:

$$\mathrm{g}_m : \mathrm{T}_m M \to \mathrm{T}^*_m M, \qquad X_m \mapsto \mathrm{g}_m(X_m, \cdot). \tag{4.4.1}$$

The isomorphisms g_m combine to a vector bundle isomorphism $\mathrm{g} : \mathrm{T}M \to \mathrm{T}^*M$, which induces isomorphisms

$$\mathrm{g} : \bigwedge^k \mathrm{T}M \to \bigwedge^k \mathrm{T}^*M, \qquad \mathrm{g}(X_1 \wedge \cdots \wedge X_k) = \mathrm{g}(X_1) \wedge \cdots \wedge \mathrm{g}(X_k). \tag{4.4.2}$$

The metric g induces natural symmetric non-degenerate bilinear forms g_m and g_m^{-1} on $\bigwedge^k \mathrm{T}_m M$ and $\bigwedge^k \mathrm{T}^*_m M$, respectively:

$$\mathrm{g}_m(X_1 \wedge \cdots \wedge X_k, Y_1 \wedge \cdots \wedge Y_k) := \det\big(\mathrm{g}_m(X_i, Y_j)\big), \tag{4.4.3}$$

[10]Denoted by the same letter.

[11]Some authors prefer to call $t = r - s$ the signature.

$$g_m^{-1}(\alpha, \beta) := g_m\big(g_m^{-1}(\alpha), g_m^{-1}(\beta)\big). \tag{4.4.4}$$

Remark 4.4.3 Let us write down the isomorphisms (4.4.1) and (4.4.2) in a local chart (U, κ). We denote

$$g_{ij} := g\big(\partial_i^\kappa, \partial_j^\kappa\big), \qquad g^{ij} := g^{-1}\big(d\kappa^i, d\kappa^j\big).$$

Then,

$$g = g_{ij} d\kappa^i \overset{s}{\otimes} d\kappa^j, \qquad g^{-1} = g^{ij} \partial_i^\kappa \overset{s}{\otimes} \partial_j^\kappa, \tag{4.4.5}$$

with $\overset{s}{\otimes}$ denoting the symmetric tensor product. Since, as mappings, $g \circ g^{-1} = \mathrm{id}$, we have

$$g^{ij} g_{jk} = \delta^i{}_k. \tag{4.4.6}$$

In these notations, the isomorphism (4.4.1) and its inverse take the form

$$g\big(X^i \partial_i^\kappa\big) = X^i g_{ij} d\kappa^j, \qquad g^{-1}\big(\alpha_i d\kappa^i\big) = \alpha_i g^{ij} \partial_j^\kappa. \tag{4.4.7}$$

In the physics literature, one usually writes $X_j = X^i g_{ij}$ and $\alpha^j = \alpha_i g^{ij}$ and one says that by the help of the metric one raises or lowers indices. Similarly, we denote

$$g_{IJ} := g\big(\partial_I^\kappa, \partial_J^\kappa\big), \qquad g^{IJ} := g^{-1}\big(d\kappa^I, d\kappa^J\big).$$

We have

$$g_{IJ}(m) = \det\big(g_{i_r j_s}(m)\big), \qquad g_{IJ} g^{JK} = \delta_I^K,$$

and we can use g_{IJ} and g^{IJ} to raise and lower indices:

$$X_I = g_{IJ} X^J, \qquad \alpha^I = g^{IJ} \alpha_J.$$

Then,

$$g^{-1}(\alpha) = \alpha^J \partial_J^\kappa, \qquad g^{-1}(\alpha, \beta) = \alpha^I \beta^J g_{IJ}. \tag{4.4.8}$$

If we assume M to be oriented, there exists a distinguished volume form: choose an oriented orthonormal basis $\{e_i\}$ in $T_m M$ and define

$$(v_g)_m(X_1, \ldots, X_n) := \det\big(g_m(X_i, e_j)\big), \quad X_i \in T_m M. \tag{4.4.9}$$

This definition is independent of the choice of the basis, because for another oriented orthonormal basis $\{e'_i\}$ one has $e'_i = A_i{}^j e_j$ with $\det A = +1$. Since $\{e_i\}$ can be extended to an orthonormal local frame in TM, $(v_g)_m$ depends smoothly on m.

Definition 4.4.4 The differential n-form v_g defined by (4.4.9) is called the canonical volume form of (M, g).

Let $\{e_i\}$ be an oriented orthonormal local frame in $\mathrm{T}M$ and let $\{\vartheta^j\}$ be the dual frame in T^*M. Then,

$$\mathrm{g} = \eta_{ij}\vartheta^i \overset{s}{\otimes} \vartheta^j \tag{4.4.10}$$

and

$$\mathrm{g}(e_i) = \eta_{ij}\vartheta^j, \qquad \mathrm{g}^{-1}\big(\vartheta^i\big) = \eta^{ij}e_j. \tag{4.4.11}$$

Analogously,

$$\mathrm{g}(e_I) = \eta_{IJ}\vartheta^J, \qquad \mathrm{g}^{-1}\big(\vartheta^I\big) = \eta^{IJ}e_J. \tag{4.4.12}$$

Moreover,

$$\mathsf{v}_\mathrm{g}(e_1,\dots,e_n) = \det(\eta_{ij}) = (-1)^s, \qquad \mathrm{g}^{-1}(\mathsf{v}_\mathrm{g}) = e_1 \wedge \dots \wedge e_n. \tag{4.4.13}$$

Thus, denoting $I_n = \{1,\dots,n\}$, we obtain

$$\mathsf{v}_\mathrm{g} = (-1)^s \vartheta^{I_n}. \tag{4.4.14}$$

Remark 4.4.5 We determine the representative of v_g in a local chart (U,κ). There exists a positive definite function f such that

$$\mathsf{v}_\mathrm{g} = f\,\mathrm{d}\kappa^1 \wedge \dots \wedge \mathrm{d}\kappa^n \equiv f\,\mathrm{d}\kappa^{I_n}. \tag{4.4.15}$$

Writing $e_i = e_i{}^j \partial^\kappa_j$ and $\vartheta^i = \vartheta^i{}_j \mathrm{d}\kappa^j$ and using (4.4.14), we obtain

$$\mathsf{v}_\mathrm{g} = (-1)^s \vartheta^1{}_{i_1} \cdots \vartheta^n{}_{i_n} \mathrm{d}\kappa^{i_1} \wedge \dots \wedge \mathrm{d}\kappa^{i_n} = (-1)^s \det\vartheta\,\mathrm{d}\kappa^{I_n}.$$

Hence, $f(m) = (-1)^s \cdot \det(\vartheta(m)) > 0$. Moreover, from $e_i{}^k\vartheta^j{}_k = \delta_i{}^j$ and $\mathrm{g}_{lm}e_i{}^l e_j{}^m = \eta_{ij}$ we conclude

$$\det e \cdot \det\vartheta = 1, \qquad \det\mathrm{g} \cdot (\det e)^2 = (-1)^s, \tag{4.4.16}$$

and thus

$$\det\vartheta = \frac{1}{\det e} = \pm\sqrt{|\det\mathrm{g}|}. \tag{4.4.17}$$

In order to have $f > 0$, we must choose $\det\vartheta = (-1)^s\sqrt{|\det\mathrm{g}|}$. Then,

$$\mathsf{v}_\mathrm{g} = \sqrt{|\det\mathrm{g}|}\mathrm{d}\kappa^1 \wedge \dots \wedge \mathrm{d}\kappa^n = \frac{1}{n!}\sqrt{|\det\mathrm{g}|}\varepsilon_{i_1\dots i_n}\mathrm{d}\kappa^{i_1} \wedge \dots \wedge \mathrm{d}\kappa^{i_n},$$

with $\varepsilon_{i_1\dots i_n}$ denoting the completely antisymmetric tensor of rank n. In the physics literature, one often uses the Levi-Civita tensor $\varepsilon^{LC}_{i_1\dots i_n} := \sqrt{|\det\mathrm{g}|}\varepsilon_{i_1\dots i_n}$. Then, the volume form reads

$$\mathsf{v}_\mathrm{g} = \frac{1}{n!}\varepsilon^{LC}_{i_1\dots i_n}\mathrm{d}\kappa^{i_1} \wedge \dots \wedge \mathrm{d}\kappa^{i_n}. \tag{4.4.18}$$

Exercises

4.4.1 Let (M, g) be an oriented pseudo-Riemannian manifold and let $m \in M$. Show that every oriented orthonormal basis in $\mathrm{T}_m M$ can be extended to an oriented orthonormal local frame in $\mathrm{T}M$ at m.

4.5 Hodge Duality

In this section, we discuss the Hodge star operator and the operator dual to the exterior derivative in the sense of Hodge. Finally, we build the bridge to classical vector analysis. Throughout this section, let (M, g) be an n-dimensional oriented pseudo-Riemannian manifold with signature (r, s).

The isomorphism (4.4.2) and the canonical volume form v_g induce the following natural linear mapping.

Definition 4.5.1 (Hodge star operator) The linear mapping

$$*\colon \Omega^k(M) \to \Omega^{n-k}(M), \qquad *\alpha := (-1)^s \mathrm{g}^{-1}(\alpha) \lrcorner \mathsf{v}_\mathrm{g}, \tag{4.5.1}$$

is called the Hodge star operator of (M, g). The $(n-k)$-form $*\alpha$ is called the Hodge dual of α.

Obviously,

$$*1 = (-1)^s \mathsf{v}_\mathrm{g}. \tag{4.5.2}$$

Moreover, from (4.4.13) we read off

$$*\mathsf{v}_\mathrm{g} = 1. \tag{4.5.3}$$

For $I \subset I_n$, let I^c denote the complement. Formula (4.1.5) implies

$$e_I \lrcorner \vartheta^{I_n} = \rho_{I,I^c} \vartheta^{I^c} \equiv \operatorname{sign}\begin{pmatrix} I_n \\ I\,I^c \end{pmatrix} \vartheta^{I^c}.$$

Using this, together with (4.4.12) and (4.4.14), we obtain

$$*\vartheta^I = (-1)^s \mathrm{g}^{-1}\big(\vartheta^I\big) \lrcorner (-1)^s \vartheta^{I_n} = \eta^{IJ} e_J \lrcorner \vartheta^{I_n} = \operatorname{sign}\begin{pmatrix} I_n \\ J\,J^c \end{pmatrix} \eta^{IJ} \vartheta^{J^c}. \tag{4.5.4}$$

Let (U, κ) be an oriented local chart. Then,

$$\begin{aligned}
*\alpha &= (-1)^s \mathrm{g}^{-1}(\alpha) \lrcorner \mathsf{v}_\mathrm{g} \\
&= (-1)^s \big(\alpha^I \partial_I^\kappa\big) \lrcorner \sqrt{|\det \mathrm{g}|}\, \mathrm{d}\kappa^{I_n} \\
&= (-1)^s \sqrt{|\det \mathrm{g}|}\, \alpha^I \operatorname{sign}\begin{pmatrix} I_n \\ I\,I^c \end{pmatrix} \mathrm{d}\kappa^{I^c} \\
&= (-1)^s \alpha^I \varepsilon^{LC}_{I I^c} \mathrm{d}\kappa^{I^c}.
\end{aligned} \tag{4.5.5}$$

Remark 4.5.2 One can show that the Hodge dual $*\alpha$ of $\alpha \in \Omega^k(M)$ is the unique $(n-k)$-form such that

$$\mathrm{g}^{-1}(*\alpha, \beta)\mathrm{v_g} = \alpha \wedge \beta \tag{4.5.6}$$

for all $\beta \in \Omega^{n-k}(M)$ (Exercise 4.5.1). Hence, this formula can be taken as an alternative definition for the Hodge star operator.

Proposition 4.5.3 *For* $\alpha, \beta \in \Omega^k(M)$,

$$* * \alpha = (-1)^{k(n-k)+s}\alpha, \tag{4.5.7}$$

$$\mathrm{g}^{-1}(*\alpha, *\beta) = (-1)^s \mathrm{g}^{-1}(\alpha, \beta), \tag{4.5.8}$$

$$\alpha \wedge *\beta = (-1)^s \mathrm{g}^{-1}(\alpha, \beta)\mathrm{v_g}. \tag{4.5.9}$$

Proof Since the $*$-operator is linear, it is enough to show (4.5.7) for $\alpha = \vartheta^I$. The latter follows from (4.5.4) (Exercise 4.5.2). Then, using (4.5.6) and (4.5.7), we calculate

$$\begin{aligned}\mathrm{g}^{-1}(*\alpha, *\beta)\mathrm{v_g} &= \alpha \wedge *\beta \\ &= (-1)^{k(n-k)}(*\beta) \wedge \alpha \\ &= (-1)^{k(n-k)}\mathrm{g}^{-1}(* * \beta, \alpha)\mathrm{v_g} \\ &= (-1)^s \mathrm{g}^{-1}(\alpha, \beta)\mathrm{v_g}.\end{aligned}$$

This identity entails both Formulae (4.5.8) and (4.5.9). □

A k-form α on M is said to be square-integrable if the n-form $\alpha \wedge *\alpha$ is integrable. On the subspace of square-integrable k-forms we define an inner product by

$$(\alpha, \beta) := \int_M \alpha \wedge *\beta. \tag{4.5.10}$$

This inner product is positive definite, and hence a scalar product, iff (M, g) is Riemannian, that is, iff $s = 0$. It satisfies (Exercise 4.5.3)

$$(*\alpha, *\beta) = (-1)^s(\alpha, \beta). \tag{4.5.11}$$

Remark 4.5.4 For an oriented local chart (U, κ), Formulae (4.4.7) and (4.5.9) imply

$$\alpha \wedge *\beta = (-1)^s \sqrt{|\det \mathrm{g}|}\alpha^I \beta^J \mathrm{g}_{IJ} \mathrm{d}\kappa^{I_n}. \tag{4.5.12}$$

Definition 4.5.5 Let (M, g) be an oriented pseudo-Riemannian manifold.

1. The operator $\mathrm{d}^* : \Omega^k(M) \to \Omega^{k-1}(M)$ defined by

$$\big(\mathrm{d}^*\alpha, \beta\big) := (\alpha, \mathrm{d}\beta)$$

   for all $\beta \in \Omega^{k-1}(M)$ is called the Hodge dual of the exterior derivative d.
2. The operator $\Box\colon\ \Omega^k(M) \to \Omega^k(M)$ defined by

$$\Box := \mathrm{d}\mathrm{d}^* + \mathrm{d}^*\mathrm{d}$$

   is called the Hodge-Laplace operator of (M, g).

Proposition 4.5.6 *For* $\alpha \in \Omega^k(M)$,

$$\mathrm{d}^*\alpha = (-1)^{n(k-1)+s+1} * \mathrm{d} * \alpha. \tag{4.5.13}$$

Proof Since M has no boundary, Stokes' Theorem implies

$$\begin{aligned}\big(\beta, \mathrm{d}^*\alpha\big) &= \int_M \mathrm{d}\beta \wedge *\alpha = \int_M \beta \wedge (-1)^k \mathrm{d} * \alpha \\ &= \int_M \beta \wedge *\big\{(-1)^{n(k-1)+s+1} * \mathrm{d} * \alpha\big\}.\end{aligned}$$

□

Remark 4.5.7 The Hodge dual d^* satisfies $(\mathrm{d}^*)^2 = 0$, but in contrast to the exterior derivative, it is not an anti-derivation on $\Omega^*(M)$.

Example 4.5.8 (Classical vector analysis) Let (M, g) be $\mathbb{R}^n$ with the Euclidean metric. With respect to the standard orientation, the volume form is given by $\mathsf{v}_\mathrm{g} = \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$. The natural isomorphism

$$\mathfrak{X}(M) \to \Omega^1(M), \qquad X \mapsto \mathrm{g} \circ X$$

relates classical vector analysis to the theory of differential forms. The classical differential operators of gradient and divergence may be defined in a coordinate-free manner as follows.

1. For $f \in C^\infty(M)$, the gradient of f is the vector field

$$\operatorname{grad} f = \mathrm{g}^{-1} \circ (\mathrm{d}f). \tag{4.5.14}$$

2. For $X \in \mathfrak{X}(M)$, the divergence of X is the function

$$\operatorname{div} X = *\mathrm{d}\big(*(\mathrm{g} \circ X)\big) \equiv -\mathrm{d}^*(\mathrm{g} \circ X). \tag{4.5.15}$$

In case $n = 3$, there is a third classical differential operator on vector fields, the curl:

3. For $X \in \mathfrak{X}(M)$, the curl of X is the vector field

$$\operatorname{curl} X = \mathrm{g}^{-1} \circ \big(*\mathrm{d}(\mathrm{g} \circ X)\big) \equiv \mathrm{d}^*(X \lrcorner \mathrm{v_g}). \tag{4.5.16}$$

Formulae (4.5.14)–(4.5.16) make sense on any oriented (pseudo-)Riemannian manifold (M, g).[12] Thus, they can be taken to define the operators of gradient, divergence and, in three dimensions, of curl on M. In the case of three dimensions, these operators fit into the following commutative diagram:

$$\begin{array}{ccccccccccc}
0 & \longrightarrow & C^\infty(M) & \xrightarrow{\mathrm{d}} & \Omega^1(M) & \xrightarrow{\mathrm{d}} & \Omega^2(M) & \xrightarrow{\mathrm{d}} & \Omega^3(M) & \longrightarrow & 0 \\
 & & \uparrow{\scriptstyle =} & & \uparrow{\scriptstyle \mathrm{g}} & & \uparrow{\scriptstyle *\mathrm{g}} & & \uparrow{\scriptstyle *} & & \\
0 & \longrightarrow & C^\infty(M) & \xrightarrow{\mathrm{grad}} & \mathfrak{X}(M) & \xrightarrow{\mathrm{curl}} & \mathfrak{X}(M) & \xrightarrow{\mathrm{div}} & C^\infty(M) & \longrightarrow & 0.
\end{array}$$

This means that grad, curl and div establish a complex parallel to the de Rham complex. In particular, $\operatorname{curl} \circ \operatorname{grad} = 0$ and $\operatorname{div} \circ \operatorname{curl} = 0$.

Example 4.5.9 (Hodge-Laplace operator on Minkowski space) Let (M, η) be Minkowski space, that is, $M = \mathbb{R}^4$ with inner product

$$\eta_{ij} = \begin{pmatrix} +1 & 0 \\ 0 & -\mathbb{1}_3 \end{pmatrix}.$$

For $f \in C^\infty(M)$, we find

$$\big(\mathrm{d}\mathrm{d}^* + \mathrm{d}^*\mathrm{d}\big) f = \mathrm{d}^*\mathrm{d} f = (-1)^{4(1-1)+3+1} * \mathrm{d} * \mathrm{d} f,$$

that is,

$$\Box f = *\mathrm{d} * \mathrm{d} f. \tag{4.5.17}$$

In the standard coordinates ct, x_1, x_2, x_3 on M,

$$\Box f = \left(-\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \triangle\right) f, \qquad \triangle = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}.$$

That is, up to the sign,[13] $\Box$ coincides with the wave operator.

Exercises

4.5.1 Prove Remark 4.5.2.

4.5.2 Prove Formula (4.5.7) by showing that it holds for $\alpha = \vartheta^I$.

4.5.3 Prove Formula (4.5.11).

[12] A sign $(-1)^s$ has to be inserted on the right hand sides of Formulae (4.5.14) and (4.5.15).

[13] The proper sign would be obtained for the signature $(-\,+\,+\,+)$.

4.5.4 Consider $M=\mathbb{R}^n$ endowed with the Euclidean metric and let $X=X^i\partial_i \in \mathfrak{X}(\mathbb{R}^n)$. Prove the following formulae.
(a) $\mathrm{g}\circ X = X^1\mathrm{d}x^1+\cdots+X^n\mathrm{d}x^n$,
(b) $\operatorname{grad} f = \partial_1 f\partial_1+\cdots+\partial_n f\partial_n$,
(c) $\operatorname{div} X=\partial_i X^i$,
(d) $\operatorname{curl} X = (\partial_2X^3-\partial_3X^2)\partial_1+(\partial_3X^1-\partial_1X^3)\partial_2+(\partial_1X^2-\partial_2X^1)\partial_3$, $(n=3)$.

4.5.5 Consider $M=\mathbb{R}^n$, endowed with the Euclidean metric. Show that application of $\Box$ to a function $f\in C^\infty(\mathbb{R}^n)$ yields the Laplace operator, that is,

$$\Box f=\sum_{i=1}^n \frac{\partial^2 f}{\partial(x^i)^2}\equiv \triangle f.$$

Moreover, show that $\triangle f=\operatorname{div}\circ\operatorname{grad} f$.

4.5.6 Let U be a contractible open subset of $\mathbb{R}^3$. Prove the following.
(a) If $\operatorname{curl} X=0$ on U, there exists $f\in C^\infty(U)$ such that $X=\operatorname{grad} f$.
(b) If $\operatorname{div} X=0$ on U, there exists $Y\in\mathfrak{X}(\mathbb{R}^3)$ such that $X=\operatorname{curl} Y$.

4.5.7 Derive the integral theorems of classical vector analysis from Stokes' Theorem for differential forms.
(a) Theorem of Gauß: let $X\in\mathfrak{X}(\mathbb{R}^n)$. For any n-dimensional manifold $M\subset\mathbb{R}^n$ with boundary,

$$\int_M(\operatorname{div} X)\mathrm{v}_M=\int_{\partial M}(X\cdot\mathfrak{n})\mathrm{v}_{\partial M}.$$

Here, v_M denotes the volume form on M, $\mathrm{v}_{\partial M}$ the induced volume form on ∂M and $\mathfrak{n}$ is the (normalized) normal vector field on the boundary, pointing outwards.
(b) Classical Theorem of Stokes: let $X\in\mathfrak{X}(\mathbb{R}^3)$. For any 2-dimensional oriented surface,

$$\int_M\big((\operatorname{curl} X)\cdot\mathfrak{n}\big)\mathrm{v}_M=\int_{\partial M}(X\cdot\mathfrak{t})\mathrm{v}_{\partial M}.$$

Here, $\mathfrak{n}$ denotes the (normalized) normal vector field, compatible with the orientation of the surface, and $\mathfrak{t}$ denotes the (normalized) tangent vector field of the boundary, compatible with the induced orientation.

4.5.8 Let (M,η) be Minkowski space, cf. Example 4.5.9. The Levi-Civita tensor ε^{LC}_{ijkl}, written down in the standard basis, is defined as the signature of the permutation $(1234)\mapsto(ijkl)$. Show that ε^{LC}_{ijkl} fulfils

$$\varepsilon_{LC}^{ijkl}\varepsilon^{LC}_{mnpl}=-\delta^{ijk}_{mnp},\qquad \varepsilon_{LC}^{ijkl}\varepsilon^{LC}_{mnkl}=-2!\delta^{ij}_{mn},\qquad \varepsilon_{LC}^{ijkl}\varepsilon^{LC}_{mjkl}=-3!\delta^{i}_{m},$$

where $\delta^{i_1\ldots i_r}_{j_1\ldots j_r}=\operatorname{sign}\binom{i_1\ldots i_r}{j_1\ldots j_r}$ if $j_1,\ldots,j_r$ is a permutation of $i_1,\ldots,i_r$ and $\delta^{i_1\ldots i_r}_{j_1\ldots j_r}=0$ otherwise.

4.6 Maxwell's Equations

We consider the classical Maxwell theory of electrodynamics. This is a relativistic field theory on Minkowski space (M, η), with the metric η defined in Example 4.5.9. We use a reduced system of units by setting $c = 1$ and $\mu_0 = \varepsilon_0 = 1$. We limit our attention to the vacuum case, that is, we set the magnetic permeability μ and the relative permittivity ε equal to 1, but we allow for sources.

To start with, let us recall the standard relativistic formulation of Maxwell theory in a chosen Lorentz[14] system[15] $\{x^\mu\}$ in M. The electromagnetic field is described by the field strength tensor

$$f_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}, \tag{4.6.1}$$

where $\mathbf{E} = (E_x, E_y, E_z)$ and $\mathbf{B} = (B_x, B_y, B_z)$ are the electric field and the magnetic induction, respectively. In the above system of units, in the vacuum we have $\mathbf{D} = \mathbf{E}$ and $\mathbf{B} = \mathbf{H}$. Thus, the electrodynamic displacement tensor, built from $\mathbf{D}$ and $\mathbf{H}$, coincides with $f_{\mu\nu}$. The charge density ρ and the current density $\mathbf{j}$ constitute the 4-covector of the relativistic current density

$$j_\mu = (\rho, -\mathbf{j}). \tag{4.6.2}$$

In these notations, Maxwell's equations take the form

$$\partial_\mu f_{\nu\rho} + \partial_\nu f_{\rho\mu} + \partial_\rho f_{\mu\nu} = 0, \tag{4.6.3}$$

$$\partial_\mu f^{\mu\nu} = j^\nu. \tag{4.6.4}$$

Equations (4.6.3) are equivalent to the homogeneous Maxwell equations

$$\operatorname{curl} \mathbf{E} + \frac{\partial}{\partial t}\mathbf{B} = 0, \qquad \operatorname{div} \mathbf{B} = 0,$$

and Eqs. (4.6.4) correspond to

$$\operatorname{curl} \mathbf{B} - \frac{\partial}{\partial t}\mathbf{E} = \mathbf{j}, \qquad \operatorname{div} \mathbf{E} = \rho.$$

The transformation to another Lorentz frame $\{x'^\mu\}$ is given by $x'^\mu = \Lambda^\mu{}_\nu x^\nu$, with Λ denoting a proper orthochronous Lorentz transformation, that is,

$$\Lambda^T \eta \Lambda = \eta, \qquad \det \Lambda = 1, \qquad \Lambda^0{}_0 \geq 1. \tag{4.6.5}$$

[14]Named after the Dutch physicist Hendrik Lorentz (1853–1928).

[15]As usual in the physics literature we use Greek indices here.

The corresponding transformation laws for the field strength tensor and the current density are

$$f'^{\mu\nu} = \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\kappa} f^{\rho\kappa}, \qquad j'^{\mu} = \Lambda^{\mu}{}_{\nu} j^{\nu}.$$

The transformation law for $f_{\mu\nu}$ and j_{μ} is obtained by lowering the indices by $\eta_{\mu\nu}$.

Now, we pass to a coordinate-free description. For that purpose, we define the following differential forms on M:

$$f := \frac{1}{2} f_{\mu\nu} \mathrm{d}x^{\mu} \wedge \mathrm{d}x^{\nu}, \qquad j := j_{\mu}\mathrm{d}x^{\mu}. \tag{4.6.6}$$

In this notation, Maxwell's equations (4.6.3) and (4.6.4) take the form

$$\mathrm{d}f = 0, \tag{4.6.7}$$

$$\mathrm{d}^* f = -j, \tag{4.6.8}$$

with $\mathrm{d}^* = *\mathrm{d}*$ when acting on 2-forms (Exercise 4.6.1).

Remark 4.6.1

1. Applying the Hodge operator to (4.6.8) and using (4.5.7), we obtain $\mathrm{d} * f = - * j$. Application of the exterior derivative then yields $\mathrm{d} * j = 0$ or

$$\mathrm{d}^* j = 0. \tag{4.6.9}$$

   This is the continuity equation for the current density. It reflects the charge conservation law. In a Lorentz system, this equation reads $\partial_{\mu} j^{\mu} = 0$, that is, $\partial_t \rho + \mathrm{div}\,\mathbf{j} = 0$.
2. Since Minkowski space is contractible, the Poincaré Lemma applies: due to $\mathrm{d}f = 0$, there exists a differential 1-form A, called the electromagnetic potential, such that

$$f = \mathrm{d}A. \tag{4.6.10}$$

   In view of (4.6.6), in a Lorentz system this reads $f_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$. Since $\mathrm{d}^2 = 0$, the potential A is determined by f up to the differential of a smooth function, that is, for every $\lambda \in C^{\infty}(M)$, the 1-form

$$A' = A + \mathrm{d}\lambda \tag{4.6.11}$$

   yield the same field strength f as A. The mapping $A \mapsto A'$ is called a gauge transformation. To fix A, one may impose gauge conditions like the so-called Lorenz[16] gauge,

$$\mathrm{d}^* A = 0, \tag{4.6.12}$$

[16] Named after the Danish mathematician and physicist Ludvig Valentin Lorenz (1829–1891).

which in a Lorentz system reads $\partial_\mu A^\mu = 0$. This gauge fixing is not complete, that is, it does not fix A uniquely. If charges and currents are absent, in addition we may impose $A^0 = 0$, which then yields the so-called Coulomb gauge:[17]

$$\operatorname{div} \mathbf{A} = 0, \tag{4.6.13}$$

see Exercise 4.6.3.

3. Let us write down the Maxwell equations in terms of the potential in the Lorenz gauge. Equation (4.6.7) is trivially fulfilled and (4.6.8) implies

$$-j = \mathrm{d}^* f = \mathrm{d}^* \mathrm{d} A = \big(\mathrm{d}^* \mathrm{d} + \mathrm{d}\mathrm{d}^*\big) A = \square A.$$

Thus, we obtain a homogeneous wave equation,

$$\square A = -j. \tag{4.6.14}$$

4. One can check that $f \wedge *f$ and $f \wedge f$ are Lorentz-invariant 4-forms on M. In a Lorentz frame they read

$$f \wedge *f = \frac{1}{2} f_{\mu\nu} f^{\mu\nu} \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^3 = -\big(\mathbf{B}^2 - \mathbf{E}^2\big) \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^3, \tag{4.6.15}$$

$$f \wedge f = \frac{1}{2} f_{\mu\nu} (*f)^{\mu\nu} \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^3 = -(\mathbf{E} \cdot \mathbf{B}) \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^3 \tag{4.6.16}$$

(Exercise 4.6.4). Since $f \wedge f = \mathrm{d}(A \wedge \mathrm{d}A)$, the invariant $f \wedge f$ is a total divergence, hence its integral over M vanishes. Consequently, as a natural candidate for the Lagrange density L of the electromagnetic field there remains $f \wedge *f$. In the case where charges and currents are absent, we put

$$L = -\frac{1}{2} f \wedge *f. \tag{4.6.17}$$

Then, the physical action becomes

$$S = -\frac{1}{2} \int_M f \wedge *f = -\frac{1}{2} (f, f) \tag{4.6.18}$$

and Hamilton's variational principle yields the field equation[18] $\mathrm{d}^* f = 0$. Indeed, denoting the variation by δ, we compute

$$0 = -\delta S = (\delta f, f) = (\delta \mathrm{d} A, f) = (\mathrm{d} \delta A, f) = \big(\delta A, \mathrm{d}^* f\big). \tag{4.6.19}$$

Since this equation must hold for any δA, we conclude $\mathrm{d}^* f = 0$.

[17] More generally, the Coulomb gauge condition can also be imposed if charges and currents are present, but A^0 cannot be set equal to zero then. Instead, it is determined dynamically.

[18] The equation $\mathrm{d} f = 0$ is of geometric nature, it does not follow from a variational principle.

5. An electromagnetic field which fulfils

$$f \wedge f = 0, \qquad f \wedge * f = 0 \tag{4.6.20}$$

is called isotropic. Using the complex-valued 2-form $F := f - \mathrm{i} * f$, these two conditions can be summarized in $F \wedge F = 0$. By (4.6.15) and (4.6.16), they are equivalent to

$$\mathbf{E} \cdot \mathbf{B} = 0 \quad \text{and} \quad \mathbf{E}^2 - \mathbf{B}^2 = 0. \tag{4.6.21}$$

In particular, plane waves are isotropic. Since the conditions (4.6.20) are Lorentz-invariant, a nonzero isotropic electromagnetic field has the following properties.

– There does not exist a Lorentz system such that **E** and **B** are parallel.
– There does not exist a Lorentz system such that **E** or **B** vanish separately.

6. The invariant form (4.6.7) and (4.6.8) of the Maxwell equations immediately generalizes to an arbitrary oriented Lorentzian spacetime manifold, that is, with a pseudo-Riemannian metric g of signature $(+ - - -)$. Since it is only the Hodge star operator which is modified, in local coordinates, Eq. (4.6.7) is still given by (4.6.3). To analyse Eq. (4.6.8), we denote $|\mathrm{g}| \equiv |\det(\mathrm{g})|$ and use the identities

$$\mathrm{g}^{\alpha\beta}\mathrm{g}^{\rho\chi}\mathrm{g}^{\sigma\gamma}\mathrm{g}^{\iota\kappa}\varepsilon_{\beta\chi\gamma\kappa} = \big(\det(\mathrm{g})\big)^{-1}\varepsilon^{\alpha\rho\sigma\iota}, \tag{4.6.22}$$

$$\varepsilon_{\mu\nu\rho\sigma}\varepsilon^{\iota\alpha\rho\sigma} = 2\big(\delta^\iota_\mu\delta^\alpha_\nu - \delta^\alpha_\mu\delta^\iota_\nu\big), \tag{4.6.23}$$

together with $\det(\mathrm{g}) = (-1)^s|\det(\mathrm{g})|$, to calculate

$$\mathrm{d}^* f = *\mathrm{d}*\left(\frac{1}{2} f_{\mu\nu}\mathrm{d}\kappa^\mu \wedge \mathrm{d}\kappa^\nu\right) = -\frac{1}{\sqrt{|\mathrm{g}|}}\partial_\mu\big(f^{\mu\nu}\sqrt{|\mathrm{g}|}\big)\mathrm{g}_{\nu\lambda}\mathrm{d}\kappa^\lambda, \tag{4.6.24}$$

(Exercise 4.6.6). Thus, in local coordinates, (4.6.8) reads

$$\frac{1}{\sqrt{|\mathrm{g}|}}\partial_\mu\big(f^{\mu\nu}\sqrt{|\mathrm{g}|}\big) = j^\nu. \tag{4.6.25}$$

7. The gauge transformation given by (4.6.11) exhibits an additional symmetry of the Maxwell theory, which is obviously not related to a change of the Lorentz system. Thus, the question arises whether one can give a geometric interpretation to gauge transformations. Indeed, it turns out that A may be interpreted as the local representative of a connection form[19] ω in a principal $U(1)$-bundle P over spacetime. In this picture, the field strength f is the local representative of the curvature form of ω and Eq. (4.6.7) coincides with the Bianchi identity. A gauge transformation (4.6.11) may be interpreted either as a vertical automorphism of P (active transformation) or as a change of local bundle coordinates

[19] The mathematical tools needed for understanding the following interpretation will be presented in detail in part II of this book.

(passive transformation). In the latter interpretation, Eq. (4.6.11) describes the change of a reference frame.

Exercises

4.6.1 Prove that the Maxwell equations (4.6.3) and (4.6.4) are equivalent to (4.6.7) and (4.6.8).

4.6.2 Analyse Formula (4.6.9) in a Lorentz frame.

4.6.3 Prove that in the case where charges and currents are absent, the Coulomb gauge is complete, cf. Remark 4.6.1/2. In detail, show that
   (a) for every 2-form f, there exists a unique 1-form A satisfying $A^0 = 0$, (4.6.13) and (4.6.10),
   (b) for every 1-form A, there exists a smooth function λ such that $A' = A + \mathrm{d}\lambda$ satisfies $A'^0 = 0$ and (4.6.13).

4.6.4 Verify Eqs. (4.6.15) and (4.6.16).

4.6.5 Prove the statements of Remark 4.6.1/5.

4.6.6 Prove the identity (4.6.22) (the identity (4.6.23) was already proved in Exercise 4.5.8) and confirm Formula (4.6.24). Write down the continuity equation for an arbitrary Lorentzian spacetime manifold in local coordinates.

4.7 Pfaffian Systems and Differential Ideals

The notion of Pfaffian system is dual to the notion of distribution as developed in Sect. 3.5. It plays an important role in the theory of first order partial differential equations and in many physical applications, notably in thermodynamics and in mechanical systems with constraints. For an exhaustive presentation of this topic we refer to [64], Sect. IV.C.

For a subset $\varDelta \subset \mathrm{T}^*M$, let $\varOmega^1_\varDelta(M)$ denote the set of 1-forms on M with values in $\varDelta$.

Definition 4.7.1 (Pfaffian system) A Pfaffian system on M is a subset $\varDelta$ of T^*M such that for all $m \in M$ the following holds.

1. $\varDelta_m := \varDelta \cap \mathrm{T}^*_m M$ is a linear subspace.
2. For every $\beta \in \varDelta_m$, there exists $\alpha \in \varOmega^1_\varDelta(M)$ such that $\alpha_m = \beta$.

The function which assigns to $m \in M$ the dimension of $\varDelta_m$ is called the rank of $\varDelta$. If the rank is constant, $\varDelta$ is called regular. Otherwise, it is called singular.

As we did with distributions in Sect. 3.5, we follow the terminology used e.g. in [181]. Moreover, Remark 3.5.2 about distributions carries over to Pfaffian systems in an obvious way. In particular, the rank is locally non-decreasing.

Remark 4.7.2 A Pfaffian system Δ is regular of rank r iff it is an r-dimensional vertical vector subbundle of T^*M. In this case, it admits local frames built from local 1-forms on M. Let $\{\vartheta^1, \dots, \vartheta^r\}$ be such a local frame over $U \subset M$. If the local r-form $\vartheta = \vartheta^1 \wedge \cdots \wedge \vartheta^r$ is closed, the restriction of Δ to U coincides with the annihilator of the characteristic distribution of ϑ (Exercise 4.7.1).

Definition 4.7.3 (Integral manifold) Let Δ be a Pfaffian system on M. A connected submanifold (N, ψ) of M is called an integral manifold of Δ through $m \in M$ if $m \in \psi(N)$ and

$$\psi'(\mathrm{T}_p N) = \Delta^0_{\psi(p)} \tag{4.7.1}$$

for all $p \in N$, where $\Delta^0_{\psi(p)} \subset \mathrm{T}_{\psi(p)}M$ denotes the annihilator of $\Delta_{\psi(p)}$. The Pfaffian system Δ is said to be integrable if for every $m \in M$ there exists an integral manifold of Δ through m.

Remark 4.7.4 Assume that Δ is regular, that is, it is a vertical subbundle of T^*M. Then, its annihilator $\Delta^0 \subset \mathrm{T}M$ is a regular distribution and the ranks of Δ and Δ^0 add up to the dimension of M. By (4.7.1), a regular Pfaffian system Δ is integrable iff so is the distribution Δ^0. In this case, the integral manifolds of Δ and Δ^0 coincide. Similarly, if $D \subset \mathrm{T}M$ is a regular distribution, then the annihilator D^0 is a regular Pfaffian system.

Example 4.7.5

1. Every subset $A \subset \Omega^1(M)$ generates a Pfaffian system Δ, with Δ_m being defined as the linear span of the set $\{\alpha_m : \alpha \in A\}$, and every Pfaffian system can be generated this way.
2. The Pfaffian system on $\mathbb{R}^2$ spanned by the 1-forms $\mathrm{d}x$ and $x\mathrm{d}y$ is singular, because it has rank 1 on the y-axis and rank 2 outside. Similarly, the Pfaffian system spanned by $\mathrm{d}x$ and $y\mathrm{d}y$ is singular, because it has rank 1 on the x-axis and rank 2 outside. The first system is integrable, with integral manifolds being the y-axis and the single points outside. In contrast, the second system is not integrable.

Pfaffian systems naturally arise in the theory of partial differential equations. Consider the first order system

$$\frac{\partial y^j}{\partial x^i} = f_i^j\left(x^1, \dots, x^p, y^1, \dots, y^{n-p}\right), \quad 1 \le i \le p,\ 1 \le j \le n-p, \tag{4.7.2}$$

in the unknowns $y^j \in C^\infty(\mathbb{R}^p)$, with $f_i^j \in C^\infty(\mathbb{R}^n)$. Since the graph of the combined mapping $y = (y^1, \dots, y^{n-p}) : \mathbb{R}^p \to \mathbb{R}^{n-p}$ is a p-dimensional embedded submanifold in $\mathbb{R}^n$ and since y solves (4.7.2) iff

$$\mathrm{d}y^j - \sum_{i=1}^{p} f_i^j\bigl(x^1, \ldots, x^p, y^1, \ldots, y^{n-p}\bigr)\mathrm{d}x^i = 0, \quad 1 \le j \le n-p, \tag{4.7.3}$$

solutions of (4.7.2) can be interpreted geometrically as integral manifolds of the Pfaffian system Δ on $\mathbb{R}^n$, spanned by the 1-forms[20]

$$\vartheta^j := \mathrm{d}x^{p+j} - \sum_{i=1}^{p} f_i^j\bigl(x^1, \ldots, x^n\bigr)\mathrm{d}x^i, \quad 1 \le j \le n-p.$$

If the ϑ^j are linearly independent, Δ is regular. Then, by Remark 4.7.4, Δ is integrable iff so is the annihilator $\Delta^0 \equiv \ker \vartheta^1 \cap \cdots \cap \ker \vartheta^{n-p}$. In the special case $p = n-1$, we obtain the following simple solvability criterion.

Proposition 4.7.6 *A regular Pfaffian system Δ of rank 1, spanned by $\vartheta \in \Omega^1(M)$, is integrable iff $\mathrm{d}\vartheta(X_1, X_2) = 0$ for all $X_1, X_2 \in \mathfrak{X}^{\Delta^0}(M)$.*

Proof Proposition 4.1.6 implies

$$\mathrm{d}\vartheta(X_1, X_2) = -\vartheta\bigl([X_1, X_2]\bigr).$$

Since $\Delta^0 = \ker \vartheta$, the Frobenius Theorem 3.5.12 yields the assertion. □

Now, we are going to derive an integrability criterion for regular Pfaffian systems in the general case. For a Pfaffian system Δ, let $\Omega^*_\Delta(M)$ denote the subspace of $\Omega^*(M)$ spanned by the forms $\alpha \in \Omega^r(M)$, $r \ge 1$, which satisfy $\alpha_m(X_1, \ldots, X_r) = 0$ for all $X_1, \ldots, X_r \in \Delta^0_m$ and $m \in M$. $\Omega^*_\Delta(M)$ is a two-sided ideal in the associative algebra $\Omega^*(M)$.

Definition 4.7.7 (Differential ideal) An ideal $J \subset \Omega^*(M)$ is called a differential ideal if $\mathrm{d}J \subset J$.

Theorem 4.7.8 (Frobenius Theorem for Pfaffian systems) *A regular Pfaffian system Δ on M is integrable iff $\Omega^*_\Delta(M)$ is a differential ideal.*

Proof By Remark 4.7.4 and the Frobenius Theorem 3.5.12 for regular distributions, we must show that $\mathrm{d}\Omega^*_\Delta(M) \subset \Omega^*_\Delta(M)$ is a necessary and sufficient condition for the regular distribution Δ^0 to be involutive. That it is necessary follows from Proposition 4.1.6. To see that it is sufficient, let r denote the rank of Δ and let $m \in M$.

[20]It is common to refer to (4.7.2), rather than to Δ, as a Pfaffian system and to the 1-forms ϑ^j as Pfaffian forms.

Choose $\vartheta^1, \ldots, \vartheta^r \in \Omega^1_\Delta(M)$ such that $\{\vartheta^1_m, \ldots, \vartheta^r_m\}$ is a basis in Δ_m. According to Proposition 4.1.6, for $X, Y \in \mathfrak{X}^{\Delta^0}(M)$,

$$\big(\mathrm{d}\vartheta^i(X,Y)\big)(m) = -\vartheta^i_m\big([X,Y]_m\big), \quad 1 \le i \le r.$$

If $\mathrm{d}\Omega^*_\Delta(M) \subset \Omega^*_\Delta(M)$, the left hand side vanishes, hence $[X,Y]_m \in \Delta^0_m$. Since m was arbitrary, it follows that Δ^0 is involutive. □

Next, we derive local criteria for a regular Pfaffian system to be integrable. We need

Lemma 4.7.9 *Let Δ be a regular Pfaffian system of rank r on M, let $m \in M$ and let $\{\vartheta^1, \ldots, \vartheta^r\}$ be a local frame in Δ at m. Then, m admits an open neighbourhood U with the property that for every k-form α in $\Omega^*_\Delta(M)$ there exist local $(k-1)$-forms $\beta_1, \ldots, \beta_r$ over U such that*

$$\alpha_{\restriction U} = \sum_{j=1}^r \vartheta^j \wedge \beta_j.$$

Proof On some neighbourhood U of m, the local frame $\{\vartheta^1, \ldots, \vartheta^r\}$ in Δ can be complemented by local 1-forms $\vartheta^{r+1}, \ldots, \vartheta^n$ to a local frame in T^*M at m. Let $\{X_1, \ldots, X_n\}$ be the dual local frame in $\mathrm{T}M$. Then, $\{X_{r+1}, \ldots, X_n\}$ is a local frame in Δ^0. Expand

$$\alpha_{\restriction U} = \sum_{i_1 < \cdots < i_k} \alpha_{i_1 \ldots i_k} \vartheta^{i_1} \wedge \cdots \wedge \vartheta^{i_k}.$$

Since $\alpha \in \Omega^*_\Delta(M)$, we have that $\alpha_{i_1 \ldots i_k} = \alpha_{\restriction U}(X_{i_1}, \ldots, X_{i_k}) = 0$ whenever $r < i_1$. Hence, $\alpha_{\restriction U} = \sum_{j=1}^r \vartheta^j \wedge \beta_j$ with $\beta_j := \sum_{j < i_2 < \cdots < i_k} \alpha_{j i_2 \ldots i_k} \vartheta^{i_2} \wedge \cdots \wedge \vartheta^{i_k}$. □

Proposition 4.7.10 *Let Δ be a regular Pfaffian system of rank r on M. The following statements are equivalent.*

1. *Δ is integrable.*
2. *For every $m \in M$, there exist a local frame $\{\vartheta^1, \ldots, \vartheta^r\}$ in Δ at m and local 1-forms γ^i_j, $i, j = 1, \ldots, r$, such that for all $i = 1, \ldots, r$ one has $\mathrm{d}\vartheta^i = \sum_{j=1}^r \vartheta^j \wedge \gamma^i_j$.*
3. *For every $m \in M$, there exists a local frame $\{\vartheta^1, \ldots, \vartheta^r\}$ in Δ at m such that for all $i = 1, \ldots, r$ one has $\mathrm{d}\vartheta^i \wedge \vartheta^1 \wedge \cdots \wedge \vartheta^r = 0$.*

Proof 1 ⇒ 2: Choose a local frame $\{\vartheta^1, \ldots, \vartheta^r\}$ in Δ over a neighbourhood U of m. Since Δ is integrable, so is the Pfaffian system $\Delta_U := \Delta \cap (\mathrm{T}^*U)$ on U. Hence, Theorem 4.7.8 implies $\mathrm{d}\vartheta^i \in \Omega^*_{\Delta_U}(U)$. Application of Lemma 4.7.9 to the manifold U and the Pfaffian system Δ_U yields the assertion.

2 ⇒ 1: According to Theorem 4.7.8, it suffices to show that $\Omega^*_\Delta(M)$ is a differential ideal. Thus, let $\alpha \in \Omega^*_\Delta(M)$ be of degree k and let $m \in M$. Let $\{\vartheta^1, \ldots, \vartheta^r\}$

be the local frame in Δ over a neighbourhood U of m and let γ^i_j, $i, j = 1, \dots, r$, be the local 1-forms over U, existing by assumption. By Lemma 4.7.9, $\alpha_{\restriction U} = \sum_{i=1}^r \vartheta^i \wedge \beta_i$ for appropriate local $(k-1)$-forms β_i on U. Then,

$$\mathrm{d}(\alpha_{\restriction U}) = \sum_{i,j=1}^{r} \vartheta^j \wedge \gamma^i_j \wedge \beta_i - \sum_{i=1}^{r} \vartheta^i \wedge \mathrm{d}\beta_i$$

and hence $(\mathrm{d}\alpha)_m(X_1, \dots, X_k) = 0$ for all $X_i \in \Delta^0_m$. Since m was arbitrary, we get $\mathrm{d}\alpha \in \Omega^*_\Delta(M)$.

2 $\Rightarrow$ 3: This is obvious.

3 $\Rightarrow$ 2: We may complement the local frame $\{\vartheta^1, \dots, \vartheta^r\}$ in Δ by local 1-forms $\vartheta^{r+1}, \dots, \vartheta^n$ to a local frame in T^*M at m. Expanding $\mathrm{d}\vartheta^i = \sum_{j<k} \alpha^i_{jk} \vartheta^j \wedge \vartheta^k$, we obtain

$$0 = \sum_{j<k} \alpha^i_{jk} \vartheta^j \wedge \vartheta^k \wedge \vartheta^1 \wedge \cdots \wedge \vartheta^r.$$

It follows that $\alpha^i_{jk} = 0$ for all $r < j$. Hence, $\mathrm{d}\vartheta^i = \sum_{j=1}^r \vartheta^j \wedge \gamma^i_j$ with $\gamma^i_j = \sum_{j<k} \alpha^i_{jk} \vartheta^k$. □

Example 4.7.11 Let $M = \mathbb{R}^3$ and let Δ be the Pfaffian system spanned by a nowhere-vanishing 1-form ϑ on M. According to Proposition 4.7.10/3, Δ is integrable iff $\mathrm{d}\vartheta \wedge \vartheta = 0$. Writing $\vartheta = \vartheta_i \mathrm{d}x^i$, we have

$$\mathrm{d}\vartheta = (\partial_1\vartheta_2 - \partial_2\vartheta_1)\mathrm{d}x^1 \wedge \mathrm{d}x^2 + (\partial_2\vartheta_3 - \partial_3\vartheta_2)\mathrm{d}x^2 \wedge \mathrm{d}x^3 + (\partial_3\vartheta_1 - \partial_1\vartheta_3)\mathrm{d}x^3 \wedge \mathrm{d}x^1$$

and $\mathrm{d}\vartheta \wedge \vartheta = 0$ is equivalent to

$$(\partial_2\vartheta_3 - \partial_3\vartheta_2)\vartheta_1 + (\partial_3\vartheta_1 - \partial_1\vartheta_3)\vartheta_2 + (\partial_1\vartheta_2 - \partial_2\vartheta_1)\vartheta_3 = 0.$$

Thus, Δ is integrable iff the vector field $\sum_i \vartheta_i \partial_i$ is perpendicular to its curl.

To conclude this section, we derive an integrability criterion for regular Pfaffian systems of rank one in terms of so-called integrating factors. Since integrability is a local property, it is no loss of generality to assume the Pfaffian system to be spanned by a nowhere-vanishing 1-form.

Definition 4.7.12 Let $\vartheta \in \Omega^1(M)$. A nowhere-vanishing function $f \in C^\infty(M)$ is called an integrating factor for ϑ if the 1-form $f\vartheta$ is exact.

Proposition 4.7.13 *The Pfaffian system Δ spanned by a single nowhere-vanishing 1-form ϑ on M is integrable iff for every $m \in M$ there exists an open neighbourhood U and an integrating factor $f \in C^\infty(U)$ for $\vartheta_{\restriction U}$. For every $g \in C^\infty(U)$ satisfying $f\vartheta_{\restriction U} = \mathrm{d}g$, the level set components of g are integral manifolds of Δ.*

Proof First, assume that for every $m \in M$ there exists an open neighbourhood U and an integrating factor $f \in C^\infty(U)$ for $\vartheta_{\restriction U}$. Then,

$$0 = \mathrm{d}(f\vartheta_{\restriction U}) = \mathrm{d}f \wedge \vartheta_{\restriction U} + f\mathrm{d}\vartheta_{\restriction U},$$

hence $\mathrm{d}\vartheta_{\restriction U} = \vartheta_{\restriction U} \wedge \frac{\mathrm{d}f}{f}$, and Proposition 4.7.10/2 yields that Δ is integrable. Conversely, if Δ is integrable, so is the regular distribution Δ^0 of rank $n-1$ on M. Hence, according to Theorem 3.5.10/4, for every $m \in M$ there exists a local chart (U, κ) at m such that $\{\partial_1^\kappa, \ldots, \partial_{n-1}^\kappa\}$ is a local frame in Δ^0 at m. Then, $\vartheta = h\mathrm{d}\kappa^n$ for some $h \in C^\infty(U)$. Since ϑ is nowhere-vanishing, so is h. Thus, $f := h^{-1} \in C^\infty(U)$ is an integrating factor for $\vartheta_{\restriction U}$.

Next, let $g \in C^\infty(U)$ be such that $f\vartheta_{\restriction U} = \mathrm{d}g$ and let Σ be a level set component of g. Since $\mathrm{d}g = f\vartheta_{\restriction U}$ is nowhere-vanishing, every point of U is regular for g. Thus, Σ is an embedded submanifold of dimension $n-1$ of U (and hence of M). Let $\iota : \Sigma \to M$ denote the natural inclusion mapping. Due to

$$\iota^*\vartheta_{\restriction U} = \frac{1}{f_{\restriction\Sigma}}\iota^*(f\vartheta_{\restriction U}) = \frac{1}{f_{\restriction\Sigma}}\iota^*(\mathrm{d}g) = \frac{1}{f_{\restriction\Sigma}}\mathrm{d}(g_{\restriction\Sigma}) = 0,$$

one has $\mathrm{T}_m\Sigma \subset \Delta_m^0$ for all $m \in \Sigma$. By counting dimensions we obtain equality. Hence, Σ is an integral manifold of Δ. □

Example 4.7.14 (Ideal gas) The concept of integrating factors is important in thermodynamics. Let us consider one mol of an ideal gas, described by the variables V (volume), T (temperature) and p (pressure) which fulfil the thermal equation of state,

$$pV = RT. \tag{4.7.4}$$

Note that (4.7.4) defines a 2-dimensional surface $A \subset \mathbb{R}^3$, which for example can be parameterized by the variables (V, T) with $V > 0$ and $T > 0$. A change of state is represented by a curve in A and the heat exchange δQ with the environment during this process is obtained by integrating the 1-form

$$\vartheta = c_V\mathrm{d}T + p(V, T)\mathrm{d}V$$

along this curve. Thus, an adiabatic change of state, that is, a process for which there is no heat exchange with the environment, corresponds to an integral manifold of the Pfaffian system Δ on A defined by ϑ. By Proposition 4.7.10, Δ is integrable, because $\mathrm{d}\vartheta \wedge \vartheta = 0$ holds trivially on a 2-dimensional manifold. Then, by Proposition 4.7.13, there exist local integrating factors f for ϑ. To find them, it suffices to consider the condition $\mathrm{d}(f\vartheta) = 0$. This yields

$$0 = c_V\frac{\partial f}{\partial V}\mathrm{d}V \wedge \mathrm{d}T + \frac{\partial f}{\partial T}p\mathrm{d}T \wedge \mathrm{d}V + f\frac{\partial p}{\partial T}\mathrm{d}T \wedge \mathrm{d}V$$

and with (4.7.4) we conclude

$$c_V \frac{\partial f}{\partial V} - \frac{RT}{V} \frac{\partial f}{\partial T} = \frac{R}{V} f.$$

Any nowhere-vanishing function f which fulfils this differential equation is an integrating factor. Let us consider the following examples:

1. $f(T, V) = V^{R/c_V}$: a potential g of $f\vartheta$ can be obtained by applying the procedure described in Sect. 4.3:

$$g(T, V) = c_V \left(T V^{R/c_V} - T_0 V_0^{R/c_V} \right).$$

   The integral manifolds of the Pfaffian system Δ are given by the level set components of g in A,

$$T V^{R/c_V} = \text{const.}$$

2. $f(T, V) = T^{-1}$: the potential g coincides with the entropy

$$S(T, V) = c_V \ln \frac{T}{T_0} + R \ln \frac{V}{V_0},$$

   and the integral manifolds of Δ provided by this integrating factor correspond to reversible processes. For the ideal gas, with $\mathrm{d}U = c_V \mathrm{d}T$, we obtain

$$\mathrm{d}S = \frac{c_V}{T} \mathrm{d}T + \frac{R}{V} \mathrm{d}V = \frac{1}{T} \mathrm{d}U + \frac{p}{T} \mathrm{d}V,$$

   that is,

$$\mathrm{d}U = T \mathrm{d}S - p \mathrm{d}V.$$

   The integrability of the 1-form $\vartheta = c_V \, \mathrm{d}T + p \, \mathrm{d}V$ is equivalent to the second law of thermodynamics or, more precisely, it yields the mathematical foundation of this law for the ideal gas.

Exercises

4.7.1 Prove the statements of Remark 4.7.2.

4.8 Constraints in Classical Mechanics

The motion of a system of N particles in the configuration space $\mathbb{R}^{3N}$ is often restricted by constraints, that is, in addition to the external forces acting upon the system there are constraining forces, which ensure that the system evolves in accordance with the constraints. For simplicity, let us restrict our attention to time-independent constraints. Let $\mathbf{r}_\alpha$ denote the position of the α-th particle. There are two qualitatively different cases.

(a) The constraints are defined by a system of equations

$$f^a(\mathbf{r}_1, \dots, \mathbf{r}_N) = 0, \quad a = 1, \dots, s, \tag{4.8.1}$$

where $f^a : \mathbb{R}^{3N} \to \mathbb{R}$ are smooth functions which are assumed to be functionally independent. Constraints of this type are called holonomic. The Level Set Theorem 1.2.1 implies that

$$Q = \{(\mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathbb{R}^{3N} : f^a(\mathbf{r}_1, \dots, \mathbf{r}_N) = 0\} \tag{4.8.2}$$

is a $(3N - s)$-dimensional embedded submanifold of $\mathbb{R}^{3N}$, called the constraint manifold or the reduced configuration space of the system. We have

$$\bigcap_a \ker(\mathrm{d}f^a)_{(\mathbf{r}_1,\dots,\mathbf{r}_N)} = \ker f'_{(\mathbf{r}_1,\dots,\mathbf{r}_N)},$$

where $f = (f^1, \dots, f^s)$. Thus, the Pfaffian system generated by the exact 1-forms $\mathrm{d}f^a$ is integrable and the integral manifolds are the level set components of f.

(b) The constraints are defined by a system of equations

$$f^a(\mathbf{r}_1, \dots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \dots, \dot{\mathbf{r}}_N) = 0, \quad a = 1, \dots, s, \tag{4.8.3}$$

with f^a being smooth functions which depend both on positions and velocities. Constraints of this type are called nonholonomic. While, in general, the dependence on the velocities may be nonlinear, in practice one rather meets the case where this dependence is linear. This is referred to as the case of linear nonholonomic constraints. Let us confine ourselves to this situation. In general, there will be both holonomic and genuinely nonholonomic constraints. Let Q denote the constraint surface defined by the first ones and let $q^1, \dots, q^f$ be local coordinates on Q. The remaining constraints read $\mu^a_i \dot{q}^i = 0$, or

$$\mu^a = \mu^a_i \mathrm{d}q^i = 0, \quad a = 1, \dots, s. \tag{4.8.4}$$

The 1-forms μ^a generate a Pfaffian system Δ which is not integrable, because otherwise the constraints would be holonomic. The corresponding distribution Δ^0 is called the constraint distribution. Allowed trajectories $t \to q^i(t)$ have to fulfil $\dot{\mathbf{q}}(t) \in \Delta^0$ for all t or, equivalently,

$$\langle \mu^a(\mathbf{q}(t)), \dot{\mathbf{q}}(t) \rangle = 0 \tag{4.8.5}$$

for all t and a. We conclude that linear nonholonomic constraints are given by a non-integrable Pfaffian system. This type of constraints does not lead to a reduction of the configuration space.

Example 4.8.1 We consider a disk of radius R rolling in upright position on a horizontal plane $\bar{P}$ without gliding. Denote the Cartesian coordinates on $\bar{P}$ by (x, y).

At each time t, the disk is located in a plane $P(t)$ which intersects $\bar{P}$ along a line $l(t)$ and touches the plane $\bar{P}$ at a point $A(t)$. Let us denote the Cartesian coordinates of this point by (x_A, y_A) and let ϕ be the angle between the line $l(t)$ and the x-coordinate axis. Then, the position of the disk is completely described by the coordinates (x_A, y_A, ϕ, ψ), with ψ denoting the angle of rotation in the plane $P(t)$. Consequently, the configuration space of the system is $Q = \mathbb{R} \times \mathbb{R} \times S^1 \times S^1$. The non-gliding condition imposes the following constraints on the system: the infinitesimal displacement of the touching point is $R\mathrm{d}\psi$ and its projections onto the x- and y-axes are given by $\mathrm{d}x_A = R\cos(\phi)\mathrm{d}\psi$ and $\mathrm{d}y_A = R\sin(\phi)\mathrm{d}\psi$, respectively. Thus, we find the velocity-dependent constraints

$$\dot{x}_A = R\cos(\phi)\dot{\psi}, \qquad \dot{y}_A = R\sin(\phi)\dot{\psi}. \tag{4.8.6}$$

They correspond to the Pfaffian forms

$$\mu^1 = \mathrm{d}x_A - R\cos(\phi)\mathrm{d}\psi, \qquad \mu^2 = \mathrm{d}y_A - R\sin(\phi)\mathrm{d}\psi.$$

According to point 3 of Proposition 4.7.10, this system is not integrable. Thus, we deal with a genuinely nonholonomic system. This fact is in accordance with our intuition, which tells us that at any touching point $A(t)$ the intersection line $l(t)$ may have an arbitrary direction. For a detailed discussion of further examples we refer to Benenti [39].

Now, let us come back to a system of N particles with masses m_α and position vectors $\mathbf{r}_\alpha$ which are acted upon by external forces $\mathbf{F}_\alpha$. Assume that this system is subject to linear nonholonomic constraints defined by 1-forms μ^a and let Δ be the Pfaffian system generated by these forms. Let us denote the corresponding constraining forces by $\mathbf{Z}_\alpha$. Newton's second law reads

$$m_\alpha \ddot{\mathbf{r}}_\alpha = \mathbf{F}_\alpha + \mathbf{Z}_\alpha, \quad \alpha = 1, \dots, N. \tag{4.8.7}$$

There is a fundamental principle, called d'Alembert's principle, which in geometric terms[21] states that the constraining forces $\mathbf{Z}_\alpha$, viewed as 1-forms,[22] belong to the annihilator $(\Delta^0)^0 = \Delta$ of the constraint distribution Δ^0. Thus, we have

$$Z_\alpha = \sum_a \lambda_a \mu^a_\alpha$$

and the system of Eqs. (4.8.7) takes the form of the so-called Lagrange equations of the first kind:

$$m_\alpha \ddot{\mathbf{r}}_\alpha = \mathbf{F}_\alpha + \sum_a \lambda_a \boldsymbol{\mu}^a_\alpha, \tag{4.8.8}$$

[21] In mechanics, this principle is usually spelled out by saying that the constraining forces do no virtual work.

[22] We identify vectors and covectors on $\mathbb{R}^{3N}$ via the Euclidean metric.

with the Lagrange multipliers λ_a to be determined from Eqs. (4.8.5). It sometimes happens that the external forces are given by a single potential function V which depends on both positions and velocities:

$$\mathbf{F}_\alpha = -\frac{\partial V}{\partial \mathbf{r}_\alpha} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V}{\partial \dot{\mathbf{r}}_\alpha}\right).$$

In this case, rewriting

$$m_\alpha \ddot{\mathbf{r}}_\alpha = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{\mathbf{r}}_\alpha}\right)$$

with $T(\dot{\mathbf{r}}_\alpha) = \frac{1}{2}\sum_\alpha m_\alpha \dot{\mathbf{r}}_\alpha^2$ denoting the kinetic energy of the system and defining the Lagrangian function

$$L\colon \mathrm{T}\mathbb{R}^{3N} \to \mathbb{R}, \qquad L(\mathbf{r}_\alpha, \dot{\mathbf{r}}_\alpha) := T(\dot{\mathbf{r}}_\alpha) - V(\mathbf{r}_\alpha, \dot{\mathbf{r}}_\alpha), \tag{4.8.9}$$

we arrive at the so-called Lagrange equations of the second kind:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\mathbf{r}}_\alpha}\right) - \frac{\partial L}{\partial \mathbf{r}_\alpha} = \sum_a \lambda_a \boldsymbol{\mu}_\alpha^a. \tag{4.8.10}$$

As already noted, in the case of holonomic constraints, the dynamics reduces to the constraint submanifold Q given by (4.8.2). Let us choose local coordinates q^i with $i = 1, \dots, 3N - s$ on Q. If we assume that the so-called generalized forces

$$F_i = \sum_\alpha \mathbf{F}_\alpha \cdot \frac{\partial \mathbf{x}_\alpha}{\partial q^i}$$

possess a potential[23] V, the Lagrangian equations of the second kind take the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0, \tag{4.8.11}$$

where again $L = T - V$ is the Lagrangian function of the system. One can cast the Lagrangian equations (4.8.10) and (4.8.11) into a coordinate-free form by means of the symplectic structure on $\mathrm{T}Q$ induced from the natural symplectic structure on T^*Q, see Sect. 9.1.

Finally, we stress that, nowadays, nonholonomic constraints constitute a huge field of research in mathematical physics. There are at least two important streams to be mentioned:

(a) nonholonomic systems with symmetries, with the aim of understanding robotics, trajectory tracking, dynamic stability, feedback stabilization and control. For this line of research we refer to Koon and Marsden [170], where the reader will

[23] E.g. if the original external forces possess a potential.

find a lot of further references. We also recommend the papers of Bates, Cushman, Kemppainen and Śniatycki, see [35] and [70], and a paper of Marle [193] for an introduction to the subject,

(b) efforts to develop a general theory (which for the time being is still lacking) with mathematical techniques like jets and algebroids. For this line of research we refer to the paper [147] by Iglesias, Marrero, De Diego and Sosa, where the reader can find a lot of further references. In this context, we also mention the classical paper [298] by Vershik and Gershkovich.

Chapter 5
Lie Groups

In this chapter, we give an introduction to the theory of Lie groups. In Sect. 5.1, we discuss the basic notions and provide the reader with a number of examples. In particular, we take up the classical groups which have already been introduced in Sect. 1.2. Next, in Sect. 5.2, we come to the notion of Lie algebra of a Lie group. We consider a number of examples, again with some emphasis on the Lie algebras of classical Lie groups. Section 5.3 is devoted to an important tool, the exponential mapping. This mapping constitutes a link between the Lie group and its Lie algebra which turns out to be useful both for the study of the local structure of Lie groups and for the study of their representations. Next, in Sect. 5.4, we discuss a number of important representations—the adjoint and the coadjoint representations of the Lie group and the corresponding derived representations of its Lie algebra. Using the adjoint representation of the Lie algebra, one can construct a natural symmetric bilinear form on the Lie algebra, the Killing form, which is invariant under the adjoint representation of the group. Next, in Sect. 5.5, we discuss the concept of left-invariant forms which in particular yields a unique (up to multiplication by a number) left-invariant volume form on every Lie group. The latter gives rise to the so-called Haar measure. We discuss the relation with Ad-invariant scalar products on the Lie algebra and conclude, in particular, that every compact Lie group admits a bi-invariant Riemannian metric. The final two sections are devoted to the theory of Lie subgroups and to homogeneous spaces. Concerning the latter, we discuss three important examples in detail: Stiefel manifolds, Graßmann manifolds and flag manifolds.

5.1 Basic Notions and Examples

The notion of Lie group arises naturally by combining the algebraic structure of a group with the differentiable structure of a smooth manifold and requiring that the two structures are compatible.

For a group G and elements $a, b \in G$, we denote the product by ab, the unit element by $\mathbb{1}$ (or sometimes also be e) and the inverse of a by a^{-1}. The assignment

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*,
Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7_5, © Springer Science+Business Media Dordrecht 2013

$a \mapsto a^{-1}$ defines a bijective mapping inv : $G \to G$, called the inversion mapping of G.

Definition 5.1.1 (Lie group) A Lie group is a set which carries the algebraic structure of a group and the differentiable structure of a smooth manifold such that the mapping

$$G \times G \to G, \qquad (a, b) \mapsto ab^{-1} \tag{5.1.1}$$

is smooth. A homomorphism of Lie groups is a mapping which is both a group homomorphism and smooth.

Remark 5.1.2

1. The inversion mapping $g \to g^{-1}$ is smooth, because it is the restriction of the mapping (5.1.1) to the submanifold $\{\mathbb{1}\} \times G$. Moreover, the multiplication mapping is smooth, because it can be written as the composition of (5.1.1) with the inversion mapping. Using the Inverse Mapping Theorem one can show that in Definition 5.1.1, instead of (5.1.1), it suffices to require that the multiplication mapping be smooth (Exercise 5.1.1).
2. Notions associated with manifolds, like dimension, compactness, connectedness, tangent and cotangent bundles etc., carry over to Lie groups in an obvious way. To some extent, this applies to notions associated with groups, too. For example, a Lie group is Abelian if so is the underlying group, or the order of an element is the order of this element of the underlying group. There are exceptions, however, see Example 5.2.20.
3. Analogously to Definition 5.1.1 one defines the notion of real and complex analytic Lie group by requiring the multiplication mapping to be real or complex analytic, respectively. It turns out that every Lie group is in fact real analytic. That is, the differentiable structure of an arbitrary Lie group G contains a unique real analytic structure [243]. For this reason, it is common in the literature to define Lie groups to be real analytic manifolds.

Example 5.1.3 (Lie groups)

1. Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$. Every $\mathbb{K}$-vector space V is a Lie group with respect to its additive structure and the manifold structure of the real vector space obtained from V by field restriction. The dimension is $\dim_{\mathbb{R}}(\mathbb{K}) \cdot \dim(V)$. Lie groups of this form are sometimes called vector Lie groups.
2. Every finite or countable group, equipped with the discrete topology and the trivial zero-dimensional manifold structure, is a Lie group. In particular, the cyclic groups of finite order and the group of integers are Lie groups.
3. Let G_1, G_2 be Lie groups. The direct product $G_1 \times G_2$ carries the group structure of the direct product of the underlying groups, given by the multiplication mapping $(a_1, a_2)(b_1, b_2) = (a_1 b_1, a_2 b_2)$, and the differentiable structure of the direct product of the underlying manifolds. These two structures combine to a Lie group structure (Exercise 5.1.2).

Example 5.1.4 (Classical Lie groups) Let A be a finite-dimensional associative algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. View A as a real algebra by restricting scalar multiplication to $\mathbb{R}$. Let $G \subset A$ be an initial submanifold which forms a group under the multiplication induced from A. Then, the multiplication mapping of G is the restriction in domain and range of the multiplication in A. Since the latter is bilinear, it is smooth. Hence, G is a Lie group. This argument implies that the following groups are Lie groups.

1. $\mathbb{K} \setminus \{0\}$ for $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{H}$, with multiplication induced from $\mathbb{K}$.
2. The unit spheres in $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{H}$ with respect to the natural Hermitian inner product defined by conjugation and multiplication. They are groups under the multiplication induced from $\mathbb{K}$ and coincide with the classical groups O(1) for $\mathbb{K} = \mathbb{R}$, U(1) for $\mathbb{K} = \mathbb{C}$ and Sp(1) for $\mathbb{K} = \mathbb{H}$, cf. Example 1.2.6. Hence, as manifolds, these classical groups are diffeomorphic to, respectively, S^0, S^1 and S^3.[1]
3. The general linear group $\mathrm{GL}(n, \mathbb{K})$, because according to Example 1.1.14 it is open in $\mathrm{M}_n(\mathbb{K})$. The dimension of $\mathrm{GL}(n, \mathbb{K})$ is given by the dimension of $\mathrm{M}_n(\mathbb{K})$ over the reals, that is, $\dim_{\mathbb{R}}(\mathbb{K}) \cdot n^2$. For $n = 1$, this reproduces point 1 above. By the same argument, the automorphism group $\mathrm{GL}(V)$ of a $\mathbb{K}$-vector space V with composition of mappings as multiplication is open in the associative algebra $\mathrm{End}(V)$ and hence a Lie group as well.
4. The classical groups of Example 1.2.6, because each of them can be represented as a level set of a smooth function on some $\mathrm{M}_n(\mathbb{K})$ at a regular value and is hence an embedded submanifold of the latter.

Example 5.1.5 (Lie group homomorphisms)

1. Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. The determinant $\det : \mathrm{M}_n(\mathbb{K}) \to \mathbb{K}$ restricts to a group homomorphism
$$\det : \mathrm{GL}(n, \mathbb{K}) \to \mathbb{K} \setminus \{0\}. \tag{5.1.2}$$
The latter is obtained by restriction of an n-linear mapping
$$\mathrm{M}_n(\mathbb{K}) \times \overset{n}{\cdots} \times \mathrm{M}_n(\mathbb{K}) \to \mathbb{K}$$
in domain to the submanifold $(\mathrm{GL}(n, \mathbb{K}), \Delta_n)$, where $\Delta_n(a) = (a, \dots, a)$, and in range to the open submanifold $\mathbb{K} \setminus \{0\}$. Hence, it is smooth and thus a Lie group homomorphism.
2. Let V, W be vector spaces over $\mathbb{K}$ and let $\varphi : V \to W$ be an isomorphism. The mapping $\mathrm{GL}(V) \to \mathrm{GL}(W)$, defined by $a \mapsto \varphi \circ a \circ \varphi^{-1}$, is a Lie group isomorphism. In particular, if $\dim V = n$, $W = \mathbb{K}^n$ and φ assigns to $v \in V$ its coefficients with respect to a chosen basis, the above mapping yields an isomorphism

[1] For the unit spheres of $\mathbb{R}$ and $\mathbb{C}$ there are simpler arguments ensuring smoothness of the multiplication mapping. In the first case, the group is finite and in the second case, in terms of the angle coordinate, multiplication is given by addition.

from $\mathrm{GL}(V)$ onto $\mathrm{GL}(n, \mathbb{K})$, and this isomorphism assigns to an automorphism of V its matrix with respect to this basis.

3. A Lie group homomorphism $\varphi : \mathbb{R} \to G$ is called a one-parameter subgroup. The image of a one-parameter subgroup is an Abelian subgroup of G. It turns out that all one-parameter subgroups are integral curves of certain vector fields, see Remark 5.3.2.
4. By definition of the group and manifold structures on the direct product of Lie groups G_1, G_2, the natural projections $G_1 \times G_2 \to G_i$, $i = 1, 2$ are group homomorphisms and smooth. Hence, they are Lie group homomorphisms.

We continue with deriving some basic facts about Lie groups. For that purpose, note that every $a \in G$ induces the following mappings of G:

$$\mathrm{L}_a(b) := ab, \qquad \mathrm{R}_a(b) := ba, \qquad \mathrm{C}_a(b) := \mathrm{L}_a \circ \mathrm{R}_{a^{-1}}(b) \equiv aba^{-1}. \tag{5.1.3}$$

They are called, respectively, left translation, right translation[2] and conjugation by a. For all $a, b \in G$ we have

$$\mathrm{L}_a \circ \mathrm{L}_b = \mathrm{L}_{ab}, \qquad \mathrm{R}_a \circ \mathrm{R}_b = \mathrm{R}_{ba}, \qquad \mathrm{L}_a \circ \mathrm{R}_b = \mathrm{R}_b \circ \mathrm{L}_a. \tag{5.1.4}$$

Thus, L_a and R_a are diffeomorphisms with inverse $\mathrm{L}_{a^{-1}}$ and $\mathrm{R}_{a^{-1}}$, respectively. For all $a, b, c \in G$, we have $\mathrm{C}_a(bc) = \mathrm{C}_a(b)\mathrm{C}_a(c)$ and $\mathrm{C}_a \circ \mathrm{C}_b = \mathrm{C}_{ab}$, that is, C_a is an automorphism of G for every a, and the assignment $a \mapsto \mathrm{C}_a$ defines a group homomorphism from G to the group of automorphisms of G. Automorphisms of G of the form C_a are called inner automorphisms.

Proposition 5.1.6 (Parallelizability) *Let G be a Lie group. The mappings*

$$\chi_{\mathrm{L}}, \chi_{\mathrm{R}} : G \times \mathrm{T}_{\mathbb{1}}G \to \mathrm{T}G, \qquad \chi_{\mathrm{L}}(a, X) := \mathrm{L}'_a X, \qquad \chi_{\mathrm{R}}(a, X) := \mathrm{R}'_a X, \tag{5.1.5}$$

are vertical vector bundle isomorphisms and hence they define global trivializations of $\mathrm{T}G$. *In particular, Lie groups are parallelizable.*

Proof Let $\mu : G \times G \to G$ denote the multiplication mapping of G, $s_0 : G \to \mathrm{T}G$ the zero section and $j : \mathrm{T}_{\mathbb{1}}G \to \mathrm{T}G$ the natural inclusion mapping. The mapping χ_{L} is obtained by composing $s_0 \times j : G \times \mathrm{T}_{\mathbb{1}}G \to \mathrm{T}G \times \mathrm{T}G$ with the natural isomorphism $\mathrm{T}G \times \mathrm{T}G \to \mathrm{T}(G \times G)$ and the tangent mapping of μ. Hence, χ_{L} is smooth. Since it preserves the fibres, and since it is fibrewise linear and projects to id_G, it is a vector bundle morphism. Since χ_{L} is obviously bijective, it is a vertical vector bundle isomorphism. The argument for χ_{R} is completely analogous. □

According to Example 2.4.1, the vertical vector bundle isomorphisms χ_{L} and χ_{R} yield dual isomorphisms $\chi_{\mathrm{L}}^{\mathrm{T}}$, $\chi_{\mathrm{R}}^{\mathrm{T}} : \mathrm{T}^*G \to G \times \mathrm{T}_{\mathbb{1}}^*G$. A brief calculation

[2] Or left and right multiplication, respectively.

(Exercise 5.1.3) shows that the inverse isomorphisms $(\chi_{\mathrm{L}}^{\mathrm{T}})^{-1}, (\chi_{\mathrm{R}}^{\mathrm{T}})^{-1} : G \times \mathrm{T}_{\mathbb{1}}^{*} G \to \mathrm{T}^{*} G$ are given by

$$\left(\chi_{\mathrm{L}}^{\mathrm{T}}\right)^{-1}(a, \eta) = \eta \circ (\mathrm{L}_{a^{-1}})'_a, \qquad \left(\chi_{\mathrm{R}}^{\mathrm{T}}\right)^{-1}(a, \eta) = \eta \circ (\mathrm{R}_{a^{-1}})'_a. \tag{5.1.6}$$

The isomorphisms χ_{L}^{-1}, χ_{R}^{-1} and $\chi_{\mathrm{L}}^{\mathrm{T}}$, $\chi_{\mathrm{R}}^{\mathrm{T}}$ will be referred to as, respectively, the left and the right trivialization of $\mathrm{T}G$ and $\mathrm{T}^{*}G$.

Proposition 5.1.7 *A connected Lie group is generated, as a group, by any neighbourhood of the unit element.*

Proof Let U be a neighbourhood of $\mathbb{1}$ in G. The group generated by U is $\tilde{G} := \bigcup_{n=1}^{\infty} U^n$ and we have to show that $\tilde{G} = G$. Without loss of generality, we may assume that U is open. Then, $\tilde{G}$ is open, hence $G \setminus \tilde{G}$ is closed. On the other hand, since L_a is a homeomorphism, $a\tilde{G}$ is open for all $a \in G$, hence

$$G \setminus \tilde{G} = \bigcup_{a \notin \tilde{G}} a\tilde{G}$$

is open. Since $G \setminus \tilde{G} \neq G$ and since G is connected, this implies $G \setminus \tilde{G} = \varnothing$. This yields the assertion. □

Remark 5.1.8 (Identity component) The connected component of G containing $\mathbb{1}$ will be denoted by G_0 and will be referred to as the identity component of G. Let $a \in G_0$. Since $\mathrm{L}_{a^{-1}}$ is a homeomorphism of G, $\mathrm{L}_{a^{-1}}(G_0)$ is a connected component. Since it contains $\mathbb{1}$, it coincides with G_0. It follows that $a^{-1}b \in G_0$ for all $a, b \in G_0$. Hence, G_0 is a subgroup. By a similar argument, one can show that G_0 is normal. Since it is an open submanifold of G, it is a normal Lie group.

Remark 5.1.9 Let G, H be Lie groups and let $\varphi : G \to H$ be a group homomorphism. Then, for all $a \in G$,

$$\varphi = \mathrm{L}_{\varphi(a)} \circ \varphi \circ \mathrm{L}_{a^{-1}}. \tag{5.1.7}$$

Since $\mathrm{L}_{a^{-1}}$ and $\mathrm{L}_{\varphi(a)}$ are diffeomorphisms of G and H, respectively, we obtain the following.

1. For a group homomorphism of Lie groups to be a Lie group homomorphism it suffices to be smooth in some neighbourhood of $\mathbb{1}$. Indeed, if U is such a neighbourhood, then (5.1.7) implies $\varphi_{\restriction aU} = \mathrm{L}_{\varphi(a)} \circ \varphi_{\restriction U} \circ (\mathrm{L}_{a^{-1}})_{\restriction aU}$.
2. For a Lie group homomorphism to be an immersion or a submersion it suffices to be an immersion or a submersion at $\mathbb{1}$. Indeed, differentiation of (5.1.7) at $a \in G$ yields $\varphi'_a = (\mathrm{L}_{\varphi(a)})'_{\mathbb{1}_H} \circ \varphi'_{\mathbb{1}_G} \circ (\mathrm{L}_{a^{-1}})'_a$ and the statement follows, because $(\mathrm{L}_{a^{-1}})'_a$ and $(\mathrm{L}_{\varphi(a)})'_{\mathbb{1}_H}$ are bijective.

To conclude this section, we discuss a couple of specific homomorphisms of some classical Lie groups which are of particular relevance for physics.

Example 5.1.10 (Isomorphism of SU(2) and Sp(1)) From Example 1.2.6, we recall that

$$\mathrm{SU}(2) = \left\{a \in \mathrm{GL}(2, \mathbb{C}) : a^\dagger = a^{-1}, \det(a) = 1\right\},$$
$$\mathrm{Sp}(1) = \left\{\mathbf{a} \in \mathbb{H} : \|\mathbf{a}\| = 1\right\}.$$

According to Remark 1.1.13, the mapping $\mathbb{H} \to \mathrm{M}_2(\mathbb{C})$, given by

$$\mathbf{a} = a_0\mathbf{1} + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} \mapsto \begin{bmatrix} a_0 + \mathrm{i}a_1 & a_2 + \mathrm{i}a_3 \\ -a_2 + \mathrm{i}a_3 & a_0 - \mathrm{i}a_1 \end{bmatrix}, \tag{5.1.8}$$

where $a_0, \ldots, a_3$ are real numbers, is an injective homomorphism of real algebras from $\mathbb{H}$ to $\mathrm{M}_2(\mathbb{C})$. One can check that the image of Sp(1) under this homomorphism coincides with SU(2) (Exercise 5.1.4). Hence, by restriction, (5.1.8) induces a group isomorphism from Sp(1) onto SU(2) which by Proposition 1.6.10 is smooth in both directions and hence a Lie group isomorphism. Since Sp(1) coincides with the unit sphere in $\mathbb{H}$, this shows in particular that SU(2) is diffeomorphic to the sphere S^3.

Example 5.1.11 (Universal coverings of SO(3) and SO(4)) Consider the isometric isomorphism of real vector spaces

$$\lambda : \mathbb{R}^4 \to \mathbb{H}, \qquad \lambda(\mathbf{x}) := x_0\mathbf{1} + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}. \tag{5.1.9}$$

Every pair of quaternions $(\mathbf{a}, \mathbf{b})$ defines a real linear mapping $\mathbb{H} \to \mathbb{H}$ by $\mathbf{q} \mapsto \mathbf{a}\mathbf{q}\overline{\mathbf{b}}$ and hence a linear mapping $\phi(\mathbf{a}, \mathbf{b})$ of $\mathbb{R}^4$ by

$$\phi(\mathbf{a}, \mathbf{b})\mathbf{x} := \lambda^{-1}\left(\mathbf{a}\lambda(\mathbf{x})\overline{\mathbf{b}}\right) \quad \text{for all } \mathbf{x} \in \mathbb{R}^4. \tag{5.1.10}$$

By assigning $\phi(\mathbf{a}, \mathbf{b})$ to $(\mathbf{a}, \mathbf{b})$ one obtains a mapping $\phi : \mathbb{H} \oplus \mathbb{H} \to \mathrm{M}_4(\mathbb{R})$, where $\mathbb{H} \oplus \mathbb{H}$ denotes the direct product of algebras with multiplication defined as usual by $(\mathbf{a}_1, \mathbf{b}_1)(\mathbf{a}_2, \mathbf{b}_2) := (\mathbf{a}_1\mathbf{a}_2, \mathbf{b}_1\mathbf{b}_2)$. The mapping ϕ is a homomorphism of real algebras and, hence, it is in particular smooth.

First, consider $\phi(\mathbf{a}, \mathbf{a})$ for $\mathbf{a} \in \mathrm{Sp}(1)$. Due to $\|\mathbf{a}\|^2 = 1$, $\phi(\mathbf{a}, \mathbf{a})$ is orthogonal. Since it leaves invariant the subspace $\mathbb{R} \times \{0\} \subset \mathbb{R}^4$, it also leaves invariant the subspace $\{0\} \times \mathbb{R}^3 \subset \mathbb{R}^4$. Hence, restriction of ϕ in domain to the submanifold $\{(\mathbf{a}, \mathbf{a}) : \mathbf{a} \in \mathrm{Sp}(1)\}$ of $\mathbb{H} \oplus \mathbb{H}$ and in range to the embedded submanifold of $\mathrm{M}_4(\mathbb{R})$ consisting of the matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}, \quad c \in \mathrm{O}(3),$$

yields a Lie group homomorphism (denoted by the same symbol) $\phi : \mathrm{Sp}(1) \to \mathrm{O}(3)$. The defining equation (5.1.10) reduces to

$$\phi(\mathbf{a})\mathbf{x} := \lambda^{-1}\left(\mathbf{a}(x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k})\overline{\mathbf{a}}\right) \quad \text{for all } \mathbf{x} \in \mathbb{R}^3. \tag{5.1.11}$$

For $\mathbf{a} = a_0\mathbf{1} + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, an explicit calculation yields

$$\left[\phi(\mathbf{a})\right]_{ij} = \left(a_0^2 - a_1^2 - a_2^2 - a_3^2\right)\delta_{ij} + 2(a_i a_j - a_0\varepsilon_{ijk}a_k). \tag{5.1.12}$$

We determine the kernel and the image of ϕ. According to (5.1.11), $\mathbf{a} \in \ker(\phi)$ iff $\mathbf{a}$ commutes with all quaternions. Then $\mathbf{a} = a\mathbf{1}$, $a \in \mathbb{R}$, hence $\ker(\phi) = \{\pm\mathbf{1}\}$, the centre of Sp(1). To find $\mathrm{im}(\phi)$, one first shows that ϕ is an immersion (Exercise 5.1.5). Then, since Sp(1) and SO(3) have the same dimension, ϕ is a submersion and hence, by Remark 1.5.16, an open mapping. See Exercise 5.1.6 for an alternative argument proving openness. It follows that $\mathrm{im}(\phi)$ contains a neighbourhood of $\mathbb{1}$ in O(3). Since $\mathrm{im}(\phi)$ is a subgroup, Proposition 5.1.7 implies that it contains the identity component SO(3) of O(3). Since Sp(1) is connected, we finally obtain $\mathrm{im}(\phi) = \mathrm{SO}(3)$.

Note that since Sp(1) coincides with the unit sphere in $\mathbb{H}$ and since $\ker(\phi) = \{\pm\mathbf{1}\}$ implies that the preimage of an element of SO(3) under ϕ consists of antipodal points, ϕ induces a bijection from the projective space $\mathbb{R}\mathrm{P}^3$ onto SO(3). Since both ϕ and the projection $\mathrm{S}^3 \to \mathbb{R}\mathrm{P}^3$ are submersions, Remark 1.5.16 yields that this bijection is in fact a diffeomorphism.[3]

Next, consider $\phi(\mathbf{a}, \mathbf{b})$ for $\mathbf{a}, \mathbf{b} \in \mathrm{Sp}(1)$. Due to $\|\mathbf{a}\|^2 = \|\mathbf{b}\|^2 = 1$, $\phi(\mathbf{a}, \mathbf{b})$ is orthogonal. Hence, by restriction in domain and range, ϕ induces a Lie group homomorphism (again denoted by the same symbol) $\phi : \mathrm{Sp}(1) \times \mathrm{Sp}(1) \to \mathrm{O}(4)$ with defining equation (5.1.10). Arguing as before one finds $\ker(\phi) = \{(\mathbf{1}, \mathbf{1}), (-\mathbf{1}, -\mathbf{1})\}$ and $\mathrm{im}(\phi) = \mathrm{SO}(4)$ (Exercise 5.1.7).

Let us add that via the algebra homomorphism (5.1.8), all of the above has an equivalent formulation in terms of two-dimensional complex matrices. This formulation is obtained by replacing Sp(1) by SU(2), $\mathbb{H}$ by the real vector space spanned by $\mathbb{1}_2$ and the traceless skew-Hermitian matrices, equipped with the scalar product $\langle A, B\rangle = \frac{1}{2}\,\mathrm{tr}(A^\dagger B)$, and the subspace of $\mathbb{H}$ spanned by $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ by the subspace of traceless skew-Hermitian matrices.

Remark 5.1.12 It remains to explain why this example runs under the name of universal covering. Let G and $\tilde{G}$ denote either SO(3) and Sp(1) or SO(4) and $\mathrm{Sp}(1) \times \mathrm{Sp}(1)$. Since $\phi : \tilde{G} \to G$ is both an immersion and a submersion, by the Inverse Mapping Theorem, it is a local diffeomorphism. Since the preimages of points of G consist of two points in $\tilde{G}$, ϕ is a two-fold covering of G. Since $\tilde{G}$, as (a product of two copies of) a 3-sphere, is simply connected, it follows by covering theory that ϕ is universal in the sense that for any covering $\psi : H \to G$ there exists a covering $\tilde{\psi} : \tilde{G} \to H$ such that $\psi \circ \tilde{\psi} = \phi$. If ψ is in addition a Lie group homomorphism, then so is $\tilde{\psi}$. Thus, ϕ is universal among the coverings of G by Lie group homomorphisms. It is, therefore, referred to as the universal covering homomorphism of G, and $\tilde{G}$ as the universal covering group of G. The universal covering plays a crucial role in the representation theory of Lie groups, one of the reasons being that the set of representations of $\tilde{G}$ contains the set of representations of G as a subset, see Exercise 5.1.8.

[3]The relation between SO(3) and $\mathbb{R}\mathrm{P}^3$ can be made still more explicit by characterizing rotations by an axis and by an angle of rotation about this axis between 0 and π. We encourage the reader to verify that this yields an identification of SO(3) with the closed ball of radius π in $\mathbb{R}^3$ with antipodal points on the boundary identified.

Example 5.1.13 (Universal covering of $\mathrm{SO}(3,1)_0$) The construction is essentially analogous to that for SO(3) and SO(4). Details are left to the reader as Exercise 5.1.7. Let $\mathrm{S}_2(\mathbb{C})$ denote the real vector space of two-dimensional complex Hermitian matrices. The mapping

$$\lambda : \mathbb{R}^4 \to \mathrm{S}_2(\mathbb{C}), \qquad \lambda(\mathbf{x}) := \begin{bmatrix} x_0 + x_1 & -\mathrm{i}x_2 + x_3 \\ \mathrm{i}x_2 + x_3 & x_0 - x_1 \end{bmatrix},$$

is an isomorphism of real vector spaces. It is isometric with respect to the quadratic forms $\mathbf{x} \mapsto x_0^2 - x_1^2 - x_2^2 - x_3^2$ on $\mathbb{R}^4$ and $A \mapsto \det(A)$ on $\mathrm{S}_2(\mathbb{C})$. For every $a \in \mathrm{M}_2(\mathbb{C})$, the assignment $A \mapsto aAa^\dagger$ defines a linear mapping of $\mathrm{S}_2(\mathbb{C})$ and hence a linear mapping $\phi(a)$ of $\mathbb{R}^4$ by

$$\phi(a)\mathbf{x} := \lambda^{-1}\big(a\lambda(\mathbf{x})a^\dagger\big).$$

The assignment $a \mapsto \phi(a)$, in turn, defines a mapping $\phi : \mathrm{M}_2(\mathbb{C}) \to \mathrm{M}_4(\mathbb{R})$. This mapping is real homogeneous of degree two, hence smooth, and satisfies $\phi(ab) = \phi(a)\phi(b)$. Since for every $a \in \mathrm{SL}(2,\mathbb{C})$ we have $\det(aAa^\dagger) = \det(A)$, ϕ restricts to a Lie group homomorphism (denoted by the same symbol) $\phi : \mathrm{SL}(2,\mathbb{C}) \to \mathrm{O}(3,1)$. By similar arguments as in Example 5.1.11 one finds $\ker(\phi) = \{\pm\mathbb{1}\}$ (the centre of $\mathrm{SL}(2,\mathbb{C})$), $\mathrm{im}(\phi) = \mathrm{SO}(3,1)_0$ (the identity component of $\mathrm{SO}(3,1)$)[4] and that ϕ is a covering homomorphism. To prove that it is universal one has to check that $\mathrm{SL}(2,\mathbb{C})$ is simply connected. To see this, one may use that polar decomposition yields a diffeomorphism $\mathrm{U}(2) \times \mathrm{S}_2(\mathbb{C}) \to \mathrm{GL}(2,\mathbb{C})$, given by $(a, A) \mapsto a\mathrm{e}^A$, see Exercise 5.1.9. Restriction of this diffeomorphism to the submanifold SU(2) of U(2) times the subspace of traceless elements of $\mathrm{S}_2(\mathbb{C})$ induces a diffeomorphism $\mathrm{S}^3 \times \mathbb{R}^3 \to \mathrm{SL}(2,\mathbb{C})$. Hence, ϕ is the universal covering homomorphism and $\mathrm{SL}(2,\mathbb{C})$ is the universal covering group of $\mathrm{SO}(3,1)_0$.

Let us add that under the isomorphism between Hermitian matrices and skew-Hermitian matrices given by multiplication by i, the restriction of ϕ to the subgroup $\mathrm{SU}(2) \subset \mathrm{SL}(2,\mathbb{C})$ reproduces the Lie group homomorphism $\mathrm{Sp}(1) \to \mathrm{SO}(3)$ of Example 5.1.11, transported to $\mathrm{M}_2(\mathbb{C})$ according to the remarks made there.

Exercises

5.1.1 Use the Inverse Mapping Theorem to show that in Definition 5.1.1 it suffices to require that the multiplication mapping be smooth.

5.1.2 Prove that the direct product of Lie groups as defined in Example 5.1.3/3 is a Lie group.

5.1.3 Verify that the mapping (5.1.6) is the inverse of the dual of the vector bundle isomorphism (5.1.5).

5.1.4 Show that the image of Sp(1) under the mapping (5.1.8) coincides with SU(2).

---

[4] In physics, the elements of $\mathrm{SO}(3,1)_0$ are referred to as proper orthochronous Lorentz transformations.

5.1.5 Prove that the mapping ϕ of Example 5.1.11 is an immersion.

5.1.6 Show that $\{R \in \mathrm{SO}(3) : \mathrm{tr}(R) \neq -1\}$ is a dense subset whose elements fulfil $R = \phi(\mathbf{a})$, with

$$\mathbf{a} = \frac{(1 + \mathrm{tr}(R))\mathbf{1} + (R_{23} - R_{32})\mathbf{i} + (R_{31} - R_{13})\mathbf{j} + (R_{12} - R_{21})\mathbf{k}}{2\sqrt{1 + \mathrm{tr}(R)}} \in \mathrm{Sp}(1).$$

Use this to prove that the Lie group homomorphism $\phi : \mathrm{Sp}(1) \to \mathrm{O}(3)$ of Example 5.1.11 is an open mapping.
Hint. Write the defining equation for $R = \phi(\mathbf{a})$ in the form $\sum_{l=1}^{3} R_{kl}\tau_l = \mathbf{a}\tau_k\bar{\mathbf{a}}$, where $\tau_1 = \mathbf{i}$, $\tau_2 = \mathbf{j}$, and $\tau_3 = \mathbf{k}$ and use the relation $\sum_{l=1}^{3} \tau_l \mathbf{q} \tau_l = \mathbf{q} - 4q_0\mathbf{1}$ for all $\mathbf{q} \in \mathbb{H}$.

5.1.7 This exercise complements Examples 5.1.11 and 5.1.13.
(a) Prove (5.1.12). Use this formula to compute $\phi(\cos\frac{\alpha}{2}\mathbf{1} + \sin\frac{\alpha}{2}\mathbf{q})$ for $\mathbf{q} = \mathbf{i}, \mathbf{j}, \mathbf{k}$.
(b) Fill in the details for the universal covering homomorphisms of Examples 5.1.11 and 5.1.13.

5.1.8 Let G, $\tilde{G}$ and H be Lie groups and let $\phi : \tilde{G} \to G$ be a covering homomorphism, see Remark 5.1.12. Show that Lie group homomorphisms $\psi : G \to H$ correspond bijectively to Lie group homomorphisms $\tilde{\psi} : \tilde{G} \to H$ satisfying $\ker(\phi) \subset \ker(\tilde{\psi})$. (This applies in particular to representations, i.e., Lie group homomorphisms $G \to \mathrm{GL}(V)$, where V is some $\mathbb{K}$-vector space.)

5.1.9 Show that the mappings $\mathrm{U}(n) \times \mathrm{S}_n(\mathbb{C}) \to \mathrm{GL}(n, \mathbb{C})$ and $\mathrm{O}(n) \times \mathrm{S}_n(\mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$, given by $(a, A) \mapsto a\mathrm{e}^A$, are diffeomorphisms. (The inverse mappings are referred to as polar decomposition of $\mathrm{GL}(n, \mathbb{C})$ and $\mathrm{GL}(n, \mathbb{R})$, respectively.)
Hint. Show that a and A can be reconstructed from $b = a\mathrm{e}^A$ by means of the formulae $A = \ln\sqrt{b^\dagger b}$ and $a = b\sqrt{b^\dagger b}^{-1}$. Moreover, prove that the assignment of (a, A) to b is smooth. The square root $\sqrt{b^\dagger b}$ is defined by the condition that it acts on the eigenspaces of $b^\dagger b$ as multiplication by the positive square root of the corresponding eigenvalue of $b^\dagger b$.

5.2 The Lie Algebra of a Lie Group

In this section, we will construct the Lie algebra associated with a Lie group. For the notion of Lie algebra and Lie algebra homomorphism, see Definition 3.1.2. Let a Lie group G be given.

Definition 5.2.1 (Left-invariant vector fields) A vector field X on G is called left-invariant if $\mathrm{L}_{a*}X = X$ for all $a \in G$.

Written pointwise, the defining condition reads $\mathrm{L}'_a X_b = X_{ab}$ for all $a, b \in G$. Due to (5.1.4), it is equivalent to

$$X_a = \mathrm{L}'_a X_{\mathbb{1}} \quad \text{for all } a \in G. \tag{5.2.1}$$

Proposition 5.2.2 *The left-invariant vector fields on G form a Lie subalgebra of $\mathfrak{X}(G)$.*

Proof This follows from the linearity of the transport operator and Proposition 3.1.5/4. □

Definition 5.2.3 The Lie subalgebra of $\mathfrak{X}(G)$ of left-invariant vector fields is called the Lie algebra of G.

In the following, we will write $\mathfrak{g}$ for the Lie algebra of G and $X, Y, \ldots$ for its elements.[5]

Proposition 5.2.4 *The mapping $\mathfrak{g} \to \mathrm{T}_{\mathbb{1}}G$, defined by $X \mapsto X_{\mathbb{1}}$, is an isomorphism of real vector spaces.*

Proof The mapping is linear. Due to (5.2.1), it is injective. To see that it is surjective, let $X_{\mathbb{1}} \in \mathrm{T}_{\mathbb{1}}G$ and let χ_{L} denote the left trivialization of $\mathrm{T}G$, see (5.1.5). The mapping $G \to \mathrm{T}G$, defined by $a \mapsto \chi_{\mathrm{L}}(a, X_{\mathbb{1}})$, is a left-invariant vector field which takes the value $X_{\mathbb{1}}$ at $\mathbb{1}$. □

Remark 5.2.5

1. Combining the natural isomorphism $\mathfrak{g} \to \mathrm{T}_{\mathbb{1}}G$ with the identical mapping of G and composing this with the left and right trivializations (5.1.5) of $\mathrm{T}G$, one can express these trivializations in terms of left-invariant vector fields:

$$\chi_{\mathrm{L}}, \chi_{\mathrm{R}} : G \times \mathfrak{g} \to \mathrm{T}G, \qquad \chi_{\mathrm{L}}(a, X) = X_a, \quad \chi_{\mathrm{R}}(a, X) = \mathrm{C}'_{a^{-1}} X_a. \tag{5.2.2}$$

2. The proofs of the following statements are left to the reader (Exercise 5.2.1). Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{K}$ and let $\{e_1, \ldots, e_n\}$ be a basis of $\mathfrak{g}$. Expansion of the commutators of the basis elements,

$$[e_i, e_j] = c^k_{ij} e_k, \quad i, j = 1, \ldots, n,$$

(summation convention) yields n^3 elements c^k_{ij} of $\mathbb{K}$, called the structure constants of $\mathfrak{g}$ relative to the given basis. In terms of the structure constants, the defining properties of a Lie algebra, cf. Definition 3.1.2, read as follows.

Anticommutativity: $c^k_{ij} + c^k_{ji} = 0$ for all $i, j, k = 1, \ldots, n$.
Jacobi identity: $c^k_{ij} c^m_{kl} + c^k_{jl} c^m_{ki} + c^k_{li} c^m_{kj} = 0$ for all $i, j, l, m = 1, \ldots, n$.

Conversely, every system c^k_{ij} of elements of $\mathbb{K}$ with these properties defines a Lie algebra. If $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras over $\mathbb{K}$ and if one can find a basis in $\mathfrak{g}$ and

[5] Starting from the next chapter, we will preferably write $A, B, \ldots$ for the elements of the Lie algebra of a Lie group, because we will often deal with situations where further vector fields occur.

a basis in $\mathfrak{h}$ such that the structure constants of $\mathfrak{g}$ and $\mathfrak{h}$ relative to these bases coincide, the assignment of the appropriate basis vectors to one another defines a Lie algebra isomorphism between $\mathfrak{g}$ and $\mathfrak{h}$.

Example 5.2.6 (Classical Lie groups) Let V be a finite-dimensional $\mathbb{K}$-vector space and let $G \subset \mathrm{End}(V)$ be a classical Lie group. Let $\mathfrak{g}$ be the Lie algebra of G. We determine the smooth mappings $X : G \to \mathrm{End}(V)$ which correspond to the elements of $\mathfrak{g}$ under the natural representation of vector fields on G of Remark 2.7.15. According to Proposition 3.2.13/3, left-invariance implies

$$\Phi_t^X(a) = \Phi_t^X \circ \mathrm{L}_a(\mathbb{1}) = \mathrm{L}_a \circ \Phi_t^X(\mathbb{1}) = a\Phi_t^X(\mathbb{1})$$

and hence

$$X(a) = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} \Phi_t^X(a) = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} a\Phi_t^X(\mathbb{1}) = aX(\mathbb{1}).$$

Hence, the elements of $\mathfrak{g}$ are represented by the mappings

$$X^A : G \to \mathrm{End}(V), \qquad X^A(a) = aA \tag{5.2.3}$$

with $A \in \mathrm{T}_{\mathbb{1}}G \subset \mathrm{End}(V)$. Via this representation of $\mathfrak{g}$, the assignment $A \mapsto X^A$ coincides with the inverse of the natural isomorphism of Proposition 5.2.4. We compute the commutator in $\mathfrak{g}$ in terms of the mappings X^A. Since the equation for the integral curves of X^A is $\dot{\gamma}(t) = \gamma(t)A$, the flow of X^A is given by

$$\Phi_t^{X^A}(a) = a\mathrm{e}^{tA}, \quad a \in G. \tag{5.2.4}$$

Using (3.1.4), for $A, B \in \mathrm{T}_{\mathbb{1}}G$ and $a \in G$ we obtain

$$\left[X^A, X^B\right](a) = a[A, B].$$

This yields

$$\left[X^A, X^B\right] = X^{[A,B]} \quad \text{for all } A, B \in \mathrm{T}_{\mathbb{1}}G. \tag{5.2.5}$$

Evaluation of (5.2.5) at $a = \mathbb{1}$ shows that, in particular, $\mathrm{T}_{\mathbb{1}}G$ is closed under the commutator of endomorphisms. Let $\mathfrak{gl}(V)$ denote the real vector space underlying $\mathrm{End}(V)$, equipped with the commutator of endomorphisms as a multiplication. Then, $\mathfrak{gl}(V)$ is a real Lie algebra, $\mathrm{T}_{\mathbb{1}}G$ is a Lie subalgebra and (5.2.5) means that the natural isomorphism $\mathrm{T}_{\mathbb{1}}G \to \mathfrak{g}$ of Proposition 5.2.4 is an isomorphism of Lie algebras. Henceforth, we will identify $\mathfrak{g}$ with the Lie subalgebra $\mathrm{T}_{\mathbb{1}}G$ of $\mathfrak{gl}(V)$. To conclude, we take $V = \mathbb{K}^n$ and list the subspaces $\mathrm{T}_{\mathbb{1}}G = \ker(f'_{\mathbb{1}})$ of $\mathfrak{gl}(n, \mathbb{K})$ for the classical groups of Example 1.2.6.[6] Computations are left to the reader (Exer-

[6] Several of these Lie algebras are in addition, and more naturally, complex Lie algebras, notably $\mathfrak{gl}(n, \mathbb{C})$ itself and $\mathfrak{sl}(n, \mathbb{C})$. This corresponds to the fact that the respective classical Lie groups are in addition, and again more naturally, complex analytic Lie groups.

cise 5.2.2). In the following, $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

$$\begin{aligned}
\mathfrak{sl}(n, \mathbb{K}) &= \{A \in M_n(\mathbb{K}) \colon \operatorname{tr} A = 0\}, \\
\mathfrak{o}(n, m) = \mathfrak{so}(n, m) &= \{A \in M_{n+m}(\mathbb{R}) \colon \mathbb{1}_{n,m} A + A^T \mathbb{1}_{n,m} = 0\}, \\
\mathfrak{o}(n, \mathbb{C}) = \mathfrak{so}(n, \mathbb{C}) &= \{A \in M_n(\mathbb{C}) \colon A + A^T = 0\}, \\
\mathfrak{u}(n, m) &= \{A \in M_{n+m}(\mathbb{C}) \colon \mathbb{1}_{n,m} A + A^\dagger \mathbb{1}_{n,m} = 0\}, \\
\mathfrak{su}(n, m) &= \{A \in M_{n+m}(\mathbb{C}) \colon \mathbb{1}_{n,m} A + A^\dagger \mathbb{1}_{n,m} = 0, \operatorname{tr} A = 0\}, \\
\mathfrak{sp}(n, \mathbb{K}) &= \{A \in M_{2n}(\mathbb{K}) \colon A^T J_n + J_n A = 0\}, \\
\mathfrak{sp}(n, m) &= \{A \in M_{n+m}(\mathbb{H}) \colon \mathbb{1}_{n,m} A + A^\dagger \mathbb{1}_{n,m} = 0\}.
\end{aligned}$$

Remark 5.2.7 The vector field X^A of Example 5.2.6 is the restriction to G of the linear vector field on $\operatorname{End}(V)$ which is generated by the linear mapping $\mathrm{R}_A : \operatorname{End}(V) \to \operatorname{End}(V)$, $\mathrm{R}_A(B) := BA$. According to Example 3.2.8, the flow of this vector field is given by the family of vector space automorphisms $\mathrm{e}^{t\mathrm{R}_A}$ of $\operatorname{End}(V)$, where $t \in \mathbb{R}$. A brief computation shows $\mathrm{e}^{t\mathrm{R}_A}(B) = B\mathrm{e}^{tA}$ for all $B \in \operatorname{End}(V)$. For $B = a$, this reproduces (5.2.4).

Example 5.2.8 From Example 5.2.6 we read off that the Lie algebras $\mathfrak{sp}(1)$, $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are spanned, respectively, by the bases

$$\mathrm{I}_1^{\mathbb{H}} = \frac{1}{2}\mathbf{i}, \qquad \mathrm{I}_2^{\mathbb{H}} = \frac{1}{2}\mathbf{j}, \qquad \mathrm{I}_3^{\mathbb{H}} = \frac{1}{2}\mathbf{k},$$

$$\mathrm{I}_1^{\mathbb{C}} = \frac{1}{2}\begin{bmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{bmatrix}, \qquad \mathrm{I}_2^{\mathbb{C}} = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad \mathrm{I}_3^{\mathbb{C}} = \frac{1}{2}\begin{bmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix},$$

$$\mathrm{I}_1^{\mathbb{R}} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \mathrm{I}_2^{\mathbb{R}} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \mathrm{I}_3^{\mathbb{R}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}.$$

Since these bases satisfy

$$[\mathrm{I}_i^{\mathbb{K}}, \mathrm{I}_j^{\mathbb{K}}] = \varepsilon_{ij}{}^k \mathrm{I}_k^{\mathbb{K}}, \quad \mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{H},$$

they define isomorphisms between these Lie algebras. In addition, since the vector product in $\mathbb{R}^3$ is given by

$$\mathbf{e}_i \times \mathbf{e}_j = \varepsilon_{ij}{}^k \mathbf{e}_k,$$

these isomorphisms relate the respective Lie bracket to the vector product in $\mathbb{R}^3$. Thus, the Lie algebras $\mathfrak{sp}(1)$, $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are isomorphic to one another and to the Lie algebra $(\mathbb{R}^3, \times)$. In particular, for further use we note the explicit form of this isomorphism for $\mathfrak{so}(3)$,

$$\varphi : \mathbb{R}^3 \to \mathfrak{so}(3), \quad \mathbf{x} \mapsto \varphi(\mathbf{x}) := \begin{bmatrix} 0 & x^1 & -x^2 \\ -x^1 & 0 & x^3 \\ x^2 & -x^3 & 0 \end{bmatrix}. \tag{5.2.6}$$

Then,

$$\varphi(\mathbf{x} \times \mathbf{y}) = \big[\varphi(\mathbf{x}), \varphi(\mathbf{y})\big], \quad \mathbf{x} \cdot \mathbf{y} = -\operatorname{tr}\big(\varphi(\mathbf{x})\varphi(\mathbf{y})\big). \tag{5.2.7}$$

Remark 5.2.9 We comment on the proof of the parallelizability of the spheres S^1 and S^3, cf. Proposition 2.3.17. Identifying S^1 with $\mathrm{U}(1)$ and $\mathrm{TU}(1)$ with a submanifold of $\mathrm{U}(1) \times \mathbb{C}$, the Lie algebra $\mathfrak{u}(1) \equiv \mathrm{T}_{\mathbb{1}}\mathrm{U}(1)$ corresponds to the imaginary axis and the vector field used in the proof of the parallelizability of S^1 is the left-invariant vector field generated by the complex imaginary unit i. Similarly, $\mathrm{S}^3 = \mathrm{Sp}(1)$ and the vector fields used in the proof of the parallelizability of S^3 are the left-invariant vector fields generated by the quaternionic imaginary units **i**, **j** and **k**.

Example 5.2.10 (Vector groups) Let V be a $\mathbb{K}$-vector space, viewed as a Lie group. A computation analogous to that for the general linear group of Example 5.2.6 yields that, via the natural representation of vector fields on V by smooth mappings $V \to V$, left-invariant vector fields correspond to constant mappings and that the commutator is trivial (Exercise 5.2.4). Hence, via this representation, the natural isomorphism of Proposition 5.2.4 identifies the Lie algebra of V with the real Lie algebra whose underlying vector space is obtained from V by field restriction and whose multiplication is trivial.

Example 5.2.11 (Discrete groups) Let G be a discrete Lie group. Since the manifold structure of G has dimension zero, the tangent bundle has dimension zero as well. Hence, the only vector field is the zero section, which is obviously left-invariant. Thus, the Lie algebra of G is given by $\{0\}$ with the obvious multiplication.

Next, we show that every homomorphism of Lie groups induces a homomorphism of the associated Lie algebras. Let G and H be Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively, and let $\varphi\colon G \to H$ be a Lie group homomorphism. Due to $\varphi(\mathbb{1}) = \mathbb{1}$, φ' maps $\mathrm{T}_{\mathbb{1}}G$ to $\mathrm{T}_{\mathbb{1}}H$. Hence, composition with the natural vector space isomorphisms $\mathfrak{g} \to \mathrm{T}_{\mathbb{1}}G$ and $\mathfrak{h} \to \mathrm{T}_{\mathbb{1}}H$ of Proposition 5.2.4 yields a linear mapping $\mathrm{d}\varphi : \mathfrak{g} \to \mathfrak{h}$, given by

$$\big(\mathrm{d}\varphi(X)\big)_b := \mathrm{L}'_b \circ \varphi'(X_{\mathbb{1}}) \quad \text{for all } b \in H. \tag{5.2.8}$$

Remark 5.2.12 Since $\varphi' X_a = \varphi' \circ \mathrm{L}'_a(X_{\mathbb{1}}) = \mathrm{L}'_{\varphi(a)} \circ \varphi'(X_{\mathbb{1}}) = (\mathrm{d}\varphi(X))_{\varphi(a)}$, the image $\mathrm{d}\varphi(X)$ is the unique left-invariant vector field on H which is φ-related to X.

Proposition 5.2.13 *The mapping* $\mathrm{d}\varphi$ *is a homomorphism of Lie algebras.*

The mapping $\mathrm{d}\varphi$ is called the homomorphism of Lie algebras induced by φ.

Proof Let $X, Y \in \mathfrak{g}$. Since $\mathrm{d}\varphi(X)$ and $\mathrm{d}\varphi(Y)$ are φ-related to X and Y, respectively, Proposition 3.1.5/2 yields that $[\mathrm{d}\varphi(X), \mathrm{d}\varphi(Y)]$ is φ-related to $[X, Y]$. Since $[\mathrm{d}\varphi(X), \mathrm{d}\varphi(Y)]$ is left-invariant, Remark 5.2.12 implies

$$\big[\mathrm{d}\varphi(X), \mathrm{d}\varphi(Y)\big] = \mathrm{d}\varphi\big([X, Y]\big). \qquad \square$$

Remark 5.2.14

1. For Lie group homomorphisms $\varphi : G \to H$ and $\psi : H \to K$ we have
$$\mathrm{d}(\psi \circ \varphi) = \mathrm{d}\psi \circ \mathrm{d}\varphi, \qquad \mathrm{d}\,\mathrm{id}_G = \mathrm{id}_{\mathfrak{g}}. \tag{5.2.9}$$
2. If φ is an isomorphism, $\mathrm{d}\varphi$ is bijective and hence an isomorphism of Lie algebras. Moreover, since $\mathrm{d}\varphi(X)$ is φ-related to X, then $\mathrm{d}\varphi(X) = \varphi_* X$ for all $X \in \mathfrak{g}$.

Example 5.2.15 (Classical Lie groups) Let G and H be classical Lie groups. Under the natural identification of the Lie algebras of the classical Lie groups with Lie subalgebras of $\mathfrak{gl}(V)$ for appropriate $\mathbb{K}$-vector spaces V, see Example 5.2.6, the induced homomorphism $\mathrm{d}\varphi : \mathfrak{g} \to \mathfrak{h}$ coincides with the tangent mapping $\varphi'_{\mathbb{1}}$. In particular, this implies that the latter respects the commutator of vector space endomorphisms. This way, for the Lie group homomorphisms of Sect. 5.1 one obtains the following induced homomorphisms (Exercise 5.2.5).

1. For the natural inclusion mapping $j : G \to \mathrm{GL}(V)$, the induced homomorphism $\mathrm{d}j$ is given by the natural inclusion mapping $\mathfrak{g} \to \mathfrak{gl}(V)$.
2. For the matrix representation $\mathrm{GL}(V) \to \mathrm{GL}(n, \mathbb{K})$ associated with a basis in V, the induced homomorphism $\mathfrak{gl}(V) \to \mathfrak{gl}(n, \mathbb{K})$ is given by the corresponding matrix representation of endomorphisms.
3. For the determinant homomorphism $\det : \mathrm{GL}(n, \mathbb{K}) \to \mathbb{K} \setminus \{0\}$, where $\mathbb{K} = \mathbb{R}, \mathbb{C}$, the induced homomorphism $\mathrm{d}\det : \mathfrak{gl}(n, \mathbb{K}) \to \mathbb{K}$ is given by the trace,
$$\mathrm{d}\det = \mathrm{tr}\,. \tag{5.2.10}$$
4. The Lie algebra isomorphism $\mathfrak{sp}(1) \to \mathfrak{su}(2)$ induced by the Lie group isomorphism $\mathrm{Sp}(1) \to \mathrm{SU}(2)$ of Example 5.1.10 coincides with the isomorphism defined in Example 5.2.8 in terms of bases.
5. For the covering homomorphisms $\mathrm{Sp}(1) \to \mathrm{SO}(3)$, $\mathrm{Sp}(1) \times \mathrm{Sp}(1) \to \mathrm{SO}(4)$ of Example 5.1.11 and $\mathrm{SL}(2, \mathbb{C}) \to \mathrm{SO}(3, 1)_0$ of Example 5.1.13, the induced homomorphisms are bijective and hence isomorphisms. For the first one, the induced isomorphism coincides with the isomorphism of Example 5.2.8. For the second one, under the identification $\mathbb{R}^4 \cong \mathbb{H}$ of Example 5.1.11, we find that the induced isomorphism is given by
$$\mathrm{d}\varphi(A, B)\mathbf{q} = A\mathbf{q} - \mathbf{q}B, \quad A, B \in \mathfrak{sp}(1), \quad \mathbf{q} \in \mathbb{H}. \tag{5.2.11}$$

Example 5.2.16 (Direct product) Let G_1 and G_2 be Lie groups with Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, respectively. The Lie algebra of the direct product of Lie groups $G = G_1 \times G_2$ is isomorphic to the direct product of Lie algebras $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ with multiplication

$$\big[(X_1, X_2), (Y_1, Y_2)\big] = \big([X_1, Y_1], [X_2, Y_2]\big), \quad X_i, Y_i \in \mathfrak{g}_i.$$

Under this identification, for a pair of Lie group homomorphisms $\varphi_i : G_i \to H_i$, $i = 1, 2$, we have

$$\mathrm{d}(\varphi_1 \times \varphi_2) = \mathrm{d}\varphi_1 \oplus \mathrm{d}\varphi_2. \tag{5.2.12}$$

Proofs are left to the reader (Exercise 5.2.7).

Example 5.2.17 (Vector groups) Let V and W be $\mathbb{K}$-vector spaces, viewed as Lie groups, and let $\varphi : V \to W$ be a Lie group homomorphism. Then, φ is continuous and additive. Since every continuous additive mapping of real vector spaces is in fact linear (Exercise 5.2.6), φ induces a linear mapping of the real vector spaces underlying V and W. Under the identification of the Lie algebras of V and W with these real vector spaces, $\mathrm{d}\varphi$ coincides with φ.

Remark 5.2.18 There exist several isomorphisms between the classical Lie algebras in low dimensions. The isomorphism between SU(2) and Sp(1) of Example 5.1.10 and the covering homomorphisms of Examples 5.1.11 and 5.1.13 yield

$$\mathfrak{so}(3) \cong \mathfrak{su}(2) \cong \mathfrak{sp}(1), \qquad \mathfrak{so}(4) \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2), \qquad \mathfrak{so}(3,1) \cong \mathfrak{sl}(2, \mathbb{C}),$$

cf. Examples 5.2.15/4 and 5.2.15/5. Analogously to the first two isomorphisms one constructs isomorphisms

$$\mathfrak{so}(3, \mathbb{C}) \cong \mathfrak{sl}(2, \mathbb{C}), \qquad \mathfrak{so}(4, \mathbb{C}) \cong \mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C}),$$

respectively. To prove the latter, for example, one uses the isomorphism

$$F : \mathbb{C}^4 \to \mathrm{End}\big(\mathbb{C}^2\big), \qquad F(\mathbf{z}) := \begin{bmatrix} z_0 + \mathrm{i}z_1 & z_2 + \mathrm{i}z_3 \\ -z_2 + \mathrm{i}z_3 & z_0 - \mathrm{i}z_1 \end{bmatrix},$$

and considers the mapping

$$\varphi : \mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C}) \to \mathrm{M}_4(\mathbb{C}), \qquad \varphi(a, b)\mathbf{z} := F^{-1}\big(aF(\mathbf{z})b^{-1}\big).$$

Using that $\det F(\mathbf{z})$ is a non-degenerate quadratic form on $\mathbb{C}^4$, one shows that φ is the universal covering homomorphism of $\mathrm{SO}(4, \mathbb{C})$. Then, the induced Lie algebra homomorphism yields the desired isomorphism. Alternatively, this isomorphism follows by complexification of the isomorphism $\mathfrak{so}(4) \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2)$. By similar arguments [291], one obtains

$$\begin{aligned}
\mathfrak{so}(2,1) &\cong \mathfrak{su}(1,1) \cong \mathfrak{sl}(2,\mathbb{R}),\\
\mathfrak{so}(2,2) &\cong \mathfrak{sl}(4,\mathbb{R}) \cong \mathfrak{sl}(2,\mathbb{R}) \oplus \mathfrak{sl}(2,\mathbb{R}),\\
\mathfrak{so}(5) &\cong \mathfrak{sp}(2),\\
\mathfrak{so}(6) &\cong \mathfrak{su}(4),\\
\mathfrak{so}(6,\mathbb{C}) &\cong \mathfrak{sl}(4,\mathbb{C}),\\
\mathfrak{so}(2,4) &\cong \mathfrak{su}(2,2).
\end{aligned}$$

For occasional further use, we state the definitions of the following special types of Lie algebras and Lie groups [127, 145, 149]. Recall that an ideal in a Lie algebra $\mathfrak{g}$ is a linear subspace $\mathfrak{i}$ such that $[\mathfrak{i}, \mathfrak{g}] \subset \mathfrak{i}$.

Definition 5.2.19 A Lie algebra $\mathfrak{g}$ is called

1. simple if it is not Abelian and if it does not contain a nontrivial ideal,
2. semisimple if it does not contain a nonzero Abelian ideal,
3. solvable if for some n, the subspace $\mathfrak{g}^{(n)}$, defined recursively by $\mathfrak{g}^{(k+1)} = [\mathfrak{g}^{(k)}, \mathfrak{g}^{(k)}]$ and $\mathfrak{g}^{(0)} = \mathfrak{g}$, satisfies $\mathfrak{g}^{(n)} = \{0\}$,
4. nilpotent if for some n, the subspace $\mathfrak{g}_{(n)}$, defined recursively by $\mathfrak{g}_{(k+1)} = [\mathfrak{g}, \mathfrak{g}_{(k)}]$ and $\mathfrak{g}_{(0)} = \mathfrak{g}$, satisfies $\mathfrak{g}_{(n)} = \{0\}$.

A Lie group G is called, respectively, simple, semisimple, solvable or nilpotent if its Lie algebra is simple, semisimple, solvable or nilpotent.

Since ideals of $\mathfrak{g}$ are in particular Lie algebras, these notions apply to them as well. Using the Jacobi identity, by induction on n one can show that the subspaces $\mathfrak{g}^{(n)}$ and $\mathfrak{g}_{(n)}$ are ideals of $\mathfrak{g}$. It follows that every nonzero solvable or nilpotent ideal contains a nonzero Abelian ideal. Hence, a Lie algebra is semisimple iff it does not have nonzero solvable ideals, or iff it does not have nonzero nilpotent ideals.

Example 5.2.20 The classical Lie groups $\mathrm{SL}(n,\mathbb{K})$ for $n \geq 2$, $\mathrm{SU}(n)$ for $n \geq 2$, $\mathrm{SO}(n)$ for $n \neq 1, 2, 4$ and $\mathrm{Sp}(n)$ for $n \geq 1$ are simple. Note that, except for $\mathrm{SO}(n)$ with n odd, none of these groups is simple in the sense of ordinary group theory, because each of them has a nontrivial centre and hence possesses a nontrivial normal subgroup. According to Remark 5.2.18, $\mathrm{SO}(4)$ is semisimple but not simple. The subgroup of $\mathrm{GL}(n,\mathbb{K})$ of upper triangular matrices is solvable. The subgroup of $\mathrm{GL}(n,\mathbb{K})$ of strictly upper triangular matrices is nilpotent. Finally, for $n \geq 2$, as Lie algebras

$$\mathfrak{gl}(n,\mathbb{K}) \cong \mathbb{K} \oplus \mathfrak{sl}(n,\mathbb{K}), \qquad \mathfrak{u}(n) \cong \mathrm{i}\mathbb{R} \oplus \mathfrak{su}(n),$$

where $\mathbb{K}$ and $\mathrm{i}\mathbb{R}$ carry the trivial multiplication. Hence, $\mathrm{GL}(n,\mathbb{K})$ and $\mathrm{U}(n)$ are neither semisimple nor solvable or nilpotent. Proofs are left to the reader (Exercise 5.2.8).

We conclude this section with a remark on right-invariant vector fields.

Remark 5.2.21 Similarly to left-invariant vector fields, by the condition $\mathrm{R}_{a*}X = X$, one defines right-invariant vector fields on G and shows the following (Exercise 5.2.9).

1. Equivalent conditions for a vector field X to be right-invariant are $\mathrm{R}'_a X_b = X_{ba}$ for all $a, b \in G$ and $X_a = \mathrm{R}'_a X_{\mathbb{1}}$ for all $a \in G$.
2. The right-invariant vector fields on G form a Lie subalgebra of $\mathfrak{X}(G)$. The assignment $X \mapsto X_{\mathbb{1}}$ defines an isomorphism of real vector spaces from this subalgebra onto $\mathrm{T}_{\mathbb{1}}G$. Via this isomorphism, the left and right trivializations of $\mathrm{T}G$ can be expressed in terms of right-invariant vector fields by

$$\chi_{\mathrm{L}}(a, X) = C'_a X_a, \qquad \chi_{\mathrm{R}}(a, X) = X_a. \tag{5.2.13}$$

3. Let $\mathrm{inv} : G \to G$ be the inversion mapping. Since $\mathrm{L}_a \circ \mathrm{inv} = \mathrm{inv} \circ \mathrm{R}_{a^{-1}}$ and $\mathrm{inv}'_{\mathbb{1}} = -\mathrm{id}_{\mathrm{T}_{\mathbb{1}}G}$, the negative of the transport operator inv_* maps the left and right-invariant vector fields which share the same value at $\mathbb{1}$ onto one another. Thus, $-\mathrm{inv}_*$ defines an anti-isomorphism between the Lie subalgebras of left-invariant and right-invariant vector fields.
4. If G is a classical Lie group contained in $\mathrm{End}(V)$, right-invariant vector fields on G correspond to mappings $a \mapsto {}^A X(a) = Aa$ with $A \in \mathrm{End}(V)$. One can check that ${}^A X = -\,\mathrm{inv}_* X^A$. This implies $[{}^A X, {}^B X] = [B, A]$.

Exercises

5.2.1 Prove the statements of Remark 5.2.5/2.

5.2.2 For the classical Lie groups $H \subset \mathrm{M}_n(\mathbb{K})$ of Example 1.2.6, determine the subspaces $\mathrm{T}_{\mathbb{1}}H \subset \mathrm{M}_n(\mathbb{K})$. By a direct computation, check that these subspaces are Lie subalgebras. Compare your result with the list of classical Lie algebras in Example 5.2.6.

5.2.3 Construct bases for the classical Lie algebras of Example 5.2.6 and determine the corresponding structure constants.

5.2.4 Carry out the computations necessary for Example 5.2.10.

5.2.5 Complete Example 5.2.15 by computing the induced homomorphisms of Lie algebras for the Lie group homomorphisms of Sect. 5.1 involving classical Lie groups.

5.2.6 Show that every continuous additive mapping of real vector spaces is linear. *Hint*. Use that additivity implies linearity over the rationals.

5.2.7 Prove the assertions of Example 5.2.16.

5.2.8 Prove the assertions of Example 5.2.20.

5.2.9 Verify the properties of right-invariant vector fields stated in Remark 5.2.21.

5.3 The Exponential Mapping

In this section, we will construct the exponential mapping associated with a Lie group. Via this mapping, the algebraic structure of the Lie algebra encodes the local structure of the Lie group. This way, a considerable part of the theory of Lie groups can be reduced to the study of Lie algebras. Let G be a Lie group and let $\mathfrak{g}$ be the Lie algebra of G.

Proposition 5.3.1 *Every left-invariant vector field on a Lie group is complete.*

Proof Let $X \in \mathfrak{g}$ and let $\Phi : \mathscr{D} \to G$ be the flow of X. According to Proposition 3.2.13/3, left-invariance implies $\mathrm{L}_a(\mathscr{D}_t) = \mathscr{D}_t$ for all $a \in G$ and $t \in \mathbb{R}$. Hence, for every t, either $\mathscr{D}_t = G$ or $\mathscr{D}_t = \varnothing$. Since $\mathscr{D}$ is open in $G \times \mathbb{R}$, there exist $t_+ > 0$ and $t_- < 0$ such that $\mathscr{D}_{t_\pm}$ is nonempty and hence coincides with G. Then, $\Phi_{t_\pm}(\mathbb{1})$ is defined and, as an element of G, it belongs to $\mathscr{D}_{t_\pm}$, so that $\Phi_{2t_\pm}(\mathbb{1})$ is defined. Iteration of this argument yields that $\mathscr{D}_t \neq \varnothing$ and hence $\mathscr{D}_t = G$ for all $t \in \mathbb{R}$. Thus, X is complete. □

Remark 5.3.2 Let $X \in \mathfrak{g}$ and let Φ denote the flow of X. By left-invariance, for all $t, s \in \mathbb{R}$ we have

$$\Phi_{t+s}(\mathbb{1}) = \Phi_t \circ \Phi_s(\mathbb{1}) = \Phi_t \circ \mathrm{L}_{\Phi_s(\mathbb{1})}(\mathbb{1}) = \mathrm{L}_{\Phi_s(\mathbb{1})} \circ \Phi_t(\mathbb{1}) = \Phi_s(\mathbb{1})\Phi_t(\mathbb{1}).$$

Thus, the maximal integral curve through $\mathbb{1}$ defines a Lie group homomorphism $\mathbb{R} \to G$, that is, a one-parameter subgroup of G. It is not hard to show that, conversely, every one-parameter subgroup of G is the maximal integral curve through $\mathbb{1}$ of a unique left-invariant vector field on G (Exercise 5.3.1). Thus, one-parameter subgroups are in bijective correspondence with left-invariant vector fields.

Definition 5.3.3 (Exponential mapping) The exponential mapping of G is defined by

$$\exp : \mathfrak{g} \to G, \qquad \exp(X) = \Phi^X_1(\mathbb{1}),$$

where Φ^X is the flow of X.

If necessary, we will write $\exp_G$ to distinguish between the exponential mappings of different Lie groups.

Remark 5.3.4

1. Since the zero vector field has the trivial flow $\Phi^0_t = \mathrm{id}_G$ for all $t \in \mathbb{R}$,

$$\exp(0) = \mathbb{1}. \tag{5.3.1}$$

2. Let $X \in \mathfrak{g}$ and let us consider the mapping $t \mapsto \exp(tX)$. Due to the scaling property (3.2.15) of vector fields, for all $t \in \mathbb{R}$ we have

$$\exp(tX) = \Phi^X_t(\mathbb{1}), \tag{5.3.2}$$

that is, the mapping $t \mapsto \exp(tX)$ yields the maximal integral curve of X through $\mathbb{1}$. Since, according to Remark 5.3.2, the latter is a group homomorphism, for all $t, s \in \mathbb{R}$ we obtain

$$\exp\big((t+s)X\big) = \exp(tX)\exp(sX), \qquad \exp(-tX) = \big(\exp(tX)\big)^{-1}. \quad (5.3.3)$$

Furthermore, by left-invariance, (5.3.2) yields $\Phi^X_t(a) = \mathrm{L}_a \circ \Phi^X_t(\mathbb{1}) = a\exp(tX)$, hence

$$\Phi^X_t = \mathrm{R}_{\exp(tX)}. \quad (5.3.4)$$

Repeated application yields

$$\Phi^Y_s \circ \Phi^X_t(a) = \mathrm{R}_{\exp(sY)} \circ \mathrm{R}_{\exp(tX)}(a) = a\exp(tX)\exp(sY). \quad (5.3.5)$$

3. Let X be a right-invariant vector field on G and let $\mathrm{inv} : G \to G$ denote the inversion mapping. According to Remark 5.2.21, $\mathrm{inv}_* X$ is left-invariant. Hence, Proposition 3.2.13, together with (5.3.3) and (5.3.4), implies

$$\Phi^X_t(a) = \mathrm{inv} \circ \Phi^{\mathrm{inv}_* X}_t \circ \mathrm{inv}(a) = \exp(-t\,\mathrm{inv}_* X)a. \quad (5.3.6)$$

That is, the flow of X is given by left translation by $\exp(-t\,\mathrm{inv}_* X)$. In particular, X is complete.

We derive the basic properties of the exponential mapping.

Proposition 5.3.5 *The exponential mapping of G is smooth. By restriction, it induces a diffeomorphism from an open neighbourhood of 0 in $\mathfrak{g}$ onto an open neighbourhood of $\mathbb{1}$ in G.*

Proof Consider the vector field $\tilde{Y}$ on $G \times \mathfrak{g}$ defined by $\tilde{Y}(a, X) := (X_a, 0)$. Its flow is given by $\tilde{\Phi}_t(a, X) = (\Phi^X_t(a), X)$, where Φ^X denotes the flow of X. Thus,

$$\exp(X) = \Phi^X_1(\mathbb{1}) = \mathrm{pr}_G \circ \tilde{\Phi}_1(\mathbb{1}, X),$$

where $\mathrm{pr}_G : G \times \mathfrak{g} \to G$ is the natural projection. Hence, exp is smooth. To see that it restricts to a diffeomorphism between open neighbourhoods of 0 in $\mathfrak{g}$ and $\mathbb{1}$ in G, we compute the tangent mapping at $0 \in \mathfrak{g}$. Using (5.3.2), we obtain

$$\exp'_0 X = \frac{\mathrm{d}}{\mathrm{d}t}\Big\restriction_0 \exp(tX) = \frac{\mathrm{d}}{\mathrm{d}t}\Big\restriction_0 \Phi^X_t(\mathbb{1}) = X_{\mathbb{1}}, \qquad X \in \mathfrak{g},$$

that is, $\exp'_0$ coincides with the natural isomorphism $\mathfrak{g} \to \mathrm{T}_{\mathbb{1}}G$ of Proposition 5.2.4. Now, the assertion follows from the Inverse Mapping Theorem 1.5.7. □

Proposition 5.3.6 *For a Lie group homomorphism $\varphi : G \to H$, one has*

$$\varphi \circ \exp_G = \exp_H \circ \mathrm{d}\varphi.$$

Proof Let $X \in \mathfrak{g}$. According to Remark 5.2.12, $\mathrm{d}\varphi(X)$ is φ-related to X. According to Proposition 3.2.13/1, for all $t \in \mathbb{R}$ we get

$$\Phi_t^{\mathrm{d}\varphi(X)} \circ \varphi = \varphi \circ \Phi_t^X .$$

Evaluation of both sides for $t = 1$ at $\mathbb{1}_G$ yields the assertion. □

Corollary 5.3.7 *For a Lie group homomorphism $\varphi : G \to H$, the following is equivalent.*

1. *φ is an immersion (submersion).*
2. *φ has discrete kernel (is open).*
3. *$\mathrm{d}\varphi$ is injective (surjective).*

In particular, if a Lie group homomorphism is surjective and has discrete kernel, it is a covering homomorphism, see Examples 5.1.11 and 5.1.13.

Proof $1 \Rightarrow 2$: If φ is a submersion, this is due to Remark 1.5.16. If φ is an immersion, the Constant Rank Theorem 1.5.11 implies that every point of G has a neighbourhood on which φ is injective. In particular, every element of $\ker(\varphi)$ has a neighbourhood containing no other element of $\ker(\varphi)$. Thus, $\ker(\varphi)$ is discrete.

$2 \Rightarrow 3$: First, assume that φ has discrete kernel. If $\mathrm{d}\varphi$ was not injective, there would exist a nonzero $A \in \mathfrak{g}$ such that $\mathrm{d}\varphi(A) = 0$. Then, Proposition 5.3.6 would imply that

$$\mathbb{1}_H = \exp_H\big(\mathrm{d}\varphi(tA)\big) = \varphi\big(\exp_G(tA)\big)$$

and hence $\exp_G(tA) \in \ker(\varphi)$ for all $t \in \mathbb{R}$. Since $t \mapsto \exp_G(tA)$ is a non-constant curve through $\mathbb{1}_G$ and since $\exp_G$ is a local diffeomorphism, this would contradict the assumption that $\ker(\varphi)$ be discrete.

Now, assume that φ is open. Since $\mathrm{im}(\exp_G)$ is open, then $\varphi(\exp_G(\mathfrak{g}))$ and hence $\exp_H(\mathrm{d}\varphi(\mathfrak{g}))$ is open in H. Since $\exp_H$ is a local diffeomorphism, $\mathrm{d}\varphi(\mathfrak{g})$ must contain an open subset. Hence, $\mathrm{d}\varphi$ is surjective.

$3 \Rightarrow 1$: By left-invariance, for $a \in G$ and $X \in \mathfrak{g}$, we have $\varphi'_a X_a = (\mathrm{d}\varphi(X))_{\varphi(a)}$. Hence, if $\mathrm{d}\varphi$ is injective (surjective), then φ is an immersion (submersion). □

Corollary 5.3.8 *For a homomorphism of Lie groups to be an isomorphism it suffices to be bijective.*

Proof If φ is bijective, then Corollary 5.3.7 implies that it is an immersion and a submersion. It follows that it is a diffeomorphism (by the Inverse Mapping Theorem 1.5.7) and that its inverse is a group homomorphism (by elementary algebra). □

Finally, Propositions 5.1.7 and 5.3.6 imply

Corollary 5.3.9 *Let $\varphi, \psi : G \to H$ be Lie group homomorphisms. If $\mathrm{d}\varphi = \mathrm{d}\psi$, then φ and ψ coincide on the identity component of G.*

Proposition 5.3.10 *For $X, Y \in \mathfrak{g}$ such that $[X, Y] = 0$ one has*

$$\exp(X)\exp(Y) = \exp(Y)\exp(X) = \exp(X + Y).$$

In particular, if $\mathfrak{g}$ is Abelian, $\exp$ is a Lie group homomorphism from the vector group underlying $\mathfrak{g}$ to G.

Proof It suffices to show that the smooth curve $\gamma : \mathbb{R} \to G$, defined by $\gamma(t) := \exp(tX)\exp(tY)$, is an integral curve through $\mathbb{1}$ of the vector field $X + Y$. Indeed, $\gamma(0) = \mathbb{1}$. By (5.3.5),

$$\dot{\gamma}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\big(\Phi_t^Y \circ \Phi_t^X(\mathbb{1})\big).$$

According to Proposition 3.2.15, the flows of X and Y commute. Using this and the product rule (2.2.8), one obtains

$$\dot{\gamma}(t) = \frac{\mathrm{d}}{\mathrm{d}s}\restriction_0 \Phi_{t+s}^Y \circ \Phi_t^X(\mathbb{1}) + \frac{\mathrm{d}}{\mathrm{d}s}\restriction_0 \Phi_{t+s}^X \circ \Phi_t^Y(\mathbb{1}) = Y_{\gamma(t)} + X_{\gamma(t)}.$$ □

If X and Y do not commute, one has the following approximate formula.

Proposition 5.3.11 *For all $X, Y \in \mathfrak{g}$ there exists $\varepsilon > 0$ such that for all $|t| < \varepsilon$,*

$$\exp(tX)\exp(tY) = \exp\left(t(X + Y) + \frac{1}{2}t^2[X, Y] + O\big(t^3\big)\right), \tag{5.3.7}$$

where $O(t^3)$ is the value of a smooth mapping $(-\varepsilon, \varepsilon) \to \mathfrak{g}$ such that $O(t^3)/t^3$ is bounded.

Proof Let $X, Y \in \mathfrak{g}$. By Proposition 5.3.5 and by continuity of the multiplication mapping, there exists $\varepsilon > 0$ and a unique smooth curve $Z : (-\varepsilon, \varepsilon) \to \mathfrak{g}$ such that for all $|t| < \varepsilon$ one has $\exp(Z(t)) = \exp(tX)\exp(tY)$. Then, for all $f \in C^\infty(G)$ and for all $|t| < \varepsilon$, we have

$$f\big(\exp\{Z(t)\}\big) = f\big(\exp(tX)\exp(tY)\big). \tag{5.3.8}$$

We consider the Taylor expansion of both sides up to second order. On the one hand, Taylor expansion of $Z(t)$ at $t = 0$ yields $Z(t) = tZ_1 + \frac{t^2}{2}Z_2 + O(t^3)$ with $Z_1 = \dot{Z}(0)$ and $Z_2 = \ddot{Z}(0)$. Writing

$$\begin{aligned} f\big(\exp\{Z(t)\}\big) &= f\left(\exp\left\{t\left(Z_1 + \frac{s}{2}Z_2 + O\big(s^2\big)\right)\right\}\right)\Big|_{s=t} \\ &= f\big(\Phi_t^{Z_1 + \frac{s}{2}Z_2 + O(s^2)}(\mathbb{1})\big)\big|_{s=t} \end{aligned}$$

and applying the Taylor formula (3.2.13) for manifolds, we obtain

$$f\big(\exp\{Z(t)\}\big) = f(\mathbb{1}) + t\big(Z_1(f)\big)(\mathbb{1}) + \frac{t^2}{2}\big(\big(Z_2 + Z_1^2\big)(f)\big)(\mathbb{1}) + O\big(t^3\big).$$

On the other hand, by (5.3.5) and the iterated Taylor formula (3.2.14) for manifolds,

$$f\big(\exp(tX)\exp(tY)\big) = f(\mathbb{1}) + t\big\{(X+Y)(f)\big\} + \frac{t^2}{2}\big\{\big(X^2+2XY+Y^2\big)(f)\big\}(\mathbb{1}) + O\big(t^3\big).$$

We read off $Z_1 = X + Y$ and $Z_2 + Z_1^2 = X^2 + 2XY + Y^2$. Thus, $Z_2 = [X, Y]$. □

Remark 5.3.12

1. The Taylor formula (3.2.13) for manifolds and Formula (5.3.4) imply that for every $a \in G$, $f \in C^\infty(G)$ and $t \in \mathbb{R}$ there holds

$$f\big(a\exp(tX)\big) = \sum_{k=0}^{n} \frac{t^k}{k!}\big(X^k(f)\big)(a) + O\big(t^{n+1}\big). \tag{5.3.9}$$

   The case $n = 2$ and $a = \mathbb{1}$ was used in the proof of Proposition 5.3.11. In addition, recall from Remark 5.1.2/3 that the smooth structure of G contains a real analytic structure. Hence, it makes sense to speak of real analytic functions on G. For such a function one can take the limit $n \to \infty$ in (5.3.9), thus obtaining an absolutely convergent series for small t and hence, by absorbing t in X, an absolutely convergent series for small X:

$$f\big(a\exp(X)\big) = \sum_{k=0}^{\infty} \frac{1}{k!}\big(X^k(f)\big)(a). \tag{5.3.10}$$

   This is the Taylor series for real analytic functions on G. Like for the Taylor formula (3.2.13) for manifolds, by repeated application of (5.3.10) one may produce iterated versions analogous to (3.2.14).
2. Similarly to the argument in the proof, Taylor expansion of (5.3.8) and comparison of coefficients allows to successively calculate the Taylor expansion of $Z(t)$ to arbitrary order. Absorbing the parameter t in X and Y, from this expansion one obtains a formal series $Z(X, Y)$, known as the Baker-Campbell-Hausdorff series of $\mathfrak{g}$. Using the notation $\mathrm{ad}(X)$ for the mapping $\mathfrak{g} \to \mathfrak{g}$ given by[7] $\mathrm{ad}(X)Y = [X, Y]$, this series can be written in the form

$$Z(X,Y) = X + \sum_{\substack{r,s\geq 0\\ k_i+l_i>0}} (-1)^r \frac{\mathrm{ad}(X)^{k_1}\mathrm{ad}(Y)^{l_1}\ldots\mathrm{ad}(X)^{k_r}\mathrm{ad}(Y)^{l_r}\mathrm{ad}(X)^s Y}{(r+1)(s+1+\sum_{i=1}^r k_i+l_i)k_1!l_1!\ldots k_r!l_r!s!} \tag{5.3.11}$$

$$= X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}\big(\big[X,[X,Y]\big] + \big[Y,[Y,X]\big]\big) + \cdots \tag{5.3.12}$$

[7] The notation ad will be explained in the next section.

(Exercise 5.3.3). It can be shown that $Z(X, Y)$ converges on some neighbourhood of the origin in $\mathfrak{g}$ and that it satisfies

$$\exp(X)\exp(Y) = \exp\big(Z(X, Y)\big),$$

see for example [297, §2.15] or [129, §I.4]. These references also contain an alternative derivation of Formula (5.3.11).

Proposition 5.3.11 allows to express the sum and the commutator in $\mathfrak{g}$ in terms of the multiplication in G. Indeed, combining this proposition with repeated application of (5.3.3), we get

Corollary 5.3.13 *For all* $X, Y \in \mathfrak{g}$,

$$\exp(X + Y) = \lim_{n\to\infty}\left(\exp\left(\frac{1}{n}X\right)\exp\left(\frac{1}{n}Y\right)\right)^n, \tag{5.3.13}$$

$$\exp\big([X, Y]\big) = \lim_{n\to\infty}\left(\exp\left(-\frac{1}{n}X\right)\exp\left(-\frac{1}{n}Y\right)\exp\left(\frac{1}{n}X\right)\exp\left(\frac{1}{n}Y\right)\right)^{n^2}. \tag{5.3.14}$$

Example 5.3.14 (Classical Lie groups) Let V be a finite-dimensional $\mathbb{K}$-vector space and let $G \subset \mathrm{End}(V)$ be a classical Lie group. Under the identification of the Lie algebra $\mathfrak{g}$ of G with the Lie subalgebra $\mathrm{T}_{\mathbb{1}}G$ of $\mathfrak{gl}(V)$ according to Example 5.2.6, the flow of the left-invariant vector field corresponding to $A \in \mathfrak{g}$ is given by (5.2.4). Hence,

$$\exp_G(A) = \mathrm{e}^A, \quad A \in \mathfrak{g}. \tag{5.3.15}$$

That is, for the classical groups, the exponential mapping is given by the restriction of the exponential series on $\mathrm{End}(V)$ to arguments from $\mathfrak{g}$. This explains the name. For $G = \mathbb{K} \setminus \{0\}$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and $G = \mathrm{U}(1)$, this yields, in particular, the ordinary exponential functions $\mathbb{K} \to \mathbb{K} \setminus \{0\}$ and $\mathrm{i}\mathbb{R} \to \mathrm{U}(1) \subset \mathbb{C}$, respectively.

Example 5.3.15 (Vector groups) Let V be a $\mathbb{K}$-vector space, viewed as a Lie group. Under the identification of the Lie algebra of V with the vector space V, cf. Example 5.2.10, the equation for the integral curves of the left-invariant vector field corresponding to $X \in V$ is given by $\dot\gamma(t) = X$. The solution with initial condition $\gamma(0) = v \in V$ is $\gamma(t) = v + tX$. Hence, the exponential mapping of V coincides with the identical mapping id_V.

Example 5.3.16 (Direct product) Let G_1 and G_2 be Lie groups with Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, respectively. Consider the direct product Lie group $G := G_1 \times G_2$ and let $\mathrm{pr}_i : G \to G_i$ denote the natural projections. Under the natural identification of the Lie algebra of G with the direct product Lie algebra $\mathfrak{g}_1 \oplus \mathfrak{g}_2$, cf. Example 5.2.16, left-invariant vector fields on G are given by pairs (X_1, X_2) with $X_i \in \mathfrak{g}_i$ and there holds $\mathrm{d}\,\mathrm{pr}_i(X_1, X_2) = X_i$. Hence, Proposition 5.3.6 implies

$$\begin{aligned}\exp_G(X_1, X_2) &= \big(\mathrm{pr}_1 \circ \exp_G(X_1, X_2), \mathrm{pr}_2 \circ \exp_G(X_1, X_2)\big)\\ &= \big(\exp_{G_1}(X_1), \exp_{G_2}(X_2)\big).\end{aligned}$$

Example 5.3.17 (Determinant and trace) Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and let us consider the determinant homomorphism $\det : \mathrm{GL}(n, \mathbb{K}) \to \mathbb{K} \setminus \{0\}$. According to Example 5.3.14, for $A \in \mathfrak{gl}(n, \mathbb{K})$ we have $\exp(A) = \mathrm{e}^A$. Hence, (5.2.10) and Proposition 5.3.6 imply

$$\det\big(\mathrm{e}^A\big) = \mathrm{e}^{\operatorname{tr} A}. \tag{5.3.16}$$

Let us mention that this equality can also be proved by more elementary arguments which do not make use of the theory of Lie groups. One way consists in replacing A by tA and showing by direct computation that both sides of (5.3.16) satisfy the same ordinary differential equation and the same initial condition. Another way is to use the Jordan normal form of A. We leave the details to the reader (Exercise 5.3.4).

Remark 5.3.18 Since $\exp(\mathfrak{g})$ contains $\mathbb{1}$ and since it is connected,[8] it is contained in the identity connected component G_0. For $G = \mathbb{R} \setminus \{0\}$, the exponential mapping is a global diffeomorphism from $\mathbb{R}$ onto $G_0 = \mathbb{R}_+$. For the vector groups, it is trivially a diffeomorphism. In general, however, exp is neither injective nor surjective onto G_0. For example, in the case $G = \mathrm{GL}(n, \mathbb{C})$, it is surjective but not injective, whereas in the case $G = \mathrm{GL}(n, \mathbb{R})$, $n \geq 2$, it is neither injective nor surjective onto G_0, see Exercise 5.3.5. As a general fact, exp is surjective onto G_0 if G_0 is compact or Abelian. In the latter case, this can be seen as follows. As a consequence of (3.2.12) and (5.3.5), the Lie algebra of G is Abelian. According to Propositions 5.3.5 and 5.3.10, $\exp(\mathfrak{g})$ is a subgroup of G_0 which contains a neighbourhood of $\mathbb{1}$. According to Proposition 5.1.7, this implies $\exp(\mathfrak{g}) = G_0$.

Exercises

5.3.1 Show that every one-parameter subgroup of a Lie group is the maximal integral curve through the unit element of a unique left-invariant vector field, cf. Remark 5.3.2.

5.3.2 Repeated application of (5.3.7) yields

$$\exp(-tX)\exp(-tY)\exp(tX)\exp(tY) = \exp\big(t^2[X, Y] + O\big(t^3\big)\big)$$
for all $X, Y \in \mathfrak{g}$.

Use this to give a geometric interpretation of the commutator of left-invariant vector fields.

[8] As the image of a connected set under a continuous mapping.

5.3.3 Calculate the first three terms of the Baker-Campbell-Hausdorff series given in (5.3.12) by Taylor expanding (5.3.8) to third order and absorbing the parameter t in the vector fields X and Y. After this warm-up, prove the general formula (5.3.11) by induction on n.

5.3.4 Carry out the proofs of the identity (5.3.16) indicated in Example 5.3.17.

5.3.5 Show that the matrix

$$a = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$$

is contained in the identity component of $\mathrm{GL}(2, \mathbb{R})$ but not in the image of the exponential mapping.

Hint. Assume that $a = \mathrm{e}^A$ for some $A \in \mathfrak{gl}(2, \mathbb{R})$. Use (5.3.16) to show that A has eigenvalues λ and $-\lambda$. Use the Jordan normal form of A to show that, then, a has eigenvalues e^λ and $\mathrm{e}^{-\lambda}$. Deduce from this that A and hence a is diagonalizable (contradiction).

5.3.6 Show that the tangent mapping of exp at $X \in \mathfrak{g}$ is given by

$$(L_{\exp(-X)})'_{\mathbb{1}} \circ \exp'_X = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} \big(\mathrm{ad}(-X)\big)^n$$

and that $\exp'_X$ is bijective if and only if no eigenvalue of $\mathrm{ad}(X)$ is an integer multiple of $2\pi\mathrm{i}$.

5.4 Adjoint Representation and Killing Form

Let $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{H}$ and let V be a finite-dimensional $\mathbb{K}$-vector space.

Definition 5.4.1 (Representation)

1. A representation of a Lie group G on V is a Lie group homomorphism $\varphi : G \to \mathrm{GL}(V)$. A homomorphism of representations $\varphi_i : G \to \mathrm{GL}(V_i)$, $i = 1, 2$, is a linear mapping $\lambda : V_1 \to V_2$ such that $\lambda \circ \varphi_1(a) = \varphi_2(a) \circ \lambda$ for all $a \in G$.
2. A representation of a Lie algebra $\mathfrak{g}$ on V is a Lie algebra homomorphism $\varphi : \mathfrak{g} \to \mathfrak{gl}(V)$. A homomorphism of representations $\varphi_i : \mathfrak{g} \to \mathfrak{gl}(V_i)$, $i = 1, 2$, is a linear mapping $\lambda : V_1 \to V_2$ such that $\lambda \circ \varphi_1(A) = \varphi_2(A) \circ \lambda$ for all $A \in \mathfrak{g}$.

Homomorphisms of representations are also referred to as intertwining operators.

Every Lie group G possesses a natural representation on its Lie algebra $\mathfrak{g}$, constructed as follows. For every $a \in G$, the conjugation mapping C_a is a Lie group automorphism of G. According to Remark 5.2.14/2, the induced homomorphism dC_a of $\mathfrak{g}$ is an automorphism and hence, in particular, a vector space automorphism of $\mathfrak{g}$. Thus, we obtain a mapping

$$\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g}), \qquad \mathrm{Ad}(a) := \mathrm{dC}_a .$$

According to (5.2.9), Ad is a group homomorphism. To see that it is smooth, consider the chart on $\mathrm{GL}(\mathfrak{g})$ given by the matrix representation induced by a basis in $\mathfrak{g}$. For Ad to be smooth with respect to this chart it suffices that the mapping $G \to \mathfrak{g}$ given by $a \mapsto \mathrm{Ad}(a)X$ be smooth for all $X \in \mathfrak{g}$. This mapping can be written as the composition

$$G \to \mathrm{T}(G \times G) \to \mathrm{T}G \to G \times \mathfrak{g} \to \mathfrak{g}.$$

Here, the first mapping assigns to $a \in G$ the tangent vector at $(a, \mathbb{1})$ corresponding to $(0, X_{\mathbb{1}})$ under the natural identification of $\mathrm{T}(G \times G)$ with $\mathrm{T}G \times \mathrm{T}G$, the second mapping is the tangent mapping of $(a, b) \mapsto \mathrm{C}_a(b)$ and the last two mappings are the left trivialization and the natural projection onto the second factor, respectively. Thus, Ad is a Lie group homomorphism and hence a representation of G on $\mathfrak{g}$.

Next, we note that every Lie algebra possesses a natural representation on itself, given by[9]

$$\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}), \qquad \mathrm{ad}(X)Y := [X, Y] \quad \text{for all } Y \in \mathfrak{g}.$$

Indeed, the Jacobi identity implies

$$\mathrm{ad}\big([X, Y]\big) = \big[\mathrm{ad}(X), \mathrm{ad}(Y)\big],$$

hence ad is a Lie algebra homomorphism. Since $\mathrm{Ad}(a)$ is a Lie algebra automorphism for all $a \in G$, we obtain

$$\mathrm{ad}\big(\mathrm{Ad}(a)X\big) = \mathrm{Ad}(a) \circ \mathrm{ad}(X) \circ \mathrm{Ad}\big(a^{-1}\big) \quad \text{for all } a \in G \text{ and } X \in \mathfrak{g}. \tag{5.4.1}$$

Definition 5.4.2 (Adjoint representations) The representations Ad and ad are called the adjoint representations of G and $\mathfrak{g}$, respectively.

Every representation φ of G on V induces a representation of $\mathfrak{g}$ on V, given by the induced homomorphism $\mathrm{d}\varphi$. Under the natural identification of the Lie algebra of $\mathrm{GL}(V)$ with $\mathfrak{gl}(V)$,

$$\mathrm{d}\varphi(X) = \varphi'_{\mathbb{1}}(X_{\mathbb{1}}) = \frac{\mathrm{d}}{\mathrm{d}t}\Big\restriction_0 \varphi\big(\exp(tX)\big), \quad X \in \mathfrak{g}. \tag{5.4.2}$$

Proposition 5.4.3 *The adjoint representations fulfil the relation*

$$\mathrm{d}\,\mathrm{Ad} = \mathrm{ad}.$$

Proof Let $X, Y \in \mathfrak{g}$. Due to Remark 5.2.14/2, left-invariance and (5.3.4), we get

$$\mathrm{Ad}\big(\exp(tX)\big)Y = \mathrm{C}_{\exp(tX)*}Y = \mathrm{R}_{\exp(-tX)*} \circ \mathrm{L}_{\exp(tX)*}Y = \Phi^X_{-t*}Y.$$

Hence, (5.4.2) yields $\mathrm{d}\,\mathrm{Ad}(X)Y = \mathscr{L}_X Y = [X, Y] = \mathrm{ad}(X)Y$. □

[9] The notation ad has already been used in Formula (5.3.11).

Let φ be a representation of G on the finite-dimensional $\mathbb{K}$-vector space V. Since the mapping $\mathrm{End}(V) \to \mathrm{End}(V^*)$ which assigns to A the dual endomorphism A^{T} is linear and hence smooth, φ induces a representation

$$\varphi^* : G \to \mathrm{GL}\big(V^*\big), \qquad \varphi^*(a) := \varphi\big(a^{-1}\big)^{\mathrm{T}}, \tag{5.4.3}$$

called the dual representation of φ. The representation $\mathrm{d}(\varphi^*) : \mathfrak{g} \to \mathfrak{gl}(V^*)$, induced on $\mathfrak{g}$, is called the dual representation of $\mathrm{d}\varphi$. A brief calculation shows

$$\mathrm{d}\big(\varphi^*\big)(X) = -\big(\mathrm{d}\varphi(X)\big)^{\mathrm{T}} \quad \text{for all } X \in \mathfrak{g}. \tag{5.4.4}$$

Definition 5.4.4 (Coadjoint representation) The duals of the adjoint representations of G and $\mathfrak{g}$ are called the coadjoint representations of G and $\mathfrak{g}$ and are denoted by Ad^* and ad^*, respectively.

By definition,

$$\big\langle \mathrm{Ad}^*(a)\xi, Y\big\rangle = \big\langle \xi, \mathrm{Ad}\big(a^{-1}\big)Y\big\rangle \quad \text{for all } a \in G,\ Y \in \mathfrak{g} \text{ and } \xi \in \mathfrak{g}^*. \tag{5.4.5}$$

According to Proposition 5.4.3 and (5.4.4),

$$\big\langle \mathrm{ad}^*(X)\xi, Y\big\rangle = -\big\langle \xi, [X, Y]\big\rangle \quad \text{for all } X, Y \in \mathfrak{g} \text{ and } \xi \in \mathfrak{g}^*. \tag{5.4.6}$$

Next, recall that the kernel of a group homomorphism is the preimage of the unit element of the range and that the kernel of a linear mapping between vector spaces is the preimage of the origin of the range. Recall further that

(a) the centralizer $C_G(H)$ of a subset H of a group G consists of those elements of $a \in G$ satisfying $ah = ha$ for all $h \in H$, and that the centre of G is the centralizer of $H = G$,
(b) the centralizer of a subset $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ consists of those elements $X \in \mathfrak{g}$ satisfying $[X, Y] = 0$ for all $Y \in \mathfrak{h}$, and that the centre of $\mathfrak{g}$ is the centralizer of $\mathfrak{h} = \mathfrak{g}$.

Proposition 5.4.5 *The kernels of* Ad *and* Ad* *coincide with the centralizer* $C_G(G_0)$ *of the identity component* G_0 *in* G. *The kernels of* ad *and* ad* *coincide with the centre* $\mathfrak{z}$ *of* $\mathfrak{g}$.

Proof It is easy to see that $\ker(\mathrm{Ad}^*) = \ker(\mathrm{Ad})$ and $\ker(\mathrm{ad}^*) = \ker(\mathrm{ad})$. Moreover, $\ker(\mathrm{ad}) = \mathfrak{z}$ is obvious. Hence, it suffices to show $\ker(\mathrm{Ad}) = C_G(G_0)$. Since for $a \in C_G(G_0)$ there holds $\mathrm{C}_{a \restriction G_0} = \mathrm{id}_{G_0}$ and since $\mathrm{d}\,\mathrm{id}_{G_0} = \mathrm{id}_{\mathfrak{g}}$, we obtain $\ker(\mathrm{Ad}) \supset C_G(G_0)$. Conversely, if $a \in G$ satisfies $\mathrm{Ad}(a) = \mathrm{id}_{\mathfrak{g}}$, Proposition 5.3.6 implies

$$\exp(X) = \exp\big(\mathrm{Ad}(a)X\big) = \mathrm{C}_a\big(\exp(X)\big) \quad \text{for all } X \in \mathfrak{g}.$$

That is, a commutes with all elements of $\exp(\mathfrak{g})$. Since $\exp(\mathfrak{g})$ contains a neighbourhood of $\mathbb{1}$ in G_0, Proposition 5.1.7 yields $\ker(\mathrm{Ad}) \subset C_G(G_0)$. □

Proposition 5.4.5 implies that if G is Abelian, then $\mathrm{Ad}(a) = \mathrm{id}_{\mathfrak{g}}$ for all $a \in G$ and $\mathrm{ad}(X) = 0$ for all $X \in \mathfrak{g}$. This applies in particular to vector groups and to tori.

Example 5.4.6 (Classical Lie groups) Let V be a finite-dimensional $\mathbb{K}$-vector space and let $G \subset \mathrm{End}(V)$ be a classical Lie group. Let $a \in G$. Under the identification of the Lie algebra $\mathfrak{g}$ with the Lie subalgebra $\mathrm{T}_{\mathbb{1}}G$ of $\mathfrak{gl}(V)$, $\mathrm{Ad}(a) \equiv \mathrm{dC}_a$ coincides with the tangent mapping $(\mathrm{C}_a)'_{\mathbb{1}}$. For $A \in \mathfrak{g}$,

$$(\mathrm{C}_a)'_{\mathbb{1}} A = \frac{\mathrm{d}}{\mathrm{d}t}{\restriction_0} \mathrm{C}_a\big(\mathrm{e}^{tA}\big) = \frac{\mathrm{d}}{\mathrm{d}t}{\restriction_0} \big(a\mathrm{e}^{tA}a^{-1}\big) = aAa^{-1}.$$

Hence,

$$\mathrm{Ad}(a)A = aAa^{-1}, \quad a \in G, \ A \in \mathfrak{g}. \tag{5.4.7}$$

Let us mention that in this example the proof of Proposition 5.4.3 reduces to the following calculation:

$$\mathrm{d}\,\mathrm{Ad}(A)B \equiv \mathrm{Ad}'_{\mathbb{1}}(A)B = \frac{\mathrm{d}}{\mathrm{d}t}{\restriction_0} \mathrm{Ad}\big(\mathrm{e}^{tA}\big)B = \frac{\mathrm{d}}{\mathrm{d}t}{\restriction_0} \big(\mathrm{e}^{tA}B\mathrm{e}^{-tA}\big) = [A, B].$$

Example 5.4.7 Recall the Lie algebra isomorphism $\varphi : \mathbb{R}^3 \to \mathfrak{so}(3)$ from Example 5.2.8. It satisfies

$$\varphi(a\mathbf{x}) = a\varphi(\mathbf{x})a^{\mathrm{T}}, \quad a \in \mathrm{SO}(3), \ \mathbf{x} \in \mathbb{R}^3, \tag{5.4.8}$$

(Exercise 5.4.1). In view of (5.4.7) and $a^{\mathrm{T}} = a^{-1}$, this implies that φ identifies the adjoint representation of SO(3) with the identical representation on $\mathbb{R}^3$. Similarly, the Lie algebra isomorphism $\mathbb{R}^3 \to \mathfrak{sp}(1)$ of this example, also denoted by φ, satisfies

$$\varphi\big(\phi(\mathbf{a})\mathbf{x}\big) = \mathbf{a}\varphi(\mathbf{x})\bar{\mathbf{a}}, \quad \mathbf{a} \in \mathrm{Sp}(1), \ \mathbf{x} \in \mathbb{R}^3,$$

where $\phi : \mathrm{Sp}(1) \to \mathrm{SO}(3)$ is the covering homomorphism of Example 5.1.11. This follows from (5.4.8), because $\phi(\overline{\mathbf{q}}) = \phi(\mathbf{q})^{\mathrm{T}}$. As a consequence, the adjoint representation of Sp(1) can be identified with the representation of Sp(1) on $\mathbb{R}^3$ induced by ϕ. The same statement holds for SU(2). Finally, (5.2.11) implies that the Lie algebra isomorphism $\mathfrak{sp}(1) \oplus \mathfrak{sp}(1) \to \mathfrak{so}(4)$ induced by the covering homomorphism $\mathrm{Sp}(1) \times \mathrm{Sp}(1) \to \mathrm{SO}(4)$ of Example 5.1.11 identifies the representation of $\mathrm{Sp}(1) \times \mathrm{Sp}(1)$ on $\mathfrak{so}(4)$, induced via this covering homomorphism by the adjoint representation of SO(4), with the adjoint representation of $\mathrm{Sp}(1) \times \mathrm{Sp}(1)$.

Example 5.4.8 (Direct product) Let G_1 and G_2 be Lie groups with Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, respectively, and let $G = G_1 \times G_2$. Under the natural identification of $\mathfrak{g}$ with $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ of Example 5.2.16, Eq. (5.2.12) implies

$$\mathrm{Ad}_G\big((a_1, a_2)\big)(Y_1, Y_2) = \big(\mathrm{Ad}_{G_1}(a_1)Y_1, \mathrm{Ad}_{G_2}(a_2)Y_2\big), \quad a_i \in G_i, \ Y_i \in \mathfrak{g}_i.$$

An analogous formula holds for ad.

Using the adjoint representation of $\mathfrak{g}$, one can construct a natural symmetric bilinear form on $\mathfrak{g}$ which is invariant under the adjoint representation of G. Recall from linear algebra that for an abstract $\mathbb{K}$-vector space V, the trace of an endomorphism of V is defined to be the trace of the corresponding matrix with respect to an arbitrary basis in V or, equivalently, to be the sum of eigenvalues, counted with multiplicities. The trace is a linear functional on $\mathrm{End}(V)$ satisfying

$$\mathrm{tr}(AB) = \mathrm{tr}(BA) \quad \text{for all } A, B \in \mathrm{End}(V). \tag{5.4.9}$$

Definition 5.4.9 (Killing form) The Killing form of a finite-dimensional Lie algebra $\mathfrak{g}$ is the bilinear form

$$\mathsf{k}\colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}, \quad \mathsf{k}(X, Y) := \mathrm{tr}\big(\mathrm{ad}(X)\mathrm{ad}(Y)\big).$$

Proposition 5.4.10 (Properties of the Killing form) *Let G be a Lie group and let $\mathfrak{g}$ be the Lie algebra of G. The Killing form k of $\mathfrak{g}$ is symmetric*, Ad-*invariant*,

$$\mathsf{k}\big(\mathrm{Ad}(a)X, \mathrm{Ad}(a)Y\big) = \mathsf{k}(X, Y), \tag{5.4.10}$$

and satisfies

$$\mathsf{k}\big(\mathrm{ad}(Z)X, Y\big) + \mathsf{k}\big(X, \mathrm{ad}(Z)Y\big) = 0. \tag{5.4.11}$$

It is non-degenerate iff $\mathfrak{g}$ (and hence G) is semisimple.

Proof The symmetry property and Eq. (5.4.11) are due to (5.4.9). By the help of (5.4.9), Eq. (5.4.10) follows from (5.4.1). To prove the last assertion, define $\mathfrak{g}^\perp := \{X \in \mathfrak{g} : \mathsf{k}(X, \mathfrak{g}) = \{0\}\}$. Since every Abelian ideal of $\mathfrak{g}$ is contained in $\mathfrak{g}^\perp$, if k is non-degenerate, $\mathfrak{g}$ is semisimple. Conversely, assume that $\mathfrak{g}$ is semisimple. As a consequence of (5.4.11), $\mathfrak{g}^\perp$ is an ideal in $\mathfrak{g}$. This implies that the Killing form of the Lie algebra $\mathfrak{g}^\perp$ is the restriction to $\mathfrak{g}^\perp$ of the Killing form of $\mathfrak{g}$ (Exercise 5.4.2) and is hence trivial. Now, Cartan's criterion[10] yields that $\mathfrak{g}^\perp$ is a solvable ideal in $\mathfrak{g}$. Hence, if it was nonzero, it would contain a nonzero Abelian ideal. Therefore, $\mathfrak{g}^\perp = \{0\}$. □

Remark 5.4.11

1. Computation of the trace by means of a pair of dual bases $\{e_i\}$ in $\mathfrak{g}$ and $\{e^{*i}\}$ in $\mathfrak{g}^*$ yields

   $$\mathsf{k}(X, Y) = \big\langle e^{*i}, \big[X, [Y, e_i]\big]\big\rangle = X^k Y^l {c_{li}}^j {c_{kj}}^i, \tag{5.4.12}$$

   where X^k and Y^l are the expansion coefficients of X and Y, respectively, with respect to the basis $\{e_i\}$ and ${c_{kj}}^i$ are the corresponding structure constants. The

[10] A Lie algebra $\mathfrak{g}$ whose Killing form k satisfies $\mathsf{k}(X, Y) = 0$ for all $X \in \mathfrak{g}$, $Y \in [\mathfrak{g}, \mathfrak{g}]$ is solvable [145, §4.3].

symmetric covariant tensor of order 2

$$g_{kl} = c_{li}{}^{j} c_{kj}{}^{i}$$

is referred to as the Cartan-Killing tensor of G with respect to the basis $\{e_i\}$.

2. According to Proposition 5.4.10, if G is semisimple, the Killing form is non-degenerate and hence induces a linear isomorphism $F : \mathfrak{g} \to \mathfrak{g}^*$ defined by $\langle F(X), Y\rangle = \mathsf{k}(X, Y)$ for all $X, Y \in \mathfrak{g}$. Invariance under Ad implies

$$F \circ \mathrm{Ad}(a) = \mathrm{Ad}^*(a) \circ F \tag{5.4.13}$$

for all $a \in G$. Thus, F is an isomorphism of the representations Ad and Ad* of G. As a consequence, for semisimple Lie groups, the adjoint and coadjoint representations can be identified.

Example 5.4.12 For each of the semisimple classical Lie groups of Example 1.2.6, there exists $c > 0$ such that

$$\mathsf{k}(X, Y) = c\,\mathrm{tr}(XY) \quad \text{for all } X, Y \in \mathfrak{g}.$$

For example, for $\mathfrak{sl}(n, \mathbb{K})$ and $\mathfrak{su}(n)$, $n \geq 2$, the factor is $c = 2n$, for $\mathfrak{so}(n)$, $n \geq 3$, it is $c = n - 2$ and for $\mathfrak{sp}(n)$, $n \geq 1$, it is $c = 2(n + 1)$ (Exercise 5.4.3).

Exercises

5.4.1 Prove the two formulae stated in Example 5.4.7.

5.4.2 Let $\mathfrak{g}$ be a Lie algebra, let $\mathfrak{i}$ be an ideal in $\mathfrak{g}$ and let $\mathsf{k}_{\mathfrak{g}}$ and $\mathsf{k}_{\mathfrak{i}}$ denote the respective Killing forms. Show that for all $X, Y \in \mathfrak{i}$ there holds $\mathsf{k}_{\mathfrak{i}}(X, Y) = \mathsf{k}_{\mathfrak{g}}(X, Y)$.

5.4.3 Determine the factor of proportionality between the Killing form and the trace form of the identical representation for the classical groups listed in Remark 5.4.12.

5.5 Left-Invariant Differential Forms

Definition 5.5.1 A differential form ξ on G is called left-invariant if $\mathrm{L}_a^* \xi = \xi$ for all $a \in G$.

The set of left-invariant differential forms on G will be denoted by $\Omega^*(G)^G$. Written pointwise, for 1-forms, the defining condition reads $\xi_{ab} \circ (\mathrm{L}_a)'_b = \xi_b$ for all $a, b \in G$ or, equivalently,

$$\xi_a = \xi_{\mathbb{1}} \circ (\mathrm{L}_{a^{-1}})'_a \quad \text{for all } a \in G. \tag{5.5.1}$$

Proposition 5.5.2 *Let G be a Lie group with Lie algebra $\mathfrak{g}$.*

1. *$\Omega^*(G)^G$ is a differential subalgebra of $\Omega^*(G)$.*
2. *There exist natural algebra isomorphisms*

$$\Omega^*(G)^G \cong \bigwedge \mathrm{T}^*_{\mathbb{1}} G \cong \bigwedge \mathfrak{g}^*.$$

The first one is given by $\xi \mapsto \xi_{\mathbb{1}}$. The second one is induced by the isomorphism $\mathfrak{g} \to \mathrm{T}_{\mathbb{1}} G$, given by $X \mapsto X_{\mathbb{1}}$.
3. *For all $\xi \in \Omega^r(G)^G$ and $X_1, \ldots, X_r \in \mathfrak{g}$, the function $\xi(X_1, \ldots, X_r)$ on G is constant.*

Proof 1. By (2.4.24), $\Omega^*(G)^G$ is closed under exterior multiplication. By (4.1.23), it is closed under taking exterior derivatives.

2. It suffices to consider the mapping $\Omega^*(G)^G \to \bigwedge \mathrm{T}^*_{\mathbb{1}} G$, defined by $\xi \mapsto \xi_{\mathbb{1}}$. This mapping is a homomorphism of algebras, which is injective by (5.5.1). To show that it is surjective, it is enough to prove that its image contains $\mathrm{T}^*_{\mathbb{1}} G$. Thus, let $\eta \in \mathrm{T}^*_{\mathbb{1}} G$. By fixing the second argument of the inverse left trivialization of $\mathrm{T}^* G$ to be η, cf. (5.1.6), one obtains a left-invariant 1-form ξ with $\xi_{\mathbb{1}} = \eta$.

3. This is a consequence of (4.1.16). □

According to assertion 2 of Proposition 5.5.2, we will identify left-invariant differential r-forms with elements of $\bigwedge^r \mathfrak{g}^*$ without explicitly stating that. Depending on the context, for $\xi \in \bigwedge^r \mathfrak{g}^*$ and $X_1, \ldots, X_r \in \mathfrak{g}$, the expression $\xi(X_1, \ldots, \xi_r)$ will be interpreted as a function on G or as a number.

Remark 5.5.3

1. Propositions 4.1.6 and 5.5.2 imply that the exterior derivative of $\xi \in \bigwedge^r \mathfrak{g}^*$ is given by[11]

$$\mathrm{d}\xi(X_0, \ldots, X_r) = \sum_{i<j} (-1)^{i+j} \xi\big([X_i, X_j], X_0, \overset{X_i X_j}{\check{\ldots}\check{}}, X_r\big), \tag{5.5.2}$$

where $X_0, \ldots, X_r \in \mathfrak{g}$. The right hand side may be taken as an intrinsic definition of a differential on the exterior algebra $\bigwedge \mathfrak{g}^*$, thus turning the natural isomorphism of algebras $\Omega^*(G)^G \cong \bigwedge \mathfrak{g}^*$ of Proposition 5.5.2 into an isomorphism of differential algebras. In particular, this yields a characterization of the subcomplex $\Omega^*(G)^G$ of the de Rham complex of G in terms of the Lie algebra alone.
2. For $\xi \in \mathfrak{g}^*$, (5.5.2) yields

$$\mathrm{d}\xi(X, Y) = -\xi\big([X, Y]\big), \quad X, Y \in \mathfrak{g}. \tag{5.5.3}$$

[11] Cf. Proposition 4.1.6 for the notation.

Let $\{e_1, \ldots, e_n\}$ be a basis in $\mathfrak{g}$, let c^i_{jk} be the corresponding structure constants and let $\{e^{*1}, \ldots, e^{*n}\}$ be the dual basis in $\mathfrak{g}^*$. Then, (5.5.3) is equivalent to

$$\mathrm{d}e^{*i} = -\frac{1}{2} c^i_{jk} e^{*j} \wedge e^{*k}, \quad i = 1, \ldots, n. \tag{5.5.4}$$

This can be seen by evaluating both sides on the basis elements e_i (Exercise 5.5.1). Equation (5.5.4) is known as the Maurer-Cartan equation associated with the basis $\{e_1, \ldots, e_n\}$ of $\mathfrak{g}$.

3. In terms of left-invariant 1-forms, the inverse left and right trivializations $(\chi^{\mathrm{T}}_{\mathrm{L}})^{-1}$ and $(\chi^{\mathrm{T}}_{\mathrm{R}})^{-1}$ of T^*G, given by (5.1.6), read

$$\left(\chi^{\mathrm{T}}_{\mathrm{L}}\right)^{-1}(a, \xi) = \xi_a, \qquad \left(\chi^{\mathrm{T}}_{\mathrm{R}}\right)^{-1}(a, \xi) = \left(\mathrm{C}^*_a \xi\right)_a \equiv \xi_a \circ (\mathrm{C}_a)'_a.$$

Note that the 1-form $\mathrm{C}^*_a \xi$ need not be left-invariant.

Proposition 5.5.2 yields in particular that the space of left-invariant n-forms corresponds to $\bigwedge^n \mathfrak{g}^*$. Hence, it has dimension one and its elements are of the form $e^{*1} \wedge \cdots \wedge e^{*n}$ for some basis $\{e^{*i}\}$ in $\mathfrak{g}^*$. By left-invariance, every nonzero element of this space is a volume form. Hence, we obtain

Corollary 5.5.4 *On every Lie group there exists a left-invariant volume form v_G. This form is unique up to multiplication by a nonzero real number.*

It is common to write $\mathsf{v}_G(a) = \mathrm{d}a$. The Lebesgue measure associated with a left-invariant volume form on G is called a Haar measure on G. Thus, Haar measures on G are left-invariant and unique up to a constant.

Remark 5.5.5

1. Every scalar product $\langle \cdot, \cdot \rangle$ on $\mathfrak{g}$ defines a Riemannian metric g on G by

$$\mathrm{g}_a(X_a, Y_a) := \langle X, Y \rangle, \quad a \in G, \ X, Y \in \mathfrak{g}. \tag{5.5.5}$$

By construction, g is left-invariant, that is, $\mathrm{L}_{a*}\mathrm{g} = \mathrm{g}$ for all $a \in G$, and so is its volume form v_{g}. This way, the choice of a scalar product on $\mathfrak{g}$ and an orientation on G singles out a unique left-invariant volume form. Furthermore, every basis of $\mathfrak{g}$ which is orthonormal with respect to $\langle \cdot, \cdot \rangle$ provides a global frame of $\mathrm{T}G$ which is orthonormal with respect to g. By (4.4.14), in terms of the corresponding dual basis $\{e^{*i}\}$ of $\mathfrak{g}^*$, the volume form v_{g} is given by

$$\mathsf{v}_{\mathrm{g}} = e^{*1} \wedge e^{*2} \wedge \cdots \wedge e^{*n}.$$

2. If G is compact, another way to single out a unique left-invariant volume form on G is to require that the volume of G be equal to 1.

3. We determine the action of right translations on left-invariant volume forms. Recall from linear algebra that for an abstract $\mathbb{K}$-vector space V, the determinant of an endomorphism of V is defined to be the determinant of the corresponding matrix with respect to an arbitrary basis in V or, equivalently, the product of eigenvalues, counted with multiplicities. A brief computation (Exercise 5.5.2) shows that for every left-invariant volume form v_G there holds

$$\mathrm{R}_a^* \mathsf{v}_G = \det\bigl(\mathrm{Ad}\bigl(a^{-1}\bigr)\bigr)\mathsf{v}_G, \quad a \in G. \tag{5.5.6}$$

The smooth function

$$\Delta : G \to \mathbb{R}, \qquad \Delta(a) := \det\bigl(\mathrm{Ad}\bigl(a^{-1}\bigr)\bigr),$$

is called the modular function of G. For every $f \in C^\infty(G)$, integrable with respect to v_G, and for every $a \in G$, the function $\mathrm{R}_a^* f$ is integrable as well and

$$\int_G \bigl(\mathrm{R}_a^* f\bigr)\mathsf{v}_G = \Delta\bigl(a^{-1}\bigr)\int_G f\mathsf{v}_G. \tag{5.5.7}$$

For compact Lie groups, where every smooth function is integrable with respect to any volume form, the existence of left-invariant volume forms provides the powerful tool of averaging over the group. This concept can be used, for example, to construct invariants of representations, e.g. an invariant scalar product.

Proposition 5.5.6 (Invariant scalar product) *Let G be a compact Lie group, let V be a finite-dimensional $\mathbb{K}$-vector space and let φ be a representation of G on V. The vector space V admits a scalar product $\langle\cdot,\cdot\rangle$ such that*

$$\bigl\langle\varphi(a)v, \varphi(a)w\bigr\rangle = \langle v, w\rangle \quad \textit{for all } a \in G \textit{ and } v, w \in V.$$

For the induced representation $\mathrm{d}\varphi$ of $\mathfrak{g}$, we have

$$\bigl\langle \mathrm{d}\varphi(X)v, w\bigr\rangle + \bigl\langle v, \mathrm{d}\varphi(X)w\bigr\rangle = 0 \quad \textit{for all } X \in \mathfrak{g} \textit{ and } v, w \in V.$$

Thus, every finite-dimensional representation of a compact Lie group may be assumed to be orthogonal (in case $\mathbb{K} = \mathbb{R}$) or unitary (in case $\mathbb{K} = \mathbb{C}, \mathbb{H}$).

Proof Choose an arbitrary scalar product $(\cdot,\cdot)$ on V. For $v, w \in V$, define a function $f_{v,w} \in C^\infty(G)$ by $f_{v,w}(a) := (\varphi(a^{-1})v, \varphi(a^{-1})w)$. Choose a left-invariant volume form v_G on G and define

$$\langle v, w\rangle := \int_G f_{v,w}\mathsf{v}_G.$$

The integral exists, because G is compact. The defining properties of a scalar product carry over from $(\cdot,\cdot)$ to $\langle\cdot,\cdot\rangle$. Let $a \in G$. A brief computation shows

$f_{\varphi(a)v,\varphi(a)w} = \mathrm{L}^*_{a^{-1}} f_{v,w}$. Hence, by left-invariance of v_G,

$$\big\langle \varphi(a)v, \varphi(a)w \big\rangle = \int_G \big(\mathrm{L}^*_{a^{-1}} f_{v,w}\big)\mathsf{v}_G = \int_G f_{v,w}\mathsf{v}_G = \langle v, w\rangle,$$

as asserted. In view of (5.4.2), the formula for $\mathrm{d}\varphi$ follows by differentiation. □

For a vector space endomorphism A, let $\mathrm{spec}(A)$ denote the spectrum.

Corollary 5.5.7 *For a finite-dimensional $\mathbb{K}$-representation φ of a compact Lie group G,*

1. $\mathrm{spec}(\varphi(a)) \subset \mathrm{U}(1)$ *for all $a \in G$ and* $\mathrm{spec}(\mathrm{d}\varphi(X)) \subset \mathrm{i}\mathbb{R}$ *for all $X \in \mathfrak{g}$,*
2. *in case $\mathbb{K} = \mathbb{C}$, $\varphi(a)$ and $\mathrm{d}\varphi(X)$ are diagonalizable for all $a \in G$ and $X \in \mathfrak{g}$.*

Proof 1. Without loss of generality, view the vector space V carrying the representation as a real vector space. Choose a basis in V which is orthonormal with respect to some φ-invariant scalar product. With respect to this basis, $\varphi(a)$ and $\mathrm{d}\varphi(X)$ are represented by an orthogonal and a skew-symmetric matrix, respectively. Hence, the assertion follows from elementary linear algebra.

2. In case $\mathbb{K} = \mathbb{C}$, $\varphi(a)$ is unitary and $\mathrm{d}\varphi(X)$ is skew-Hermitian. Hence, this assertion follows from elementary linear algebra, too. □

Application of Corollary 5.5.7 to the adjoint representations of G and $\mathfrak{g}$ yields

Corollary 5.5.8 *Let G be a compact Lie group.*

1. *The modular function of G is given by $\Delta(a) = \pm 1$ for all $a \in G$. If G is connected, $\Delta(a) = 1$.*
2. *The Killing form k of $\mathfrak{g}$ is negative semidefinite.*
3. *If G is in addition semisimple, $-\mathsf{k}$ is an* Ad*-invariant scalar product on $\mathfrak{g}$.*

Proof 1. Let $a \in G$. On the one hand, Corollary 5.5.7/1 implies that $\Delta(a) = \det(\mathrm{Ad}(a)) \in \mathrm{U}(1)$. On the other hand, $\mathrm{Ad}(a)$ is an automorphism of the real vector space $\mathfrak{g}$, hence $\det(\mathrm{Ad}(a))$ is real. It follows that $\Delta(a) = \pm 1$. In particular, Δ is locally constant. Hence, if G is connected, $\Delta(a) = \Delta(\mathbb{1}) = 1$.

2. Let $X \in \mathfrak{g}$. By Corollary 5.5.7/1, the spectrum of $\mathrm{ad}(X)^2$ consists of the squares of certain purely imaginary numbers. Hence, $\mathsf{k}(X, X) = \mathrm{tr}(\mathrm{ad}(X)^2) \leq 0$.

3. This follows from point 2 and non-degeneracy, cf. Proposition 5.4.10. □

Recall from Remark 5.5.5/1 that every scalar product on $\mathfrak{g}$ defines a left-invariant Riemannian metric g on G via (5.5.5). If the scalar product is invariant under Ad, g is in addition right-invariant and hence bi-invariant, that is

$$\mathrm{L}_{a*}\mathrm{g} = \mathrm{R}_{a*}\mathrm{g} = \mathrm{g} \quad \text{for all } a \in G$$

(Exercise 5.5.3). Thus, Corollary 5.5.8 implies

Corollary 5.5.9 *Every compact Lie group admits a bi-invariant Riemannian metric. If the group is in addition semisimple, then the negative of the Killing form yields such a Riemannian metric.*

To conclude this section, we discuss left-invariant differential forms on G with values in $\mathfrak{g}$, that is, differential forms $\xi \in \Omega^*(G, \mathfrak{g})$ satisfying $\mathrm{L}_a^* \xi = \xi$. According to Remark 4.1.10/2, given a basis $\{e_i\}$ in $\mathfrak{g}$, every $\xi \in \Omega^r(G, \mathfrak{g})$ can be written as

$$\xi = \xi^i \otimes e_i \tag{5.5.8}$$

with ordinary differential r-forms ξ^i. The form ξ is left-invariant iff so are the ξ^i. Since $\mathfrak{g}$ is an algebra, $\Omega^*(G, \mathfrak{g})$ carries an exterior product $\Omega^{r_1}(G, \mathfrak{g}) \times \Omega^{r_2}(G, \mathfrak{g}) \to \Omega^{r_1+r_2}(G, \mathfrak{g})$, given by

$$\begin{aligned}
&[\xi_1, \xi_2](Y_1, \dots, Y_{r_1+r_2}) \\
&:= \frac{1}{r_1! r_2!} \sum_{\pi \in \mathrm{S}_{r_1+r_2}} \big[\xi_1(Y_{\pi(1)}, \dots, Y_{\pi(r_1)}), \xi_2(Y_{\pi(r_1+1)}, \dots, Y_{\pi(r_1+r_2)})\big].
\end{aligned}$$

There holds the following analogue of Proposition 5.5.2.

Proposition 5.5.10 *Let G be a Lie group with Lie algebra $\mathfrak{g}$ and let $\Omega^*(G, \mathfrak{g})^G$ denote the set of left-invariant differential forms on G with values in $\mathfrak{g}$.*

1. *$\Omega^*(G, \mathfrak{g})^G$ is a differential subalgebra of $\Omega^*(G, \mathfrak{g})$.*
2. *There exist natural algebra isomorphisms*

$$\Omega^*(G, \mathfrak{g})^G \cong \left(\bigwedge \mathrm{T}^* G\right) \otimes \mathfrak{g} \cong \left(\bigwedge \mathfrak{g}^*\right) \otimes \mathfrak{g},$$

   given by $\xi \mapsto \xi_{\mathbb{1}}$ and induced by the mapping $X \mapsto X_{\mathbb{1}}$, respectively.
3. *For all $\xi \in \Omega^*(G, \mathfrak{g})^G$ and $X_1, \dots, X_r \in \mathfrak{g}$, the $\mathfrak{g}$-valued function $\xi(X_1, \dots, X_r)$ on G is constant.*

Proof The arguments are completely analogous to those for ordinary left-invariant differential forms in the proof of Proposition 5.5.2, except for the surjectivity of the mapping $\xi \mapsto \xi_{\mathbb{1}}$, because $\Omega^*(G, \mathfrak{g})$ need not be generated as an algebra by 1-forms. Thus, let $\eta \in \bigwedge^r \mathrm{T}_{\mathbb{1}}^* G \otimes \mathfrak{g}$. Choose a basis $\{e_i\}$ in $\mathfrak{g}$ and write $\eta = \eta^i \otimes e_i$ with $\eta^i \in \bigwedge^r \mathrm{T}_{\mathbb{1}}^* G$. According to Proposition 5.5.2/1, $\eta^i = \xi^i_{\mathbb{1}}$ for certain left-invariant $\xi^i \in \Omega^r(G)$. Then, $\xi := \xi^i \otimes e_i$ is a left-invariant differential r-form with values in $\mathfrak{g}$ satisfying $\xi_{\mathbb{1}} = \eta$. □

According to Proposition 5.5.10, left-invariant differential 1-forms with values in $\mathfrak{g}$ can be identified with linear endomorphisms of $\mathfrak{g}$.

Definition 5.5.11 (Maurer-Cartan form) The left-invariant differential 1-form on G with values in $\mathfrak{g}$ corresponding to $\mathrm{id}_{\mathfrak{g}}$ is called the Maurer-Cartan form of G.

We denote the Maurer-Cartan form by Θ. By definition,

$$\langle\Theta, X\rangle = X \quad \text{for all } X \in \mathfrak{g}. \tag{5.5.9}$$

As a consequence, Θ assigns to a tangent vector Y at a the left-invariant vector field X with $X_a = Y$. Moreover, the expansion (5.5.8) of Θ with respect to a basis $\{e_i\}$ in $\mathfrak{g}$ is given by

$$\Theta = e^{*i} \otimes e_i .$$

Remark 5.5.12

1. Like for ordinary left-invariant 1-forms, and by the same argument, the exterior derivative of left-invariant 1-forms with values in $\mathfrak{g}$ is given by (5.5.2). For the Maurer-Cartan form this yields

$$\mathrm{d}\Theta(X,Y) = -\big\langle\Theta, [X,Y]\big\rangle = -[X,Y] = -\big[\Theta(X), \Theta(Y)\big]$$

   and hence

$$\mathrm{d}\Theta + \frac{1}{2}\Theta \wedge \Theta = 0. \tag{5.5.10}$$

   This is the Maurer-Cartan equation (5.5.4) in a basis-independent version.
2. For a classical Lie group $G \subset \mathrm{End}(V)$, under the identification of the Lie algebra of G with a subalgebra of $\mathfrak{gl}(V)$ and of the tangent spaces $\mathrm{T}_a G$ with subspaces of $\mathrm{End}(V)$, cf. Example 5.2.6, there holds

$$\Theta_a(Y) = a^{-1}Y, \quad a \in G,\ Y \in \mathrm{T}_a G.$$

   In the physics literature, it is common to write $\mathrm{d}a$ for the tangent mapping[12] of id_G at a. Using $(\mathrm{id}_G)'_a = \mathrm{id}_{\mathrm{T}_a G}$, one obtains $\Theta_a = a^{-1}\mathrm{d}a$.

We conclude this section with a remark on right-invariant differential forms.

Remark 5.5.13 (Right-invariant differential forms) By analogy with left-invariant differential forms one defines right-invariant differential forms on G by the condition $\mathrm{R}_a^*\xi = \xi$ and shows the following (Exercise 5.5.7).

1. Equivalent conditions for a differential 1-form ξ to be right-invariant are $\xi_{ba} \circ (\mathrm{R}_a)'_b = \xi_b$ for all $a, b \in G$ and $\xi_a = \xi_{\mathbb{1}} \circ (\mathrm{R}_{a^{-1}})'_a$ for all $a \in G$.
2. Proposition 5.5.2 remains true for right-invariant differential forms if $\mathfrak{g}$ is replaced by the subalgebra of right-invariant vector fields. In terms of right-invariant 1-forms, the inverse left and right trivializations $(\chi_{\mathrm{L}}^{\mathrm{T}})^{-1}$ and $(\chi_{\mathrm{R}}^{\mathrm{T}})^{-1}$ of T^*G, given by (5.1.6), read

$$\big(\chi_{\mathrm{L}}^{\mathrm{T}}\big)^{-1}(a,\xi) = \big(\mathrm{C}_{a^{-1}}^*\xi\big)_a \equiv \xi_a \circ (\mathrm{C}_{a^{-1}})'_a, \qquad \big(\chi_{\mathrm{R}}^{\mathrm{T}}\big)^{-1}(a,\xi) = \xi_a.$$

[12] Like writing $\mathrm{d}x$ for the first derivative at x of the function $f(x) = x$ on $\mathbb{R}$; it must not be confused with the notation $\mathrm{v}_G(a) = \mathrm{d}a$ for an invariant volume form.

3. Via $-\text{inv}^*$, the inversion mapping $\text{inv} : G \to G$ provides an isomorphism of differential algebras between the subalgebras of left and right-invariant differential forms. Under the natural isomorphisms with $\bigwedge \mathrm{T}^*_{\mathbb{1}} G$, $-\text{inv}^*$ corresponds to the identical mapping.

Exercises

5.5.1 Show that Eqs. (5.5.3) and (5.5.4) are equivalent.

5.5.2 Show that the pull-back of a left-invariant volume form on G by R_a, $a \in G$, is given by (5.5.6) and that integrable smooth functions satisfy (5.5.7).

5.5.3 Consider the Riemannian metric g on G defined via (5.5.5) by a scalar product on $\mathfrak{g}$. Prove that if the scalar product is Ad-invariant, then g is bi-invariant.

5.5.4 (a) Show that the mapping

$$R : (-\pi, \pi) \times (0, \pi) \times (-\pi, \pi) \to \mathrm{SO}(3),$$
$$R(\phi, \vartheta, \psi) := R_z(\phi) R_y(\vartheta) R_z(\psi),$$

where

$$R_z(\phi) = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad R_y(\phi) = \begin{bmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{bmatrix}$$

(rotation by an angle ϕ about the z-axis and the y-axis, respectively), induces a local chart on SO(3). What is the complement of the domain of this chart?[13]

(b) Show that the left-invariant volume form with respect to which SO(3) has unit volume is given in these coordinates by

$$\mathrm{v}_{\mathrm{SO}(3)} = \frac{1}{8\pi^2} \sin\vartheta \, \mathrm{d}\phi \wedge \mathrm{d}\vartheta \wedge \mathrm{d}\psi.$$

5.5.5 Use the result of Exercise 5.1.7(a) to show that the mapping

$$U : (-\pi, \pi) \times (0, \pi) \times (-2\pi, 2\pi) \to \mathrm{SU}(2),$$

given by

$$U(\phi, \vartheta, \psi) := \begin{bmatrix} \mathrm{e}^{\frac{\mathrm{i}}{2}(\psi+\phi)} \cos\frac{\vartheta}{2} & -\mathrm{e}^{\frac{\mathrm{i}}{2}(\psi-\phi)} \sin\frac{\vartheta}{2} \\ \mathrm{e}^{-\frac{\mathrm{i}}{2}(\psi-\phi)} \sin\frac{\vartheta}{2} & \mathrm{e}^{-\frac{\mathrm{i}}{2}(\psi+\phi)} \cos\frac{\vartheta}{2} \end{bmatrix},$$

induces a local chart on SU(2). Compute the left-invariant volume form with respect to which SU(2) has unit volume in these coordinates.

5.5.6 Find a left-invariant volume form on $\mathrm{SL}(2, \mathbb{C})$.

5.5.7 Verify the properties of right-invariant differential forms stated in Remark 5.5.13.

[13] The angles ϕ, ϑ and ψ are called Euler angles.

5.6 Lie Subgroups

Let G be a Lie group of dimension n and let $\mathfrak{g}$ be the Lie algebra of G.

Definition 5.6.1 (Lie subgroup) A Lie subgroup of G is a pair (H, φ), where H is a Lie group and $\varphi : H \to G$ is an injective and immersive Lie group homomorphism. If φ is an embedding, (H, φ) is said to be embedded. Lie subgroups (H_1, φ_1) and (H_2, φ_2) of G are said to be equivalent if there exists a Lie group isomorphism $\psi : H_1 \to H_2$ such that $\varphi_1 = \varphi_2 \circ \psi$.

Example 5.6.2

1. The classical groups of Example 1.2.6 are embedded Lie subgroups of the appropriate general linear group $\mathrm{GL}(n, \mathbb{K})$. Here, the Lie group homomorphism φ is given by the natural inclusion mapping.
2. $(\mathrm{GL}(n, \mathbb{H}), \varphi)$, with φ being induced from (1.1.2), is an embedded Lie subgroup of $\mathrm{GL}(2n, \mathbb{C})$.
3. Let $a, b \in \mathbb{R} \setminus \{0\}$ such that $\frac{a}{b}$ is irrational and let $\varphi : \mathbb{R} \to \mathrm{T}^2$, $\varphi(t) := (\mathrm{e}^{\mathrm{i}at}, \mathrm{e}^{\mathrm{i}bt})$. Then, $(\mathbb{R}, \varphi)$ is a Lie subgroup of T^2. It is not embedded.
4. Let $a, b \in \mathbb{Z} \setminus \{0\}$ be relatively prime and let $\varphi : \mathrm{U}(1) \to \mathrm{T}^2$, $\varphi(z) := (z^a, z^b)$. Then, $(\mathrm{U}(1), \varphi)$ is an embedded Lie subgroup of T^2.
5. The identity component G_0 of G is an embedded Lie subgroup of G.

Right from the start we observe that Propositions 1.6.10 and 1.6.14 remain true if the terms submanifold and C^k-mapping are replaced by Lie subgroup and Lie group homomorphism, respectively. The situation with Lie subgroups is yet simpler than that with general submanifolds, see Proposition 5.6.4 below. We start with deriving the basic properties of Lie subgroups. A distribution D on G is said to be left-invariant if $\mathrm{L}'_a D_b = D_{ab}$ for all $a, b \in G$.

Lemma 5.6.3 *The distribution on G spanned by a Lie subalgebra of $\mathfrak{g}$ is left-invariant, regular and integrable.*

Proof Left-invariance holds by construction and implies regularity. The distribution is involutive by Remark 3.5.6 and hence integrable by Corollary 3.5.12. □

Proposition 5.6.4 (Basic properties of Lie subgroups)

1. *The connected components of a Lie subgroup (H, φ) of G are maximal integral manifolds of the distribution on G spanned by the Lie subalgebra* $\mathrm{im}(\mathrm{d}\varphi)$ *of* $\mathfrak{g}$.
2. *Lie subgroups are initial submanifolds.*
3. *Lie subgroups are equivalent iff their images coincide as sets.*

Assertion 2 implies that the figure eight submanifold cannot occur as the group manifold of a Lie subgroup. Assertion 3 yields, in particular, that if a subgroup

$H \subset G$ admits a smooth manifold structure which makes it into a Lie subgroup of G, then this structure is unique.

Proof Let $\mathfrak{h}$ denote the Lie algebra of H and let D denote the distribution on G spanned by $\mathrm{im}(\mathrm{d}\varphi) = \mathrm{d}\varphi(\mathfrak{h})$. For $a \in H$ and $X \in \mathfrak{h}$, a brief computation shows $\varphi' X_a = (\mathrm{d}\varphi(X))_a$. This implies $\varphi' \mathrm{T}_a H = D_{\varphi(a)}$ for all $a \in H$. Hence, for every connected component H_i of H, $(H_i, \varphi_{\restriction H_i})$ is an integral manifold of D. It is maximal, because the vector fields in $\mathfrak{h}$ are complete. This proves assertion 1. Then, assertion 2 follows by means of Proposition 3.5.15. In view of the fact that the restriction of a group homomorphism in range to a subgroup remains a group homomorphism, assertion 3 follows from assertion 2 by the same argument as for the analogous assertion about initial submanifolds, see Remark 1.6.13/5. □

Next, we link Lie subgroups to Lie subalgebras.

Proposition 5.6.5 (Connected Lie subgroups and Lie subalgebras) *The assignment of* $\mathrm{im}(\mathrm{d}\varphi)$ *to* (H, φ) *defines a bijection between equivalence classes of connected Lie subgroups of G and Lie subalgebras of* $\mathfrak{g}$.

Proof We must show that for every Lie subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ there exists a connected Lie subgroup (H, φ) of G such that $\mathrm{im}(\mathrm{d}\varphi) = \mathfrak{h}$ and that (H, φ) is unique up to equivalence.

Existence: let D be the distribution on G spanned by $\mathfrak{h}$. By Lemma 5.6.3, D is left-invariant and integrable. Let H be the maximal integral manifold of D through $\mathbb{1}$. By left-invariance of D, for every $a \in H$, $\mathrm{L}_{a^{-1}}(H)$ is an integral manifold of D. Since it contains $\mathbb{1}$, we get $\mathrm{L}_{a^{-1}}(H) \subset H$. Hence, $a^{-1}b \in H$ for all $a, b \in H$, that is, H is a subgroup of G. Since the multiplication mapping of H is the restriction of that of G in domain to $H \times H$ and in range to H, it is smooth by Proposition 3.5.15. Hence, H is a Lie group. Since, by construction, the natural inclusion mapping $\iota : H \to G$ is a Lie group homomorphism, (H, ι) is a Lie subgroup of G. Since $\mathrm{T}_{\mathbb{1}} H = \{X_{\mathbb{1}} : X \in \mathfrak{h}\}$, the image of $\mathrm{d}\varphi$ coincides with $\mathfrak{h}$. This proves existence.

Uniqueness: let (H_1, φ_1) and (H_2, φ_2) be connected Lie subgroups of G satisfying $\mathrm{im}(\mathrm{d}\varphi_1) = \mathrm{im}(\mathrm{d}\varphi_2)$. By assertion 1 of Proposition 5.6.4, they are maximal integral manifolds through $\mathbb{1}$ of one and the same distribution on G. Hence, $\varphi_1(H_1) = \varphi_2(H_2)$. Then, assertion 3 of this proposition implies that (H_1, φ_1) and (H_2, φ_2) are equivalent. □

Finally, we derive two sufficient conditions for a subgroup of G to be a Lie subgroup. We suppress the natural inclusion mapping in the notation.

Proposition 5.6.6 (Subgroups which are submanifolds) *Let* $H \subset G$ *be a subgroup. If* H *admits a smooth structure which makes it into a submanifold of* G, *then* H *is a Lie group and a Lie subgroup of* G *with respect to this structure. The corresponding Lie subalgebra is given by* $\mathfrak{h} = \{X \in \mathfrak{g} : X_{\mathbb{1}} \in \mathrm{T}_{\mathbb{1}} H\}$.

The smooth structure is then unique by Proposition 5.6.4/3.

Proof Define $\mathfrak{h} := \{X \in \mathfrak{g} : X_{\mathbb{1}} \in \mathrm{T}_{\mathbb{1}} H\}$ and let D denote the distribution spanned by $\mathfrak{h}$. It suffices to show that H is an integral manifold of D, because then $\mathrm{L}_a(H)$ is an integral manifold of D through a for all $a \in G$, hence D is integrable and Proposition 3.5.15 yields that the restriction of the multiplication mapping of G to H is smooth. This yields the assertion.

Thus, let $a \in H$. Let $k := \dim(H)$. Since both D and H have dimension k, it suffices to show that $D_a \subset \mathrm{T}_a H$. Assume, on the contrary, that this is not the case. For a smooth curve γ in H and $b \in H$, define $\gamma^b := \mathrm{L}_b \circ \gamma$. Note that γ^b is a curve in H which is smooth in G, but we cannot assume that it is smooth in H, because a priori we do not know whether L_b is a smooth mapping of H. Choose smooth curves $\gamma_1, \dots, \gamma_k$ in H through a such that $\dot\gamma_1(0), \dots, \dot\gamma_k(0)$ form a basis in $\mathrm{T}_a H$. Since $D_a \not\subset \mathrm{T}_a H$ and since $D_a = \mathrm{L}'_a \mathrm{T}_{\mathbb{1}} H$, there exists a smooth curve γ_{k+1} in H through $\mathbb{1}$ such that $\dot\gamma^a_{k+1}(0) \notin \mathrm{T}_a H$. Then, $\gamma_1^{a^{-1}}, \dots, \gamma_k^{a^{-1}}$ are curves in H through $\mathbb{1}$, smooth in G, such that $\{\dot\gamma_1^{a^{-1}}(0), \dots, \dot\gamma_k^{a^{-1}}(0), \dot\gamma_{k+1}(0)\}$ is a linearly independent system in $\mathrm{T}_{\mathbb{1}} G$. Choose smooth curves $\gamma_{k+2}, \dots, \gamma_n$ in G through $\mathbb{1}$ whose tangent vectors at $t = 0$ complement this system to a basis in $\mathrm{T}_{\mathbb{1}} G$. Then, according to the Inverse Mapping Theorem, there exist $\varepsilon > 0$ and an open neighbourhood U of $\mathbb{1}$ in G such that the mapping

$$\Phi : (-\varepsilon, \varepsilon)^n \to U \subset G, \qquad \Phi(\mathbf{t}) := \gamma_1^{a^{-1}}(t_1) \cdots \gamma_k^{a^{-1}}(t_k) \gamma_{k+1}(t_{k+1}) \cdots \gamma_n(t_n),$$

is a diffeomorphism. Since $H \cap U$ is open in H, it is a submanifold of U, hence $(H \cap U, \Phi^{-1}_{\restriction H \cap U})$ is a submanifold of $\mathbb{R}^n$ of dimension k. On the other hand, by construction, the image $\Phi^{-1}(H \cap U)$ contains $(-\varepsilon, \varepsilon)^{k+1} \times \{0\}$. This is a contradiction, because the image of a submanifold cannot contain the image of another submanifold of higher dimension. Hence, $D_a \subset \mathrm{T}_a H$ and the proposition is proved. □

To derive the other sufficient condition for a subgroup to be a Lie subgroup we need

Lemma 5.6.7 *If $\mathfrak{m}_1$ and $\mathfrak{m}_2$ are complementary vector subspaces of $\mathfrak{g}$, there exist open neighbourhoods U_i of the origin in $\mathfrak{m}_i$, $i = 1, 2$, and V of $\mathbb{1}$ in G such that the mapping*

$$\varphi : \mathfrak{m}_1 \times \mathfrak{m}_2 \to G, \qquad \varphi(X_1, X_2) := \exp(X_1)\exp(X_2),$$

restrics to a diffeomorphism from $U_1 \times U_2$ onto V.

Proof The tangent mapping $\varphi'_0 : \mathfrak{m}_1 \times \mathfrak{m}_2 \to \mathfrak{g}$ is given by $(X_1, X_2) \mapsto X_1 + X_2$. Since $\mathfrak{m}_1$ and $\mathfrak{m}_2$ are complementary, it is bijective. Hence, the assertion follows from the Inverse Mapping Theorem 1.5.7. □

Theorem 5.6.8 (Closed subgroups) *Let $H \subset G$ be a subgroup. If H is closed in G, then it admits a smooth structure which makes it into a Lie group and an embedded Lie subgroup of G. The corresponding Lie subalgebra of $\mathfrak{g}$ is $\mathfrak{h} = \{X \in \mathfrak{g} : \exp(tX) \in H$ for all $t \in \mathbb{R}\}$.*

Proof First, we show that $\mathfrak{h}$ is a linear subspace of $\mathfrak{g}$. By construction, $tX \in \mathfrak{h}$ for all $X \in \mathfrak{h}$ and $t \in \mathbb{R}$. For $X, Y \in \mathfrak{g}$, (5.3.13) implies

$$\exp\big(t(X+Y)\big) = \lim_{n\to\infty}\left(\exp\left(\frac{t}{n}X\right)\exp\left(\frac{t}{n}Y\right)\right)^n.$$

Since $\exp(\frac{t}{n}X)\exp(\frac{t}{n}Y) \in H$ for all n and t and since H is closed, $\exp(t(X+Y)) \in H$ for all $t \in \mathbb{R}$, hence $X+Y \in \mathfrak{h}$. Next, equip H with the relative topology induced from G. Choose a neighbourhood V of $\mathbb{1}$ in G such that $\rho := \exp^{-1} : V \to \mathfrak{g}$ is defined and smooth and hence a local chart on G. We claim that V can be adjusted so that

$$\rho(V \cap H) = \rho(V) \cap \mathfrak{h}. \tag{5.6.1}$$

If so, then the family consisting of the local charts $(\mathrm{L}_a(V), \rho \circ \mathrm{L}_{a^{-1}}\restriction_{\mathrm{L}_a(V)})$, $a \in H$, satisfies conditions (E1) and (E2) of Proposition 1.7.3 and hence defines a smooth structure on H which makes H into an embedded submanifold of G. Then, Proposition 1.6.10 implies that in this structure, H is a Lie group and hence a Lie subgroup of G. The corresponding Lie subalgebra of $\mathfrak{g}$ contains $\mathfrak{h}$ and, by (5.6.1), it has the same dimension as $\mathfrak{h}$, hence it coincides with $\mathfrak{h}$. In particular, $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$.

To show (5.6.1), assume, on the contrary, that V cannot be chosen so that (5.6.1) holds. Since $\exp(\mathfrak{h}) \subset H$, this means that $V \cap \exp(\mathfrak{h})$ is properly contained in $V \cap H$, for arbitrarily small V. Thus, there is a sequence $\{a_n\}$ in H converging to $\mathbb{1}$ such that $a_n \notin \exp(\mathfrak{h})$ for all n. Choose a subspace $\mathfrak{m}$ of $\mathfrak{g}$ complementary to $\mathfrak{h}$. According to Lemma 5.6.7, for large enough n, a_n defines unique sequences $\{X_n\}$ in $\mathfrak{h}$ and $\{Y_n\}$ in $\mathfrak{m}$ by $a_n = \exp(X_n)\exp(Y_n)$ and both of these sequences converge to the respective origin. Since H is a subgroup, we obtain $\exp(Y_n) \in H$. Since $a_n \notin \exp(\mathfrak{h})$, we conclude $Y_n \neq 0$. Now, choose a norm $\|\cdot\|$ on $\mathfrak{m}$. Since the corresponding unit sphere in $\mathfrak{m}$ is compact, by possibly removing some members we may assume that the sequence $\{Y_n\|Y_n\|^{-1}\}$ converges to some $Y \in \mathfrak{m}$ with $\|Y\| = 1$. We show that $\exp(tY) \in H$ for all $t \in \mathbb{R}$ and hence $Y \in \mathfrak{h}$ (contradiction). Let $t \in \mathbb{R}$ be given. Let $\{k_n\}$ denote the sequence of integer parts corresponding to $\{t\|Y_n\|^{-1}\}$. Then, $k_n\|Y_n\| \to t$ for $n \to \infty$ and hence

$$\exp(tY) = \lim_{n\to\infty}\exp\left(k_n\|Y_n\| \cdot \frac{Y_n}{\|Y_n\|}\right) = \lim_{n\to\infty}\big(\exp(Y_n)\big)^{k_n} \in H,$$

as H is closed. This completes the proof of the theorem. □

Remark 5.6.9 Instead of Proposition 1.7.3, one may use the following argument to prove that H, equipped with the relative topology induced from G, admits a submanifold structure. The details are left to the reader (Exercise 5.6.1). By (5.3.14), $\mathfrak{h}$ is a Lie subalgebra. Let D be the distribution on G generated by $\mathfrak{h}$ and let $\tilde{H}$ be the connected Lie subgroup associated with $\mathfrak{h}$, that is, the maximal integral manifold through $\mathbb{1}$ of D. Since the subset H of G is invariant under the flow of the vector fields in $\mathfrak{h}$, for every $a \in H$ there holds $\mathrm{L}_a(\tilde{H}) \subset H$. Use the argument of the proof of Theorem 5.6.8 to show that closedness of H implies the existence of an open neighbourhood V of $\mathbb{1}$ in G such that $\tilde{H} \cap V = H \cap V$. Construct from this an open neighbourhood $\tilde{V}$ of $\tilde{H}$ in G such that $\tilde{H} = \tilde{V} \cap H$. Conclude that the maximal integral manifolds $\mathrm{L}_a(\tilde{H})$ of D, $a \in H$, are the connected components of H. Since the latter is second countable, there are at most countably many of them. This yields the desired submanifold structure of H.

Corollary 5.6.10 *A Lie subgroup is embedded iff its image is a closed subset.*

Proof Denote the Lie subgroup by (H, φ). If $\varphi(H)$ is closed, Theorem 5.6.8 provides a Lie subgroup structure on $\varphi(H)$, where $\varphi(H)$ carries the relative topology induced from G. By Proposition 5.6.4/3, the Lie subgroups (H, φ) and $\varphi(H)$ are equivalent. This means that φ is a diffeomorphism (and hence a homeomorphism) onto its image. Conversely, if (H, φ) is embedded, Proposition 5.6.4/1 and Remark 3.5.19/1 imply that $\varphi(H)$ is closed. □

Example 5.6.11

1. The kernel of a Lie group homomorphism $G \to H$ is a subgroup (by the homomorphism property) and closed (by continuity) and hence an embedded Lie subgroup of G by Theorem 5.6.8. Using Proposition 5.3.6 it is easy to see that the Lie algebra of $\ker(\varphi)$ is given by $\ker(\mathrm{d}\varphi)$ (Exercise 5.6.2).
2. In view of Theorem 5.6.8, for the classical groups of Example 1.2.6 to be Lie groups it suffices that they are closed subsets of the appropriate general linear group $\mathrm{GL}(n, \mathbb{K})$. Proposition 5.6.4/3 implies that the smooth structure so obtained coincides with that provided by the Level Set Theorem 1.2.1.

Exercises

5.6.1 Work out the alternative proof of Theorem 5.6.8 sketched in Remark 5.6.9.

5.6.2 Use Proposition 5.3.6 to prove that the Lie algebra of the kernel of a Lie group homomorphism is given by the kernel of the induced Lie algebra homomorphism, cf. Example 5.6.11/1.

5.7 Homogeneous Spaces

Let G be a Lie group. Recall from elementary algebra that every subgroup $H \subset G$ defines an equivalence relation on G by $a \sim_H b$ iff $a^{-1}b \in H$. The equivalence classes are given by the subsets aH of G, $a \in G$. They are called the left cosets of

H in G. Let G/H denote the set of equivalence classes, equipped with the quotient topology. This space will be referred to as the quotient of G by H. By definition of the quotient topology, the natural projection

$$\pi : G \to G/H, \qquad \pi(a) := aH,$$

is continuous. Every $a \in G$ defines a mapping

$$\hat{\mathrm{L}}_a : G/H \to G/H, \qquad \hat{\mathrm{L}}_a(bH) := abH,$$

fulfilling

$$\hat{\mathrm{L}}_a \circ \pi = \pi \circ \mathrm{L}_a \quad \text{for all } a \in G. \tag{5.7.1}$$

The topological structure of the quotient G/H may be quite strange, as for example in the case $G = \mathrm{T}^2$ and $H = \{(\mathrm{e}^{\mathrm{i}at}, \mathrm{e}^{\mathrm{i}bt}) : t \in \mathbb{R}\}$ with $\frac{a}{b}$ irrational. If, however, H is closed, then the quotient G/H turns out to be a smooth manifold. To prove this, we need

Lemma 5.7.1 *Let G be a Lie group and let $H \subset G$ be a subgroup.*

1. *The natural projection $\pi : G \to G/H$ is open.*
2. *The induced mappings $\hat{\mathrm{L}}_a$ are homeomorphisms of G/H.*
3. *If H is closed, then G/H is Hausdorff.*

Proof 1. Let $U \subset G$ be open. The subset $\pi(U)$ is open in G/H iff $\pi^{-1}(\pi(U))$ is open in G. The latter holds, because $\pi^{-1}(\pi(U)) = \bigcup_{a \in H} \mathrm{R}_a(U)$ and the R_a are homeomorphisms of G.

2. Due to $\hat{\mathrm{L}}_a \circ \hat{\mathrm{L}}_{a^{-1}} = \mathrm{id}_{G/H}$, it suffices to show that $\hat{\mathrm{L}}_a$ is open for all $a \in G$. Thus, let $\hat{U} \subset G/H$ be open. By (5.7.1), $\hat{\mathrm{L}}_a(\hat{U}) = \pi \circ \mathrm{L}_a(\pi^{-1}(\hat{U}))$. Since π is continuous and since L_a and π are open mappings, the assertion follows.

3. Consider the mapping $\varphi : G \times G \to G$, defined by $\varphi(a_1, a_2) := a_1^{-1} a_2$. We have $\pi(a_1) = \pi(a_2)$ iff $(a_1, a_2) \in \varphi^{-1}(H)$. Since H is closed and since φ is continuous, $\varphi^{-1}(H)$ is a closed subset of $G \times G$. Now, let $\hat{a}_1, \hat{a}_2 \in G/H$ be such that $\hat{a}_1 \neq \hat{a}_2$. Choose $a_i \in G$ such that $\hat{a}_i = \pi(a_i)$. Then, $(a_1, a_2) \notin \varphi^{-1}(H)$. Since $\varphi^{-1}(H)$ is closed, there exist neighbourhoods V_i of a_i in G such that $(V_1 \times V_2) \cap \varphi^{-1}(H) = \varnothing$. This implies $\pi(V_1) \cap \pi(V_2) = \varnothing$. Since π is open, $\pi(V_i)$ is a neighbourhood of $\hat{a}_i$, $i = 1, 2$. This yields the assertion. □

Theorem 5.7.2 (Homogeneous space) *Let G be a Lie group and let $H \subset G$ be a closed subgroup. There exists a unique smooth structure on the quotient G/H such that the natural projection $\pi : G \to G/H$ is a smooth submersion.*

Proof It suffices to prove existence, because uniqueness holds for any submersion, see Remark 1.5.16. Choose a subspace $\mathfrak{m}$ of $\mathfrak{g}$ complementary to $\mathfrak{h}$. The manifold structure on G/H will be modelled on this subspace.

According to Lemma 5.6.7, there exist open neighbourhoods $W_{\mathfrak{h}}$ and $W_{\mathfrak{m}}$ of the origins in $\mathfrak{h}$ and $\mathfrak{m}$, respectively, and an open neighbourhood V of $\mathbb{1}$ in G such that the mapping

$$\varphi : W_{\mathfrak{m}} \times W_{\mathfrak{h}} \to V, \qquad (X, Y) \mapsto \exp(X)\exp(Y),$$

is a diffeomorphism. Since H is closed and hence embedded, $\exp(W_{\mathfrak{h}})$ is open with respect to the relative topology on H induced from G. As a consequence, $W_{\mathfrak{m}}$ may be shrunk so that $(\exp(W_{\mathfrak{m}})^2 \cap H) \subset \exp(W_{\mathfrak{h}})$. Consider the mapping $\pi \circ \exp : W_{\mathfrak{m}} \to G/H$. It is

(a) injective: if $X, Y \in W_{\mathfrak{m}}$ satisfy $\exp(-X)\exp(Y) \in H$, then $\exp(-X)\exp(Y) = \exp(Z)$ for some $Z \in W_{\mathfrak{h}}$, hence $\varphi(Y, 0) = \varphi(X, Z)$ and injectivity of φ implies $X = Y$ (and $Z = 0$);
(b) open: if U is open in $W_{\mathfrak{m}}$, then $U \times W_{\mathfrak{h}}$ is open in $W_{\mathfrak{m}} \times W_{\mathfrak{h}}$. Since φ and π are open, $\pi \circ \exp(U) = \pi \circ \varphi(U \times W_{\mathfrak{h}})$ is open in G/H.

Thus, $\pi \circ \exp$ maps $W_{\mathfrak{m}}$ homeomorphically onto the open subset

$$\hat{U} := \pi \circ \exp(W_{\mathfrak{m}})$$

of G/H and hence induces a local chart $\hat{\kappa} := (\pi \circ \exp)^{-1} : \hat{U} \to \mathfrak{m}$ on G/H. Since the mappings $\hat{\mathrm{L}}_a$ are homeomorphisms,

$$\left\{\left(\hat{\mathrm{L}}_a(\hat{U}), \hat{\kappa} \circ \hat{\mathrm{L}}_{a^{-1}}\right) : a \in G\right\}$$

is an atlas on G/H. The transition mappings can be expressed in terms of exp, $\exp^{-1}$, left translations and the projection $\mathfrak{g} \to \mathfrak{m}$ defined by the vector space decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$. Details are left to the reader. It follows that the atlas so constructed defines a smooth structure on G/H. To see that π is a smooth submersion, due to $\pi(\mathrm{L}_a(V)) = \hat{\mathrm{L}}_a(\hat{U})$, it suffices to check that the restriction $\pi : \mathrm{L}_a(V) \to \hat{\mathrm{L}}_a(\hat{U})$ is a smooth submersion for all $a \in G$. The latter follows from the fact that the mapping $\hat{\kappa} \circ \hat{\mathrm{L}}_{a^{-1}} \circ \pi \circ \mathrm{L}_a \circ \varphi : W_{\mathfrak{m}} \times W_{\mathfrak{h}} \to W_{\mathfrak{m}}$ coincides with the natural projection of the direct product. □

Remark 5.7.3 According to Remark 1.5.16, being a submersion, π admits local sections. That is, for every $\hat{a} \in G/H$ there exists an open neighbourhood $\hat{U}$ and a smooth mapping $s : \hat{U} \to M$ such that $\pi \circ s = \mathrm{id}_{\hat{U}}$. In the present situation, as a byproduct of the construction of local charts on G/H in the proof of Theorem 5.7.2, one has the following distinguished class of local sections. For every subspace $\mathfrak{m}$ complementary to $\mathfrak{h}$ there exists an open neighbourhood W of the origin in $\mathfrak{m}$ such that $\exp(W) \subset G$ is an embedded submanifold, $\pi_{\restriction \exp(W)}$ is injective and $\hat{U} := \pi \circ \exp(W)$ is open in G/H. Then $s := \pi^{-1} : \hat{U} \to \exp(W)$ is a local section of π, and so is $(\hat{\mathrm{L}}_a(\hat{U}), \mathrm{L}_a \circ s \circ \hat{\mathrm{L}}_{a^{-1}})$ for all $a \in G$. This argument also shows that the natural vector space isomorphism $\mathfrak{g} \to \mathrm{T}_{\mathbb{1}}G$, given by evaluation at $\mathbb{1}$, descends to an isomorphism

$$\mathfrak{g}/\mathfrak{h} \cong \mathrm{T}_{[\mathbb{1}]}(G/H). \tag{5.7.2}$$

The existence of local sections of π implies that the equality (5.7.1) can be resolved locally for $\hat{\mathrm{L}}_a$. This entails

Corollary 5.7.4 *The mapping $\hat{\mathrm{L}}_a$ is a diffeomorphism of G/H for all $a \in G$.*

In Sect. 6.2 it will become clear that Corollary 5.7.4 provides the justification of the term homogeneous space, referred to in the headlines of this section and of Theorem 5.7.2 (see Remark 6.2.10/3). Furthermore, in Sect. 6.5 we will see that the coset manifold G/H can be interpreted as the quotient of the action of the Lie group H on the manifold G by right translation and, correspondingly, as the base manifold of a principal bundle with structure group H. Finally, we note that the construction of the coset manifold carries over in an obvious way to the space of right H-cosets in G. In this case, right translation induces diffeomorphisms $\hat{R}_a$ for all $a \in G$.

Example 5.7.5 (Stiefel manifolds) Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$. The Stiefel manifold $\mathrm{S}_{\mathbb{K}}(k,n)$ is defined to be the set of k-frames in $\mathbb{K}^n$ which are orthonormal with respect to the standard scalar product. The actual manifold structure is obtained by identifying $\mathrm{S}_{\mathbb{K}}(k,n)$ with a certain homogeneous space as follows. Any orthonormal k-frame in $\mathbb{K}^n$ can be obtained from the first k elements of the standard basis in $\mathbb{K}^n$ by a linear isometric transformation. This transformation is determined by the given frame only up to its action on the last $n-k$ elements of the standard basis. That is, two transformations produce the same k-frame iff they differ by prior application of a block matrix

$$\begin{bmatrix} \mathbb{1}_k & 0 \\ 0 & a \end{bmatrix}, \tag{5.7.3}$$

where a is an isometry of $\mathbb{K}^{n-k}$. Hence, as sets,

$$\mathrm{S}_{\mathbb{R}}(k,n) = \mathrm{O}(n)/\mathrm{O}(n-k),$$
$$\mathrm{S}_{\mathbb{C}}(k,n) = \mathrm{U}(n)/\mathrm{U}(n-k),$$
$$\mathrm{S}_{\mathbb{H}}(k,n) = \mathrm{Sp}(n)/\mathrm{Sp}(n-k),$$

where $\mathrm{O}(n-k)$, $\mathrm{U}(n-k)$ and $\mathrm{Sp}(n-k)$ refer to the corresponding subgroups of matrices of the form (5.7.3), and these equalities define the smooth structure of $\mathrm{S}_{\mathbb{K}}(k,n)$. For $k < n$, a possible sign (if $\mathbb{K} = \mathbb{R}$) or phase (if $\mathbb{K} = \mathbb{C}$) in the determinant of the isometry producing the desired k-frame can be shifted to the irrelevant part which acts on the last $n-k$ standard basis vectors. Therefore,

$$\mathrm{S}_{\mathbb{R}}(k,n) = \mathrm{SO}(n)/\mathrm{SO}(n-k), \qquad \mathrm{S}_{\mathbb{C}}(k,n) = \mathrm{SU}(n)/\mathrm{SU}(n-k).$$

An orthonormal 1-frame is just a vector of length 1. Hence,

$$\mathrm{S}_{\mathbb{K}}(1,n) = \mathrm{S}^{dn-1}, \qquad d = \dim_{\mathbb{R}}(\mathbb{K}).$$

We leave it to the reader to check that this is a diffeomorphism, indeed (Exercise 5.7.3).

Example 5.7.6 (Graßmann manifolds) The Graßmann manifold $\mathrm{G}_{\mathbb{K}}(k,n)$ is defined to be the set of k-dimensional subspaces of $\mathbb{K}^n$. As for the Stiefel manifolds, the manifold structure of $\mathrm{G}_{\mathbb{K}}(k,n)$ is obtained by identifying it with a homogeneous space. Any k-dimensional subspace of $\mathbb{K}^n$ can be obtained from the subspace spanned by the first k elements of the standard basis in $\mathbb{K}^n$ by application of a linear isometry. This transformation is determined by the given subspace up to prior application of a block matrix

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \tag{5.7.4}$$

where a and b are isometries of $\mathbb{K}^k$ and $\mathbb{K}^{n-k}$, respectively. Hence,

$$\begin{aligned} \mathrm{G}_{\mathbb{R}}(k,n) &= \mathrm{O}(n)/\big(\mathrm{O}(k)\times\mathrm{O}(n-k)\big), \\ \mathrm{G}_{\mathbb{C}}(k,n) &= \mathrm{U}(n)/\big(\mathrm{U}(k)\times\mathrm{U}(n-k)\big), \\ \mathrm{G}_{\mathbb{H}}(k,n) &= \mathrm{Sp}(n)/\big(\mathrm{Sp}(k)\times\mathrm{Sp}(n-k)\big), \end{aligned}$$

where the subgroups consist of matrices of the form (5.7.4) and the smooth structure of $\mathrm{G}_{\mathbb{K}}(k,n)$ is defined by that of the corresponding homogeneous space. The case $k=1$ reproduces the projective spaces $\mathrm{G}_{\mathbb{K}}(1,n)=\mathbb{K}\mathrm{P}^{n-1}$ of Example 1.1.15. The proof that both sides are in fact diffeomorphic is postponed until Sect. 6.5 (Example 6.5.4/3).

Example 5.7.7 (Flag manifolds) Let $0<k_1<\cdots<k_{r-1}<n$ be integers. A flag of type $(k_1,\dots,k_{r-1})$ in $\mathbb{K}^n$ is an ascending sequence of vector subspaces $V_1\subset\cdots\subset V_{r-1}\subset\mathbb{K}^n$ where $\dim(V_i)=k_i$. By defining W_1 to be V_1 and W_{i+1} to be the orthogonal complement of V_i in V_{i+1}, with every flag one can associate a decomposition $\mathbb{K}^n=W_1\oplus\cdots\oplus W_r$ into mutually orthogonal subspaces, and this defines a bijection from the set of flags of type $(k_1,\dots,k_{r-1})$ onto the set of orthogonal direct sum decompositions of $\mathbb{K}^n$ into subspaces of the dimensions $(n_1,\dots,n_r)$, where $n_1=k_1$, $n_r=n-k_{r-1}$ and $n_{i+1}=k_{i+1}-k_i$ in between. Let $\mathrm{F}_{\mathbb{K}}(n_1,\dots,n_r)$ denote either of these sets. A similar argument as for the Graßmann manifolds shows that

$$\begin{aligned} \mathrm{F}_{\mathbb{R}}(n_1,\dots,n_r) &= \mathrm{O}(n)/\big(\mathrm{O}(n_1)\times\cdots\times\mathrm{O}(n_r)\big), \\ \mathrm{F}_{\mathbb{C}}(n_1,\dots,n_r) &= \mathrm{U}(n)/\big(\mathrm{U}(n_1)\times\cdots\times\mathrm{U}(n_r)\big), \\ \mathrm{F}_{\mathbb{H}}(n_1,\dots,n_r) &= \mathrm{Sp}(n)/\big(\mathrm{Sp}(n_1)\times\cdots\times\mathrm{Sp}(n_r)\big), \end{aligned}$$

where the subgroups consist of block diagonal matrices

$$\begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_r \end{bmatrix}$$

with a_i being an element of, respectively, $\mathrm{O}(n_i)$, $\mathrm{U}(n_i)$ or $\mathrm{Sp}(n_i)$, $i = 1, \ldots, r$. As before, these equalities are used to define a smooth structure on $\mathrm{F}_{\mathbb{K}}(n_1, \ldots, n_r)$. The case $r = 1$ and $n_1 = k$ reproduces the Graßmann manifolds: $\mathrm{F}_{\mathbb{K}}(k, n-k) = \mathrm{G}_{\mathbb{K}}(k, n)$.

Next, we investigate the special case where H is a normal subgroup. This means that $aH = Ha$, or equivalently $\mathrm{C}_a(H) = H$, for all $a \in G$. In this case, G/H inherits a group structure, defined by the condition that the natural projection $\pi : G \to G/H$ be a group homomorphism.

Proposition 5.7.8 (Quotient Lie group) *Let G be a Lie group with Lie algebra $\mathfrak{g}$ and let $H \subset G$ be a closed normal subgroup with associated Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$.*

1. *G/H with the induced group structure is a Lie group.*
2. *The natural projection $\pi : G \to G/H$ is a Lie group homomorphism.*
3. *We have* $\ker(\mathrm{d}\pi) = \mathfrak{h}$ *and* $\mathrm{d}\pi$ *induces an isomorphism from $\mathfrak{g}/\mathfrak{h}$ onto the Lie algebra of G/H.*

The Lie group G/H will be referred to as the quotient Lie group of G by H and the Lie algebra of G/H will be identified with $\mathfrak{g}/\mathfrak{h}$ by means of the natural isomorphism induced by $\mathrm{d}\pi$ without further notice. Under this identification, $\mathrm{d}\pi$ corresponds to the natural projection $\mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$.

Proof 1. Since π is a group homomorphism, the multiplication mappings μ of G and $\hat{\mu}$ of G/H satisfy $\hat{\mu} \circ (\pi \times \pi) = \pi \circ \mu$. Using pairs of local sections of π, this equality can be locally resolved for $\hat{\mu}$. Hence, $\hat{\mu}$ is smooth and G/H is a Lie group.

2. This holds by construction.

3. Due to $\ker(\pi) = H$ and Example 5.6.11/1, $\ker(\mathrm{d}\pi) = \mathfrak{h}$ and the assertion follows from the homomorphism theorem for algebras.[14] □

Example 5.7.9

1. Let $\varphi : G \to H$ be a Lie group homomorphism. Then, $\ker(\varphi)$ is a closed normal subgroup of G and hence $G/\ker(\varphi)$ is a Lie group with Lie algebra $\mathfrak{g}/\ker(\mathrm{d}\varphi)$.
2. The identity component G_0 of G is a normal subgroup: for every $a \in G$, C_a is a homeomorphism of G, hence $\mathrm{C}_a(G_0)$ is a connected component of G. Since $\mathrm{C}_a(\mathbb{1}) = \mathbb{1}$, it coincides with G_0. As a connected component, G_0 is also closed and hence G/G_0 is a Lie group. The cosets coincide with the connected components of G. Hence, the quotient topology on G/G_0 is discrete. Thus, the quotient Lie group of G by G_0 is given by the group-theoretical quotient of G by G_0, equipped with the discrete smooth structure.

[14]The kernel of an algebra homomorphism is an ideal of the domain and the mapping induced by passing to the quotient is an algebra isomorphism onto its image.

Proposition 5.7.10 (Quotient homomorphism) *Let G_i be Lie groups with Lie algebras $\mathfrak{g}_i$ and let $H_i \subset G_i$ be closed normal subgroups with associated Lie subalgebras $\mathfrak{h}_i \subset \mathfrak{g}_i$, $i = 1, 2$. Let $\varphi : G_1 \to G_2$ be a Lie group homomorphism satisfying $\varphi(H_1) \subset H_2$.*

1. *There exists a unique mapping $\hat{\varphi} : G_1/H_1 \to G_2/H_2$ such that for the natural projections $\pi_i : G_i \to G_i/H_i$, $i = 1, 2$, there holds $\hat{\varphi} \circ \pi_1 = \pi_2 \circ \varphi$. This mapping is a Lie group homomorphism.*
2. *There holds $\mathrm{d}\varphi(\mathfrak{h}_1) \subset \mathfrak{h}_2$ and $\mathrm{d}\hat{\varphi}$ is given by the homomorphism $\mathfrak{g}_1/\mathfrak{h}_1 \to \mathfrak{g}_2/\mathfrak{h}_2$ induced by $\mathrm{d}\varphi$ on passing to the quotients.*

Proof 1. Due to $\varphi(H_1) \subset H_2$, the mapping $\hat{\varphi}$ is well-defined. It is obviously unique. A brief calculation, using that π_1, π_2 and φ are group homomorphisms, shows that $\hat{\varphi}$ is a group homomorphism. Since π_1 admits local sections, $\hat{\varphi}$ is smooth.

2. By (5.2.9), there holds $\mathrm{d}\hat{\varphi} \circ \mathrm{d}\pi_1 = \mathrm{d}\pi_2 \circ \mathrm{d}\varphi$. Since under the identification of the Lie algebras of G_i/H_i with the quotients $\mathfrak{g}_i/\mathfrak{h}_i$, $\mathrm{d}\pi_i$ corresponds to the natural projection $\mathfrak{g}_i \to \mathfrak{g}_i/\mathfrak{h}_i$, this implies the assertion. □

Propositions 5.7.8 and 5.7.10 provide a Lie group analogue of the homomorphism theorem of group theory which states that for a group homomorphism $\varphi : G \to H$, the image $\mathrm{im}(\varphi)$ is a subgroup of H and that the induced mapping $\hat{\varphi} : G/\ker(\varphi) \to \mathrm{im}(\varphi)$ is a group isomorphism.

Proposition 5.7.11 (Homomorphism theorem) *Let $\varphi : G \to H$ be a Lie group homomorphism. Then, $\mathrm{im}(\varphi)$ is a Lie subgroup of H and the induced mapping $\hat{\varphi} : G/\ker(\varphi) \to \mathrm{im}(\varphi)$ is a Lie group isomorphism.*

The assertion may be rephrased by stating that $(G/\ker(\varphi), \hat{\varphi})$ is a Lie subgroup of H.

Proof By Propositions 5.7.8 and 5.7.10, $G/\ker(\varphi)$ is a Lie group and the induced mapping $\hat{\varphi} : G/\ker(\varphi) \to H$ is a Lie group homomorphism. By construction, $\hat{\varphi}$ is injective and hence an immersion by Corollary 5.3.7. Thus, $(G/\ker(\varphi), \hat{\varphi})$ is a Lie subgroup of H and so is the image $\mathrm{im}(\varphi)$ with respect to the smooth structure transported by $\hat{\varphi}$. □

Example 5.7.12

1. The exponential function $\varphi : \mathbb{R} \to \mathbb{C}$, defined by $t \mapsto \mathrm{e}^{2\pi \mathrm{i} t}$ has kernel $\mathbb{Z}$ and image $\mathrm{U}(1)$. The induced mapping yields a Lie group isomorphism from $\mathbb{R}/\mathbb{Z}$ onto $\mathrm{U}(1)$. This extends to an isomorphism from $\mathbb{R}^n/\mathbb{Z}^n$ onto T^n for every n.
2. The Lie group homomorphism $\mathrm{SU}(2) \to \mathrm{SO}(3)$ of Example 5.1.11 is surjective and has the centre $\mathbb{Z}_2$ of $\mathrm{SU}(2)$ as its kernel, hence the induced mapping yields a Lie group isomorphism from $\mathrm{SU}(2)/\mathbb{Z}_2$ onto $\mathrm{SO}(3)$. Similarly, the Lie group homomorphism $\mathrm{SL}(2, \mathbb{C}) \to \mathrm{SO}(3, 1)$ of Example 5.1.13 induces a Lie group isomorphism from the quotient of $\mathrm{SL}(2, \mathbb{C})$ by its centre onto the identity component of $\mathrm{SO}(3, 1)$.

Exercises

5.7.1 Prove Corollary 5.7.4.

5.7.2 Compute the dimensions of the Stiefel, Graßmann and flag manifolds, see Examples 5.7.5–5.7.7.

5.7.3 Let $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{H}$ and let $d = \dim_{\mathbb{R}} \mathbb{K}$. Show that the Stiefel manifold $\mathrm{S}_{\mathbb{K}}(1, n)$ is diffeomorphic to the sphere S^{dn-1}, cf. Example 5.7.5.

Chapter 6
Lie Group Actions

In Sect. 6.1, we define the notions of Lie group action and G-manifold and collect some of their elementary properties. There is a variety of derived notions fitting together to a geometric structure which will be studied in this chapter. In particular, a Lie group action gives rise to a special type of vector fields, so-called Killing vector fields,[1] see Sect. 6.2. These vector fields span an integrable distribution whose integral manifolds coincide with the orbits of the group action. This way, every orbit is endowed with the structure of an initial submanifold. Starting from Sect. 6.3, we limit our attention to the important special class of proper group actions. Under this additional regularity assumption, one can prove the Tubular Neighbourhood Theorem[2] which constitutes one of the basic tools of the theory of Lie group actions, see Sect. 6.4. It states that for every orbit there exists a G-invariant neighbourhood and a diffeomorphism identifying this neighbourhood G-equivariantly with a G-invariant neighbourhood of the zero section in the normal bundle of this orbit. In particular, we study the case of a free proper action in some detail, because it gives rise to interesting bundle structures. Next, in Sect. 6.6, we study elementary properties of the orbit space of a given Lie group action, and in Sect. 6.7 we discuss invariant vector fields. The latter notion is of basic importance for the study of physical systems with symmetries, see Chap. 10. In Sect. 6.8, we make some elementary remarks on relative equilibria and relatively periodic integral curves.

6.1 Basics

Let M be a smooth manifold and let G be a Lie group. Let $\mathfrak{g}$ be the Lie algebra of G. Starting from this chapter, elements of $\mathfrak{g}$ will be denoted by $A, B, \ldots$. Depending on the context, they will be viewed either as left-invariant vector fields on G or as elements of the tangent space $\mathrm{T}_{\mathbb{1}}G$.

[1] Or fundamental vector fields.

[2] Also known as the Slice Theorem.

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*,
Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7_6,

Definition 6.1.1 (Lie group action)

1. An action of G on M is a smooth mapping $\Psi : G \times M \to M$ such that the induced mappings
$$\Psi_a : M \to M, \qquad \Psi_a(m) := \Psi(a, m),$$
satisfy $\Psi_{\mathbb{1}} = \mathrm{id}_M$ and either $\Psi_a \circ \Psi_b = \Psi_{ab}$ or $\Psi_a \circ \Psi_b = \Psi_{ba}$ for all $a, b \in G$. In the first case, Ψ is called a left action and in the second case a right action. The triple (M, G, Ψ) is referred to as a Lie group action and the pair (M, Ψ) as a G-manifold.
2. A morphism of Lie group actions (M_1, G_1, Ψ^1) and (M_2, G_2, Ψ^2) (both left or both right) consists of a smooth mapping $\varphi : M_1 \to M_2$ and a Lie group homomorphism $\varrho : G_1 \to G_2$ such that
$$\varphi \circ \Psi^1 = \Psi^2 \circ (\varrho \times \varphi). \tag{6.1.1}$$
One says that the mapping φ intertwines the actions Ψ^1 and Ψ^2. If $G_1 = G_2$ and ϱ is the identical mapping, φ is also called equivariant or, equivalently, a morphism of the G-manifolds (M_1, Ψ^1) and (M_2, Ψ^2).

Since $\Psi_a \circ \Psi_{a^{-1}} = \mathrm{id}_M$, the induced mappings Ψ_a are diffeomorphisms of M. Hence, the assignment $a \mapsto \Psi_a$ defines a group homomorphism (in the case of a left action) or a group anti-homomorphism (in the case of a right action) from G to the group of diffeomorphisms of M. Condition (6.1.1) is equivalent to
$$\Psi^2_{\varrho(a)} \circ \varphi = \varphi \circ \Psi^1_a \quad \text{for all } a \in G. \tag{6.1.2}$$
In particular, a mapping $\varphi : M_1 \to M_2$ is equivariant with respect to the G-actions Ψ^i on M_i if and only if
$$\Psi^2_a \circ \varphi = \varphi \circ \Psi^1_a \quad \text{for all } a \in G. \tag{6.1.3}$$
Analogously to the notion of morphism one defines the notion of anti-morphism from a left to a right Lie group action, or vice versa, by requiring that ϱ be a Lie group anti-homomorphism. If $G_1 = G_2 = G$ and ϱ is the inversion mapping, that is, if
$$\Psi^2_{a^{-1}} \circ \varphi = \varphi \circ \Psi^1_a \quad \text{for all } a \in G, \tag{6.1.4}$$
then φ is called anti-equivariant or an anti-morphism of G-manifolds. If G is Abelian, there is no difference between left and right actions. In the general case, a left action can be turned into a right action, and vice versa, by composing it with the mapping $G \times M \to G \times M$ given by $(a, m) \mapsto (a^{-1}, m)$. Then, id_M is anti-equivariant and hence an anti-morphism of G-manifolds.

Example 6.1.2 We encourage the reader to check the axioms of Lie group action for each of the following examples (Exercise 6.1.1).

1. Every representation $\varrho : G \to \mathrm{GL}(V)$ of G on a finite-dimensional $\mathbb{K}$-vector space V defines a left action of G on V by $\Psi(a, v) := \varrho(a)v$ for all $a \in G, v \in V$.

Since the induced diffeomorphisms $\Psi_a = \varrho(a)$ are linear, one speaks of a linear action. A homomorphism of representations of G is the same as a morphism of the associated G-manifolds.

2. The operations of left translation, right translation and conjugation, defined in (5.1.3), yield actions of a Lie group on itself. Each of them may be restricted in the first argument to a Lie subgroup $H \subset G$, thus producing the actions of H on G by left and right translation and conjugation, respectively.
3. For every closed subgroup $H \subset G$, the induced left translations $\hat{\mathrm{L}}_a : G/H \to G/H$, where $a \in G$, define an action of G on the homogeneous space G/H of left cosets. The natural projection $G \to G/H$ is equivariant and hence a morphism of G-manifolds. A similar statement holds for the induced right translations $\hat{\mathrm{R}}_a$ on the homogeneous space of right cosets.
4. The flow of a complete vector field on M is an action of $\mathbb{R}$ on M. More generally, according to Proposition 3.2.15, the flows of commuting complete vector fields $X_1, \ldots, X_r$ on M define an action of the vector group $\mathbb{R}^r$ on M by
$$\Psi_{\mathbf{t}} := \Phi^{X_1}_{t_1} \circ \cdots \circ \Phi^{X_r}_{t_r} . \qquad (6.1.5)$$
By Proposition 3.2.13/2, every diffeomorphism φ of M intertwines Ψ with the corresponding action defined by $\varphi_* X_1, \ldots, \varphi_* X_r$.
5. With every Lie group action (M, G, Ψ) there come induced actions of G on $\mathrm{T}M$ and T^*M, given by $(a, X) \mapsto \Psi'_a X$ for all $X \in \mathrm{T}M$ and
$$(a, \xi_m) \mapsto \xi_m \circ (\Psi_a)'_{\Psi_{a^{-1}}(m)}$$
for all $\xi_m \in \mathrm{T}^*_m M$, respectively. Since the induced diffeomorphisms of $\mathrm{T}M$ and T^*M are vector bundle automorphisms, these actions turn $\mathrm{T}M$ and T^*M into G-vector bundles, see Remark 6.1.3 below. The natural projections $\mathrm{T}M \to M$ and $\mathrm{T}^*M \to M$ are equivariant. All of this carries over in an obvious way to arbitrary tensor bundles over M.
6. The direct product of Lie group actions (M_i, G_i, Ψ^i), $i = 1, 2$, both left or both right, is defined to be the action of $G_1 \times G_2$ on $M_1 \times M_2$ given by
$$\big((a_1, a_2), (m_1, m_2)\big) \mapsto \big(\Psi^1_{a_1}(m_1), \Psi^2_{a_2}(m_2)\big).$$
For $i = 1, 2$, the natural projections $M_1 \times M_2 \to M_i$ and $G_1 \times G_2 \to G_i$ yield morphisms of Lie group actions. The direct product of G-manifolds (M_i, Ψ^i), $i = 1, 2$, both left or both right, is defined to be the action of G on $M_1 \times M_2$ given by
$$\big(a, (m_1, m_2)\big) \mapsto \big(\Psi^1_a(m_1), \Psi^2_a(m_2)\big).$$
The natural projections $M_1 \times M_2 \to M_i$, $i = 1, 2$, are equivariant.
7. Let (M, G, Ψ) be a Lie group action, (H, ϱ) a Lie subgroup of G and (N, φ) a submanifold of M. One says that (N, φ) is invariant under (H, ϱ) if
$$\Psi_{\varrho(H)}\big(\varphi(N)\big) \subset \varphi(N).$$
In this case, and if additionally (N, φ) is initial, the restriction $\tilde{\Psi}$ of Ψ in domain to the submanifold $(H \times N, \varrho \times \varphi)$ of $G \times M$ and in range to the initial submanifold (N, φ) of M is an action of H on N, called the restriction of (M, G, Ψ) to

(H, ϱ) and (N, φ). It is determined by the relation $\varphi \circ \tilde{\Psi} = \Psi \circ (\varrho \times \varphi)$, that is, by the condition that (φ, ϱ) be a morphism of Lie group actions from $(N, H, \tilde{\Psi})$ to (M, G, Ψ).

Remark 6.1.3 If M is endowed with an additional structure, like a Riemannian metric, one may require the diffeomorphisms Ψ_a to respect this structure. This way, one obtains, for example, the notion of Riemannian Lie group action and Riemannian G-manifold (with G acting by isometries), symplectic Lie group action and symplectic G-manifold (with G acting by symplectomorphisms, cf. Chap. 8) and G-vector bundle (with G acting by vector bundle automorphisms).

Next, we introduce the basic notions associated with a Lie group action (M, G, Ψ), collect some of their elementary properties and describe the geometric structure arising. Two points $m_1, m_2 \in M$ are said to be conjugate under Ψ if $m_2 = \Psi_a(m_1)$ for some $a \in G$. Obviously, to be conjugate is an equivalence relation on M.

Definition 6.1.4 Let (M, G, Ψ) be a Lie group action.

1. The above equivalence classes are called the orbits of G under Ψ. The set of equivalence classes, equipped with the quotient topology, is called the orbit space of (M, G, Ψ) and is denoted by M/G.
2. For $m \in M$, the induced mapping $\Psi_m : G \to M$ is called the orbit mapping of m.
3. For $m \in M$, the subgroup $G_m := \{a \in G : \Psi_a(m) = m\}$ is called the stabilizer[3] of m.
4. The kernel of Ψ is the subgroup $\ker(\Psi) := \{a \in G : \Psi_a = \mathrm{id}_M\}$ of G.

The kernel of Ψ coincides with the kernel of the corresponding (anti-)homomorphism from G to the group of diffeomorphisms of M. By definition, $\ker(\Psi) = \bigcap_{m \in M} G_m$. The orbit through $m \in M$ is denoted by $G \cdot m$. It is given by

$$G \cdot m = \big\{\Psi_a(m) \in M : a \in G\big\}. \tag{6.1.6}$$

More generally, for a subset $N \subset M$ and a subgroup $H \subset G$ we denote

$$H \cdot N := \big\{\Psi_a(m) \in M : m \in N, a \in H\big\}. \tag{6.1.7}$$

Assume that Ψ is a left action. Since for $a \in G$, the orbit mapping Ψ_m satisfies $\Psi_m(ab) = \Psi_m(a)$ iff $b \in G_m$, it induces an injective mapping

$$\hat{\Psi}_m : G/G_m \to M, \qquad \hat{\Psi}_m(aG_m) := \Psi_m(a), \tag{6.1.8}$$

where G/G_m denotes the homogeneous space of left cosets. Since the natural projection $G \to G/G_m$ is a surjective submersion, by Remark 1.5.16, $\hat{\Psi}_m$ is smooth.

[3] Another common name is isotropy group of m.

Moreover, it is equivariant with respect to the action of G on G/G_m by the induced left translations $\hat{\mathrm{L}}_a$ and its image coincides with the orbit of m.[4] Thus, for every m, $\hat{\Psi}_m$ yields an equivariant bijection between G/G_m and the orbit through m. In the next section, this bijection will be used to induce a smooth structure on $G \cdot m$.

Proposition 6.1.5 *Let (M, G, Ψ) be a Lie group action.*

1. *The action Ψ is an open mapping.*
2. *The natural projection $\pi : M \to M/G$ is an open mapping.*
3. *M/G is locally compact and second countable.*
4. *G_m is a closed subgroup of G for all $m \in M$. The mapping $a \mapsto (\Psi_a)'_m$ defines a representation[5] of G_m on $\mathrm{T}_m M$, called the isotropy representation at m.*
5. *The kernel of Ψ is a closed normal subgroup of G. The mapping*

$$\hat{\Psi} : G/\ker(\Psi) \times M \to M, \qquad \hat{\Psi} \circ (\varrho \times \mathrm{id}_M) = \Psi,$$

where $\varrho : G \to G/\ker(\Psi)$ denotes the natural projection, is an action of $G/\ker(\Psi)$ on M and (id_M, ϱ) is a morphism of Lie group actions.

Proof 1. This follows from $\Psi(V \times U) = \bigcup_{a \in V} \Psi_a(U)$ and the fact that the mappings Ψ_a are homeomorphisms.

2. Let $U \subset M$ be open. Then, $\Psi_a(U)$ is open for all $a \in G$ and hence $\pi^{-1}(\pi(U)) = \bigcup_{a \in G} \Psi_a(U)$ is open. By construction of the quotient topology, then $\pi(U)$ is open.

3. Since M is locally compact, every $m \in M$ possesses a compact neighbourhood U. Since π is open, $\pi(U)$ is a neighbourhood of $\pi(m)$ in M/G. Since the image of a compact space under a continuous mapping is compact, $\pi(U)$ is compact. Hence, M/G is locally compact. Second countability holds by definition of the quotient topology.

4. For every $m \in M$, the stabilizer G_m is the preimage of m under the orbit mapping Ψ_m. The rest follows from points 5 and 7 of Example 6.1.2.

5. Being the kernel of a group (anti-)homomorphism, $\ker(\Psi)$ is normal. Being the intersection of the stabilizers of all points of M, it is also closed. Hence, by Proposition 5.7.8, $G/\ker(\Psi)$ is a Lie group. Since $\Psi_a = \mathrm{id}_M$ for all $a \in \ker(\Psi)$, $\hat{\Psi}$ is well-defined. Since ϱ is a submersion, $\hat{\Psi}$ is smooth by Remark 1.5.16. The rest is obvious. □

Next, we observe that for every $a \in G$,

$$G_{\Psi_a(m)} = a G_m a^{-1} \text{ (left action)}, \qquad G_{\Psi_a(m)} = a^{-1} G_m a \text{ (right action)}. \tag{6.1.9}$$

---

[4]In the case of a right action, $\Psi_m(ba) = \Psi_m(a)$ iff $b \in G_m$. Thus, $\hat{\Psi}_m$ is defined on the homogeneous space of right cosets by $\hat{\Psi}_m(G_m a) := \Psi_m(a)$. It is equivariant with respect to the induced right translations $\hat{\mathrm{R}}_a$.

[5]More precisely, a left (right) representation in case Ψ is a left (right) action.

That is, the stabilizers of any two points lying on the same orbit are conjugate.[6]

Definition 6.1.6 (Orbit type) Let (M, G, Ψ) be a Lie group action. The type of an orbit O is defined to be the conjugacy class of subgroups of G given by $\{G_m : m \in O\}$. The set of conjugacy classes of subgroups of G which appear as types of orbits under Ψ is called the set of orbit types of (M, G, Ψ). For a given orbit type $[H]$, the subset of M of orbit type $[H]$ will be denoted by $M_{[H]}$.

We note that

$$M_{[H]} = \{m \in M : G_m = aHa^{-1} \text{ for some } a \in G\}. \tag{6.1.10}$$

Definition 6.1.7 For a closed subgroup $H \subset G$, define

$$M_H := \{m \in M : H = G_m\}, \tag{6.1.11}$$

$$M^H := \{m \in M : H \subset G_m\}. \tag{6.1.12}$$

M_H and M^H are called the subset of isotropy type H and the subset of fixed points under H, respectively.

By construction,

$$M_{[H]} = G \cdot M_H. \tag{6.1.13}$$

Moreover, if H is compact, one has (Exercise 6.1.2)

$$M_H = M^H \cap M_{[H]}. \tag{6.1.14}$$

Finally, we introduce the following algebraic properties of Lie group actions.

Definition 6.1.8 A Lie group action is called

1. effective if the kernel is equal to $\{\mathbb{1}\}$,
2. free if all stabilizers are equal to $\{\mathbb{1}\}$,
3. transitive if it has a single orbit.

According to Proposition 6.1.5/5, by passing to the induced action of $G/\ker(\Psi)$, every action can be replaced by an effective action which has the same orbits. Moreover, it is obvious that free actions and transitive actions have just one single orbit type.

Example 6.1.9 Proofs are left to the reader (Exercise 6.1.3).

---

[6]Subgroups H_1 and H_2 of G are said to be conjugate if $H_2 = aH_1a^{-1}$ for some $a \in G$. To be conjugate is an equivalence relation on the set of subgroups of G. The equivalence classes are called conjugacy classes. The conjugacy class of a subgroup H in G will be denoted by $[H]$.

1. The kernel of the Lie group action associated with a representation coincides with the kernel of the representation. Hence, the action is effective iff the representation is faithful. The action is neither free nor transitive, because the origin is invariant. As an example, for the fundamental representation[7] of $G = \mathrm{SO}(3)$ on $\mathbb{R}^3$, the orbits are given by the spheres about the origin and the origin itself. Therefore, the orbit space may be identified with the closed half-line $[0, \infty)$. The stabilizer of a nonzero $\mathbf{x} \in \mathbb{R}^3$ consists of the rotations about $\mathbf{x}$, hence it is isomorphic to SO(2). Since it does not depend on the norm of $\mathbf{x}$, all stabilizers of nonzero vectors belong to the same orbit type. The stabilizer of the origin is the whole group.
2. The actions of G on itself by left and right translations are effective, free and transitive. The restrictions of these actions to a proper Lie subgroup $H \subset G$ are still effective and free but no longer transitive. The orbits are the right and left cosets of H, respectively. According to Theorem 5.7.2, the orbit space G/H carries in addition a smooth structure. The action of G on itself by inner automorphisms is neither free nor transitive, because the unit element is invariant. The kernel is given by the centre. Hence, the action is effective iff the centre is trivial. The orbits are called the conjugacy classes of elements of G. The stabilizer of $a \in G$ is the centralizer of a in G. Therefore, the orbit types are given by the conjugacy classes of those subgroups of G which are centralizers of elements of G. The orbit space is usually referred to as the adjoint quotient of G.
3. The action of G on the homogeneous space G/H, where $H \subset G$ is a nontrivial closed subgroup, by the induced left translations $\widehat{\mathrm{L}}_a$ is transitive but not free. The stabilizer of a coset aH, $a \in G$, is given by $G_{aH} = aHa^{-1}$. Consequently, the kernel of the action is given by $\bigcap_{a\in G} aHa^{-1}$. It lies between the intersection of the centre of G with H and H itself, the latter being the case when H is normal.
4. For the action of $G = \mathbb{R}$ on M defined by the flow Φ of a complete vector field X, the orbit mapping of a point $m \in M$ is given by the maximal integral curve Φ_m. The stabilizer is $G_m = \mathbb{R}$ in case $X_m = 0$, $G_m = \mathbb{Z}T$ in case the maximal integral curve Φ_m is periodic with period T, or $G_m = \{0\}$ otherwise. Since G is Abelian, the orbit types correspond bijectively to the subgroups which appear as stabilizers.

Remark 6.1.10 Every Lie group action (M, G, Ψ) induces a representation of G on $C^\infty(M)$ by the algebra automorphisms $f \mapsto \Psi^*_{a^{-1}} f$, $a \in G$. A function $f \in C^\infty(M)$ is called invariant under Ψ if it is a fixed point of this representation, that is, if $\Psi^*_a f = f$ for all $a \in G$. This means that f is equivariant with respect to the trivial action of G on $\mathbb{R}$. The invariant functions form a subalgebra which will be denoted by $C^\infty(M)^G$. Analogously, Ψ induces representations of G on the various types of smooth tensor fields by $T \mapsto \Psi_{a*}T$. T is said to be invariant if it is invariant under this representation. This is equivalent to being equivariant as a mapping from M to the corresponding tensor bundle, where the latter carries the induced action of Example 6.1.2/5. More generally, a section of a G-vector bundle is

[7]Given by matrix multiplication of SO(3)-matrices with elements of $\mathbb{R}^3$.

said to be invariant if it is invariant under the transport operators of the vector bundle automorphisms provided by the G-action, and this is equivalent to being equivariant with respect to the action of G on the bundle manifold and the induced action on the base manifold.

Exercises

6.1.1 Verify the axioms of a Lie group action for the examples of 6.1.2.
6.1.2 Prove Formula (6.1.14). Try to find a counterexample for noncompact H.
6.1.3 Verify the statements of Example 6.1.9.

6.2 Killing Vector Fields

Throughout this section, let (M, G, Ψ) be a Lie group action and let $\mathfrak{g}$ be the Lie algebra of G. For $A \in \mathfrak{g}$, we have

$$\Psi_{\exp(tA)} \circ \Psi_{\exp(sA)} = \Psi_{\exp(tA)\exp(sA)} = \Psi_{\exp((t+s)A)},$$

hence the assignment $(m, t) \mapsto \Psi_{\exp(tA)}(m)$ defines a one-parameter group of diffeomorphisms of M, that is, a complete flow.

Definition 6.2.1 (Killing vector field) The vector field on M defined by the flow $\Psi_{\exp(tA)}$ is called the Killing vector field generated by A. It will be denoted by A_*.

The value of A_* at $m \in M$ may be expressed in several equivalent ways:

$$(A_*)_m = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \Psi_{\exp(tA)}(m) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \Psi_m\big(\exp(tA)\big) = \Psi'_m(A). \qquad (6.2.1)$$

Proposition 6.2.2 *Let (M, G, Ψ) be a Lie group action.*

1. *For every $a \in G$ and $A \in \mathfrak{g}$, there holds $\Psi_{a*}A_* = (\mathrm{Ad}(a)A)_*$ in case Ψ is a left action and $\Psi_{a*}A_* = (\mathrm{Ad}(a^{-1})A)_*$ in case Ψ is a right action.*
2. *The mapping $\mathfrak{g} \to \mathfrak{X}(M)$ given by $A \mapsto A_*$ is an anti-homomorphism of Lie algebras if Ψ is a left action and a homomorphism if Ψ is a right action. The kernel of this mapping coincides with the Lie algebra of* $\ker(\Psi) \subset G$. *In particular, if Ψ is effective, the mapping $A \mapsto A_*$ is injective.*
3. *For $m \in M$, the Lie algebra of G_m is given by $\mathfrak{g}_m = \{A \in \mathfrak{g} : (A_*)_m = 0\}$. In particular, if Ψ is free, then A_* cannot have critical points.*

Proof 1. Due to Proposition 3.2.13/2, the flow of $\Psi_{a*}A_*$ is $\Psi_a \circ \Psi_{\exp(tA)} \circ \Psi_{a^{-1}}$, which is equal to $\Psi_{\exp(t\mathrm{Ad}(a)A)}$ if Ψ is a left action and to $\Psi_{\exp(t\mathrm{Ad}(a^{-1})A)}$ if Ψ is a right action.

2. We give the proof for the case where Ψ is a left action. Linearity is obvious from (6.2.1). According to Proposition 3.3.2/2 and point 1, for $A, B \in \mathfrak{g}$, we have

$$[A_*, B_*]_m = (\mathscr{L}_{A_*} B_*)_m = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 (\Psi_{\exp(-tA)*} B_*)_m = \Psi'_m\left(\frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \big(\mathrm{Ad}\big\{\exp(-tA)\big\}B\big)\right).$$

Now, (5.4.2) and Proposition 5.4.3 yield $[A_*, B_*] = -[A, B]_*$. Furthermore, A belongs to the Lie algebra of $\ker(\Psi)$ iff $\Psi_{\exp(tA)} = \mathrm{id}_M$ for all t. Since $\Psi_{\exp(tA)}$ is the flow of A_*, this is equivalent to $A_* = 0$.

3. Let $A \in \mathfrak{g}$. By Theorem 5.6.8, A belongs to the Lie algebra $\mathfrak{g}_m$ of G_m iff $\Psi_{\exp(tA)}(m) = m$ for all t, that is, iff m is an equilibrium of A_*. □

Remark 6.2.3 According to point 3, m is an equilibrium of A_* for all $A \in \mathfrak{g}_m$. Thus, by Formula (3.6.2), the representation of $\mathfrak{g}_m$ on $\mathrm{T}_m M$ induced by the isotropy representation of G_m is given by

$$\mathfrak{g}_m \to \mathrm{End}(\mathrm{T}_m M), \qquad A \mapsto \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 (\Psi_{\exp(tA)})'_m = \mathrm{Hess}_m(A_*).$$

It is referred to as the isotropy representation of $\mathfrak{g}_m$.

Proposition 6.2.4 (Transformation properties)

1. *Let (M_i, G_i, Ψ^i), $i = 1, 2$, be Lie group actions and let $\varphi : M_1 \to M_2$, $\varrho : G_1 \to G_2$ define a morphism. The Killing vector field on M_1 generated by an element A of the Lie algebra of G_1 is φ-related to the Killing vector field on M_2 generated by* $\mathrm{d}\varrho(A)$.
2. *Let (M_i, Ψ^i), $i = 1, 2$, be G-manifolds and let $\varphi : M_1 \to M_2$ be equivariant. For every $A \in \mathfrak{g}$, the Killing vector fields $A_*^{M_i}$ on M_i generated by A are φ-related.*

If φ is a diffeomorphism, points 1 and 2 yield, respectively,

$$\big(\mathrm{d}\varrho(A)\big)_* = \varphi_* A_*, \qquad A_*^{M_2} = \varphi_* A_*^{M_1}. \tag{6.2.2}$$

Proof According to Proposition 5.3.6, for $m \in M_1$, we have

$$\begin{aligned}(\varphi' \circ A_*)(m) &= \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \varphi \circ \Psi^1_{\exp_{G_1}(tA)}(m)\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \Psi^2_{\varrho(\exp_{G_1}(tA))} \circ \varphi(m)\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \Psi^2_{\exp_{G_2}(t\mathrm{d}\varrho(A))} \circ \varphi(m).\end{aligned}$$

This yields the first assertion. The second one follows by letting $G_1 = G_2 = G$ and $\varrho = \mathrm{id}_G$. □

Example 6.2.5 Proofs are left to the reader (Exercise 6.2.1).

1. For the left action of G associated with a representation $\varrho : G \to \mathrm{GL}(V)$ on a finite-dimensional $\mathbb{K}$-vector space V, the Killing vector field A_* is the linear vector field on V corresponding to the endomorphism $\mathrm{d}\varrho(A)$. In particular, for the adjoint and the coadjoint actions we obtain
$$A_*^{\mathrm{Ad}} = \mathrm{ad}(A), \qquad A_*^{\mathrm{Ad}^*} = \mathrm{ad}^*(A), \quad A \in \mathfrak{g}, \tag{6.2.3}$$
and for the identical representation of a classical group $G \subset \mathrm{GL}(V)$, under the natural identification of $\mathfrak{g}$ with a Lie subalgebra of $\mathfrak{gl}(V)$, we have $A_* = A$.

2. For the action of G on itself by right translations we get $A_* = A$. That is, the homomorphism $\mathfrak{g} \to \mathfrak{X}(G)$ of Proposition 6.2.2/2 coincides with the natural inclusion mapping. For the action of G on itself by left translations, A_* is the right-invariant vector field on G generated by $A \in \mathrm{T}_{\mathbb{1}}G$. According to Remark 5.2.21/3, this means that the anti-homomorphism $\mathfrak{g} \to \mathfrak{X}(G)$ is given by the natural inclusion mapping, composed with $-\mathrm{inv}_*$. For the action of G on itself by inner automorphisms, A_* is the difference between the right and the left invariant vector fields generated by A and hence the anti-homomorphism $\mathfrak{g} \to \mathfrak{X}(G)$ is given by $-(1 + \mathrm{inv}_*)$.
3. Let $H \subset G$ be a closed subgroup and let $\pi : G \to G/H$ denote the natural projection. For the action of G on the homogeneous space G/H by induced left translations, Proposition 6.2.4/2 and the previous result yield $(A_*)_{\pi(a)} = \pi' \circ \mathrm{R}'_a A$ for all $a \in G$ and $A \in \mathfrak{g}$.
4. For the action of $G = \mathbb{R}$ on M defined by the flow of a complete vector field X, under the identification of the Lie algebra of $\mathbb{R}$ with $\mathbb{R}$ itself, the homomorphism $\mathfrak{g} \to \mathfrak{X}(M)$ is given by $s \mapsto s_* = sX$.

By analogy with Lie group actions and their morphisms one defines

Definition 6.2.6 (Lie algebra action)

1. An action of a Lie algebra $\mathfrak{g}$ on a smooth manifold M is a mapping $\psi : \mathfrak{g} \to \mathfrak{X}(M)$ which is either an anti-homomorphism or a homomorphism of Lie algebras. In the first case, ψ is called a left action and in the second case a right action. The triple $(M, \mathfrak{g}, \psi)$ is referred to as a Lie algebra action and the pair (M, ψ) as a $\mathfrak{g}$-manifold.
2. A morphism of Lie algebra actions $(M_1, \mathfrak{g}_1, \psi_1)$ and $(M_2, \mathfrak{g}_2, \psi_2)$ (both left or both right) consists of a smooth mapping $\varphi : M_1 \to M_2$ and a Lie algebra homomorphism $\varrho : \mathfrak{g}_1 \to \mathfrak{g}_2$ such that the vector fields $\psi_1(A)$ and $\psi_2 \circ \varrho(A)$ are φ-related for all $A \in \mathfrak{g}_1$. If $\mathfrak{g}_1 = \mathfrak{g}_2 = \mathfrak{g}$ and $\varrho = \mathrm{id}$, φ is called equivariant or, equivalently, a morphism of $\mathfrak{g}$-manifolds.

Note that the induced mapping $\mathfrak{g} \times M \to \mathrm{T}M$, $(A, m) \mapsto (\psi(A))_m$, is automatically smooth. Using the notion of Lie algebra action, Propositions 6.2.2/2 and 6.2.4 may be restated as follows (Exercise 6.2.3).

Corollary 6.2.7

1. *The Killing vector fields of a left or right Lie group action* (M, G, Ψ) *induce a left or right action, respectively, of the Lie algebra* $\mathfrak{g}$ *of* G *on* M.
2. *If* (φ, ϱ) *is a morphism of Lie group actions, then* $(\varphi, \mathrm{d}\varrho)$ *is a morphism of the corresponding Lie algebra actions. If* φ *is a morphism of* G*-manifolds, then it is also a morphism of the corresponding* $\mathfrak{g}$*-manifolds.*

Next, consider the distribution $D^{\mathfrak{g}}$ on M spanned by the Killing vector fields, that is,

$$D^{\mathfrak{g}}_m = \big\{(A_*)_m \in \mathrm{T}_m M : A \in \mathfrak{g}\big\}.$$

According to points 1 and 3 of Proposition 6.2.2, we have

$$D^{\mathfrak{g}}_{\Psi_a(m)} = \Psi'_a D^{\mathfrak{g}}_m \tag{6.2.4}$$

for all $a \in G$, and $\dim(D^{\mathfrak{g}}_m) = \dim(G) - \dim(G_m)$, respectively. In particular, $D^{\mathfrak{g}}$ need not be regular; this is so, however, if Ψ is free or transitive. We will show that $D^{\mathfrak{g}}$ is integrable and that the integral manifolds are given by the orbits. For that purpose, recall the mapping $\hat{\Psi}_m : G/G_m \to M$, defined by (6.1.8).

Theorem 6.2.8 (Orbit Theorem) *Let (M, G, Ψ) be a Lie group action. For every $m \in M$, $(G/G_m, \hat{\Psi}_m)$ is an initial submanifold of M whose connected components are maximal integral manifolds of $D^{\mathfrak{g}}$. In particular, $D^{\mathfrak{g}}$ is integrable.*

Proof We give the proof for the case of a left action. Let $m \in M$ and let $\pi^G : G \to G/G_m$ denote the natural projection. First, we show that $\hat{\Psi}_m$ is an injective immersion. Injectivity was already noted before. Since $\hat{\Psi}_m \circ \pi^G = \Psi_m$, for $(\hat{\Psi}_m)'_{aG_m}$ to be injective for all $aG_m \in G/G_m$ it suffices to show that

$$\ker(\Psi_m)'_a \subset \ker\big(\pi^G\big)'_a$$

for all $a \in G$. To see this, let $A \in \mathfrak{g}$ be such that $(\Psi_m)' A_a = 0$. Due to

$$(\Psi_m)' A_a = (\Psi_m)' \circ (\mathrm{L}_a)' A_{\mathbb{1}} = (\Psi_a)' \circ (\Psi_m)' A_{\mathbb{1}} = (\Psi_a)'(A_*)_m, \tag{6.2.5}$$

and since $(\Psi_a)'$ is bijective, we obtain $(A_*)_m = 0$ and hence $A \in \mathfrak{g}_m$ by Proposition 6.2.2/3. Then, $\exp(tA) \in G_m$ for all $t \in \mathbb{R}$ and hence

$$\big(\pi^G\big)'_a(A_a) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \pi^G\big(a \exp(tA)\big) = 0.$$

This proves that $\hat{\Psi}_m$ is an injective immersion.

Second, we show that the connected components of $(G/G_m, \hat{\Psi}_m)$ are maximal integral manifolds of $D^{\mathfrak{g}}$. For simplicity, we may assume that G and hence G/G_m are connected. Due to π^G being a submersion, for $(G/G_m, \hat{\Psi}_m)$ to be an integral manifold of $D^{\mathfrak{g}}$ it suffices that $\mathrm{im}((\Psi_m)'_a) = D^{\mathfrak{g}}_{\Psi_a(m)}$ for all $a \in G$. The latter is true, indeed, because using (6.2.5) and Proposition 6.2.2/1, for $A \in \mathfrak{g}$ one finds

$$(\Psi_m)' A_a = (\Psi_a)'(A_*)_m = (\Psi_{a*} A_*)_{\Psi_a(m)} = \big(\big\{\mathrm{Ad}(a)A\big\}_*\big)_{\Psi_a(m)}$$

and hence $(A_*)_{\Psi_a(m)} = (\Psi_m)'(\mathrm{Ad}(a^{-1})A)_m$. Maximality follows from the fact that the images of the connected components of the submanifolds $(G/G_m, \hat{\Psi}_m)$, $m \in M$, establish a disjoint decomposition of M. Finally, since m was arbitrary, we can conclude that $D^{\mathfrak{g}}$ is integrable. Hence, Theorem 3.5.15 yields that $(G/G_m, \hat{\Psi}_m)$ is initial. □

As a consequence of Theorem 6.2.8, according to Remark 1.6.13/5, the smooth structures induced on an orbit O by the submanifolds $(G/G_m, \hat{\Psi}_m)$, $m \in O$, coincide for all m, so that we can view O itself as a submanifold.[8] Then, Theorem 6.2.8 implies

[8]There is also a direct argument proving this, see Exercise 6.2.4.

Corollary 6.2.9 *The orbits of Ψ are initial submanifolds of M whose connected components are maximal integral manifolds of $D^{\mathfrak{g}}$.*[9]

Remark 6.2.10

1. Let (M, G, Ψ) be a Lie group action, let O be an orbit and let $m \in O$. Consider the isotropy representation of G_m on $\mathrm{T}_m M$, cf. Proposition 6.1.5/4. By Proposition 6.2.2/1,
$$(\Psi_a)'_m (A_*)_m = (\Psi_{a*} A_*)_m = \big(\big(\mathrm{Ad}(a)A\big)_*\big)_m. \tag{6.2.6}$$
Therefore, the subspace $\mathrm{T}_m O$ of $\mathrm{T}_m M$ is invariant and the isotropy representation descends to a representation on $\mathrm{N}_m O$, which for reasons that will become clear later is called the slice representation at m. By construction, the natural projection $\mathrm{T}_m M \to \mathrm{N}_m O$ intertwines the isotropy representation with the slice representation.
On the other hand, the isotropy representation restricted to $\mathrm{T}_m O$ can be brought to the following normal form. By Theorem 6.2.8, the mapping $(\Psi_m)'_{\mathbb{1}} : \mathfrak{g} \to \mathrm{T}_m M$ has image $\mathrm{T}_m O$. By point 3 of Proposition 6.2.2 it has kernel $\mathfrak{g}_m$. Hence, this mapping induces a vector space isomorphism
$$\mathrm{T}_m O \cong \mathfrak{g}/\mathfrak{g}_m. \tag{6.2.7}$$
In view of (6.2.1), Formula (6.2.6) implies that this isomorphism intertwines the representation of G_m on $\mathfrak{g}/\mathfrak{g}_m$ induced by the adjoint representation with the isotropy representation on $\mathrm{T}_m O$ (Exercise 6.2.5).
2. Let (M, G, Ψ) be a Lie group action and let O be an orbit. Since O is invariant under Ψ and since it is an initial submanifold of M, Ψ restricts to a transitive action of G on O. Since the connected components of O are integral manifolds of $D^{\mathfrak{g}}$, (6.2.4) implies that the action of G on $\mathrm{T}M$ induced by Ψ leaves invariant the submanifold $\mathrm{T}O$.[10] The restriction of this action to the invariant submanifold $(\mathrm{T}M)_{\restriction O}$ induces an action Ψ^{N} of G on the normal bundle $\mathrm{N}O = (\mathrm{T}M)_{\restriction O}/\mathrm{T}O$ by the vector bundle automorphisms[11]
$$\Psi^{\mathrm{N}}_a\big([X]\big) := \big[\Psi'_a(X)\big], \tag{6.2.8}$$
where $a \in G$ and $X \in \mathrm{T}M$. The natural projections $(\mathrm{T}M)_{\restriction O} \to \mathrm{N}O \to O$ are equivariant. Equivariance of the second projection implies that Ψ^{N} restricts to an action of G_m on $\mathrm{N}_m O$ for every $m \in O$. This action coincides with the slice representation at m.
3. A manifold carrying a transitive Lie group action is said to be homogeneous. The Orbit Theorem 6.2.8 yields that every homogeneous manifold is diffeomorphic to G/H for some Lie group G and some closed subgroup H of G. The

[9] In particular, according to Proposition 3.5.21, they form the foliation associated with $D^{\mathfrak{g}}$.

[10] Both of these statements follow as well by viewing O as the image of the submanifold $(G/G_m, \hat{\Psi}_m)$ and using that $\hat{\Psi}_m$ is G-equivariant.

[11] Smoothness follows from the existence of local sections of the submersion $(\mathrm{T}M)_{\restriction O} \to \mathrm{N}O$.

diffeomorphism provided by this theorem is obviously equivariant with respect to the action of G on G/H by the induced translations $\hat{\mathrm{L}}_a$ (in the case of a left action) or $\hat{\mathrm{R}}_a$ (in the case of a right action). Hence, it is an isomorphism of Lie group actions. In this sense, all of the homogeneous spaces discussed in Sect. 5.7 constitute homogeneous manifolds. This explains their name. In the spirit of Remark 6.1.3, the notion of homogeneous manifold may be further specialized, e.g., to that of homogeneous Riemannian manifold (a Riemannian manifold carrying a transitive Lie group action by isometries).

Exercises

6.2.1 Provide proofs for the statements of Example 6.2.5.

6.2.2 Calculate the Killing vector fields for the natural action
(a) of SO(2) on $\mathbb{R}^2$,
(b) of SO(3) on $\mathbb{R}^3$,
(c) of SU(2) on $\mathbb{C}^2$.
For SO(3) and SU(2), use the bases given in Example 5.2.8. Read off the Killing vector fields on the corresponding unit spheres in $\mathbb{R}^3$ and $\mathbb{C}^2$, respectively.

6.2.3 Verify the statements of Corollary 6.2.7.

6.2.4 Let (M, G, Ψ) be a left Lie group action. Show that for every $m \in M$ and $a \in G$, $\mathrm{R}_{a^{-1}}$ descends to a diffeomorphism $\varphi : G/G_m \to G/G_{\Psi_a(m)}$ satisfying $\hat{\Psi}_{\Psi_a(m)} \circ \varphi = \hat{\Psi}_m$. Use this for showing that the smooth structures induced on an orbit O by means of the submanifolds $(G/G_m, \hat{\Psi}_m)$, $m \in O$, do not depend on the choice of m. (This proof does not use the fact that orbits are initial submanifolds.)

6.2.5 Use Proposition 6.2.2/1 and Formula (6.2.1) to show that the vector space isomorphism $\mathfrak{g}/\mathfrak{g}_m \cong \mathrm{T}_m O$ induced by the mapping $(\Psi_m)'_{\mathbb{1}} : \mathfrak{g} \to \mathrm{T}_m M$ intertwines the representation of G_m on $\mathfrak{g}/\mathfrak{g}_m$ induced by the adjoint representation with the isotropy representation on $\mathrm{T}_m O$, cf. Remark 6.2.10/1.

6.3 Proper Actions

Lie group actions possessing the following topological property behave particularly well. For example, they allow for local models in terms of vector bundles. This has important consequences for the structure of both the manifold carrying the action and the orbit space.

Definition 6.3.1 (Proper action)

1. A continuous mapping of locally compact Hausdorff spaces (in particular, of manifolds) is proper if the preimages of compact subsets are compact.[12]

[12] More generally, this definition applies when the domain is Hausdorff and the range is locally compact Hausdorff [53]. Compactness of subsets is understood with respect to the relative topology.

2. A Lie group action (M, G, Ψ) is called proper if the following mapping is proper:

$$\Psi_{\text{ext}} : G \times M \to M \times M, \qquad \Psi_{\text{ext}}(a, m) := \big(\Psi_a(m), m\big). \tag{6.3.1}$$

We give two equivalent characterizations of proper mappings. Note that since manifolds are metrizable, one may use the criterion that a subset is compact iff every sequence contains a subsequence converging in that subset.

Lemma 6.3.2 *For a continuous mapping $\varphi : M \to N$ between manifolds, the following statements are equivalent.*

1. *The mapping φ is proper.*
2. *Any sequence in M whose image converges in N contains a convergent subsequence.*
3. *The mapping φ is closed and $\varphi^{-1}(p)$ is compact for all $p \in N$.*

Proof 1 ⇒ 2: Let $\{m_n\}$ be a sequence in M such that $\{\varphi(m_n)\}$ converges in N. Since N is locally compact, there exist $n_0 \in \mathbb{N}$ and a compact subset $K \subset N$ such that $\varphi(m_n) \in K$ for all $n \geq n_0$. Then, $m_n \in \varphi^{-1}(K)$ for all $n \geq n_0$. Since φ is proper, $\varphi^{-1}(K)$ is compact, hence $\{m_n\}$ contains a convergent subsequence.

2 ⇒ 3: The mapping φ is closed: let $K \subset M$ be closed. Let $\{p_n\}$ be a sequence in $\varphi(K)$ which converges to some $p \in N$. For every n there is $m_n \in K$ such that $p_n = \varphi(m_n)$. By point 2, $\{m_n\}$ contains a subsequence converging to some $m \in M$. Since K is closed, $m \in K$. Then, $p = \varphi(m) \in \varphi(K)$. Thus, $\varphi(K)$ is closed. Compactness of $\varphi^{-1}(p)$ for $p \in N$ is obvious.

3 ⇒ 1: Let $K \subset N$ be compact and let $\{U_i : i \in I\}$ be an open covering of $\varphi^{-1}(K)$. Let $p \in K$. Since $\{U_i : i \in I\}$ is an open covering of $\varphi^{-1}(p)$ and since the latter is compact, there is a finite subcovering $\{U_i : i \in I_p\}$. Let V_p denote the complement in N of the image under φ of the complement of $\bigcup_{i \in I_p} U_i$ in M. By construction, $p \in V_p$. Since φ is closed, V_p is open. Hence, $\{V_p : p \in K\}$ is an open covering of K. Since K is compact, there is a finite subcovering labelled by $p_1, \dots, p_r$. Then, $\{U_i : i \in I_{p_1} \cup \cdots \cup I_{p_r}\}$ is a finite subcovering of $\{U_i : i \in I\}$. □

Lemma 6.3.2 implies the following equivalent characterizations of a proper action.

Corollary 6.3.3 *For a Lie group action (M, G, Ψ), the following statements are equivalent.*

1. *The action is proper.*
2. *If $\{a_n\}$ is a sequence in G and $\{m_n\}$ is a sequence in M such that $\{m_n\}$ and $\{\Psi_{a_n}(m_n)\}$ converge, then $\{a_n\}$ contains a convergent subsequence.*
3. *The mapping Ψ_{ext} is closed and all stabilizers of Ψ are compact.*

Proper actions have the following elementary topological properties.

Proposition 6.3.4 *Let (M, G, Ψ) be a proper Lie group action.*

1. *The restriction of Ψ to the product of a closed subgroup $H \subset G$ with an H-invariant initial*[13] *submanifold is proper.*
2. *The orbit mapping Ψ_m is proper for every $m \in M$.*
3. *The orbit space M/G is Hausdorff.*

Proof 1. We apply Corollary 6.3.3/2. Without loss of generality, we may assume that the subgroup and the submanifold are given by subsets. Thus, let $H \subset G$ be a closed subgroup and let $N \subset M$ be an H-invariant initial submanifold. Let $\{a_n\}$ and $\{m_n\}$ be sequences in H and N such that $\{m_n\}$ and $\{\Psi_{a_n}(m_n)\}$ converge in N. Since N is a submanifold, the natural inclusion mapping $N \to M$ is smooth and hence continuous. Therefore, $\{m_n\}$ and $\{\Psi_{a_n}(m_n)\}$ converge in M. Since Ψ is proper, $\{a_n\}$ contains a subsequence which converges in G. Since H is closed, the limit belongs to H.

2. We apply Lemma 6.3.2/2. Let $\{a_n\}$ be a sequence in G such that $\{\Psi_m(a_n)\}$ converges. Then, the sequences $\{a_n\}$ and $\{m_n = m\}$ satisfy the assumption of Corollary 6.3.3/2. Since Ψ is proper, it follows that $\{a_n\}$ contains a convergent subsequence.

3. The argument is analogous to that for homogeneous spaces, cf. Lemma 5.7.1/3. Denote the natural projection by $\pi : M \to M/G$. Let $m_1, m_2 \in M$ be such that $\pi(m_1) \neq \pi(m_2)$. Then, $(m_1, m_2) \notin \mathrm{im}(\Psi_{\mathrm{ext}})$. Since $\mathrm{im}(\Psi_{\mathrm{ext}})$ is closed in $M \times M$, there exist open neighbourhoods U_i of m_i in M such that $(U_1 \times U_2) \cap \mathrm{im}(\Psi_{\mathrm{ext}}) = \varnothing$. Then, $\pi(U_1) \cap \pi(U_2) = \varnothing$. Since, by Proposition 6.1.5/2, the mapping π is open, $\pi(U_i)$ is a neighbourhood of $\pi(m_i)$, $i = 1, 2$. □

The Orbit Theorem 6.2.8 and point 2 of Proposition 6.3.4 imply

Corollary 6.3.5 *The orbits of a proper Lie group action are closed embedded submanifolds.*

Proof Let $m \in M$. By Proposition 6.3.4/2 and Lemma 6.3.2/3, Ψ_m is closed. Hence, the image $\Psi_m(G)$ is closed and the induced mapping $\hat{\Psi}_m : G/G_m \to M$ is closed. Since the latter is bijective onto $\Psi_m(G)$, it is open onto $\Psi_m(G)$ in the relative topology induced from M. □

Remark 6.3.6 Let (M, G, Ψ) be a proper Lie group action, let O be an orbit and let $m \in O$. The fact that G_m is compact has the following immediate consequences for the isotropy representation.

1. By Proposition 5.5.6, $\mathrm{T}_m M$ admits a G_m-invariant scalar product.
2. The orthogonal complement $(\mathrm{T}_m O)^\perp$ of $\mathrm{T}_m O$ in $\mathrm{T}_m M$ with respect to this scalar product is invariant under the isotropy representation. Hence, the isotropy representation decomposes into the direct sum

$$\mathrm{T}_m M \cong \mathrm{T}_m O \oplus (\mathrm{T}_m O)^\perp .$$

[13] The assumption that the submanifold be initial is made to ensure that the restricted action is smooth, cf. Example 6.1.2/7.

3. The natural projection $\mathrm{T}_m M \to \mathrm{N}_m O$ restricts to an isomorphism $(\mathrm{T}_m O)^\perp \to \mathrm{N}_m O$ which intertwines the isotropy representation induced on $(\mathrm{T}_m O)^\perp$ with the slice representation.
4. Points 2 and 3 imply that there exists a vector space isomorphism

$$\mathrm{T}_m M \cong \mathrm{T}_m O \oplus \mathrm{N}_m O \tag{6.3.2}$$

intertwining the isotropy representation with the direct sum of the isotropy representation on $\mathrm{T}_m O$ and the slice representation.

Next, we discuss the tool of averaging. Depending on whether Ψ is a left or a right action, choose a left or a right-invariant volume form $\mathrm{d}a$ on G, cf. Corollary 5.5.4. Let T be a smooth tensor field on M of type (p, q) with compact support. Due to Proposition 6.3.4/2, for every $m \in M$, the smooth mapping $G \to (\mathbb{T}^q_p M)_m$ given by $a \mapsto (\Psi_{a*}T)_m$ has compact support $(\Psi_m^{-1}(\mathrm{supp}(T)))^{-1} \subset G$. Hence, one can define a mapping $T^G : M \to \mathbb{T}^q_p M$ by

$$T^G_m := \int_G (\Psi_{a*}T)_m \,\mathrm{d}a,$$

where the integrand is a differential form on G of maximal degree with values in the finite-dimensional real vector space $(\mathbb{T}^q_p M)_m$. The reader may convince himself that the local representative of T^G with respect to a chart (U, κ) on M and the induced chart on $\mathbb{T}^q_p M$ is given by a system of real-valued functions on $\kappa(U)$ each of which is obtained by integrating a compactly supported smooth function on $G \times \kappa(U)$ over the first variable. Thus, T^G is a smooth tensor field of the same type as T. By construction, T^G is invariant and $\mathrm{supp}(T^G) = G \cdot \mathrm{supp}(T)$. We call T^G the average of T with respect to Ψ. Since the choice of $\mathrm{d}a$ is unique up to a nonzero real factor, so is the average. If G is compact, one may fix this factor by requiring that G have unit volume. Note that taking the average is a linear operation but it is obviously not compatible with the multiplication of tensor fields.

Proposition 6.3.7 *Let (M, G, Ψ) be a proper Lie group action.*

1. *The invariant smooth functions separate the points of M/G.*[14]
2. *Every covering of M by invariant open subsets admits a countable subordinate partition of unity by invariant smooth functions.*
3. *There exists an invariant Riemannian metric on M.*

Proof Let $\pi : M \to M/G$ denote the natural projection.

1. Let $m_1, m_2 \in M$ be such that $\pi(m_1) \neq \pi(m_2)$. Since the orbit $G \cdot m_2$ is closed, using a local chart at m_1 one can construct a smooth function $f \geq 0$ on M with compact support such that $f(m_1) = 1$ and $\mathrm{supp}(f) \cap (G \cdot m_2) = \varnothing$. Then, the average satisfies $f^G(m_1) > 0$ and $f^G(m_2) = 0$.

[14] That is, for all $m_1, m_2 \in M$ such that $G \cdot m_1 \neq G \cdot m_2$, there is $f \in C^\infty(M)^G$ with $f(m_1) \neq f(m_2)$.

2. Let $\{U_j\}$ be a covering of M by invariant open subsets. Then, $\{\pi(U_j)\}$ is an open covering of M/G. Since, by Propositions 6.1.5/3 and 6.3.4/3, M/G is locally compact, second countable and Hausdorff, there exists a countable subordinate open covering $\{V_i : i \in I \subset \mathbb{N}\}$ which is locally finite, see for example [302, Lemma 1.9]. Then, $\{\pi^{-1}(V_i) : i \in I\}$ is a locally finite covering of M by invariant open subsets, subordinate to $\{U_j\}$. According to Proposition 1.3.7, there exists a partition of unity $\{f_{i,\alpha} : (i,\alpha) \in A \subset I \times \mathbb{N}\}$ with compact supports such that $\mathrm{supp}(f_{i,\alpha}) \subset \pi^{-1}(V_i)$ for all $(i,\alpha) \in A$. By passing to the averages $f_{i,\alpha}^G$ we obtain a family of invariant smooth functions. However, these functions need no longer add up to 1 and the family of their supports need no longer be locally finite. This can be remedied by an appropriate summation as follows. For $f \in C^\infty(M)$, let $f^{(k)}(m)$ denote the k-th order tangent mapping at m, viewed as a k-linear mapping $\mathrm{T}_m M \times \cdots \times \mathrm{T}_m M \to \mathbb{R}$, and let $\|f^{(k)}(m)\|$ denote the operator norm of this mapping with respect to a chosen Riemannian metric on M, cf. Proposition 4.4.2. Define

$$C_{i,\alpha} := \sup\left\{\left\|f_{i,\alpha}^{(k)}(m)\right\| : 0 \leq k \leq \alpha\right\}, \quad f_i(m) := \sum_{\alpha=0}^{\infty} 2^{-\alpha} C_{i,\alpha}^{-1} f_{i,\alpha}^G(m).$$

The latter series converges absolutely for all $m \in M$, and so do all formal derivatives, hence f_i is well-defined and smooth.[15] By construction, f_i is invariant and $\mathrm{supp}(f_i) \subset \pi^{-1}(V_i)$. In particular, the family $\{\mathrm{supp}(f_i) : i \in I\}$ is a locally finite covering of M. Hence, $\sum_{i\in I} f_i(m)$ is well-defined for all m and by dividing $f_i(m)$ by this sum one finally obtains the desired partition of unity.

3. Choose a countable atlas $\{(U_i, \kappa_i) : i \in I\}$ on M such that the closures of the U_i are compact and carry out the construction of the symmetric second-order covariant tensor fields g_i, $i \in I$, as in the proof of Proposition 4.4.2. Averaging yields invariant tensor fields g_i^G. As noted in the proof of point 2, the open covering $\{\pi(\mathrm{supp}(\mathrm{g}_i))\}$ of M/G admits a subordinate countable, locally finite covering $\{V_j : j \in J\}$. For each j, choose i_j so that $V_j \subset \mathrm{supp}(g_{i_j})$ and define $\mathrm{g}(m) := \sum_{j\in J} \mathrm{g}_{i_j}^G(m)$. This is an invariant smooth second-order covariant tensor field on M. Since on the interior of their supports, the g_i are Riemannian metrics, and since averaging does not affect this property, g is a Riemannian metric.[16] □

Example 6.3.8

1. Due to Corollary 6.3.3/2, for an action to be proper it suffices that G be compact. If M is compact, then this condition is also necessary. In particular, the action induced by the flow of a complete vector field on a compact manifold cannot be proper.
2. The action associated with a representation of a Lie group G on a $\mathbb{K}$-vector space V is proper iff G is compact, because the stabilizer of the origin is G. This applies, in particular, to the adjoint representation.

[15] In fact, the construction is designed so that the sequence of smooth functions, given by the partial sums, converges in an appropriate topology on $C^\infty(M)$. For the latter, see e.g. [211].

[16] For an alternative proof, see Exercise 6.3.1.

3. The action of G on itself by left or right translation is proper: let $\{a_n\}$ and $\{b_n\}$ be sequences in G such that $b_n \to b$ and $a_n b_n \to \tilde{b}$. Then, $a_n = (a_n b_n) b_n^{-1} \to \tilde{b} b^{-1}$. The action of G on itself by inner automorphisms is proper iff G is compact, because the stabilizer of the unit element is G. By Proposition 6.3.4/1, these statements carry over to the corresponding induced actions of a closed subgroup on G.

Remark 6.3.9 Let (M_1, Ψ^1) and (M_2, Ψ^2) be G-manifolds and let $\varphi : M_1 \to M_2$ be equivariant. If Ψ^2 is proper, then Ψ^1 is proper, too: let $\{a_n\}$ and $\{m_n\}$ be sequences in G and M_1, respectively, such that $\{m_n\}$ and $\{\Psi^1_{a_n}(m_n)\}$ converge. Then, $\{\varphi(m_n)\}$ and $\{\Psi^2_{a_n}(\varphi(m_n))\}$ also converge and thus $\{a_n\}$ contains a convergent subsequence. This has the following consequences.

1. If an action is proper, so is the induced action on the tensor bundles.
2. The direct product of G-manifolds is proper if one of the factors is proper.

Exercises

6.3.1 Prove assertion 3 of Proposition 6.3.7, using the following ingredients: a Riemannian metric on M, a partition of unity $\{f_i : i \in I \subset \mathbb{N}\}$ such that $\mathrm{supp}(f_i)$ is compact and contained in the interior of $\mathrm{supp}(f_{i+1})$ for all i, and an invariant partition of unity subordinate to the covering of M by the interiors of $\Psi_G(\mathrm{supp}(f_i))$.

6.3.2 Let (M, G, Ψ) be a Lie group action with G compact and let $m \in M$ be a fixed point of Ψ, i.e., $G_m = G$. Show that every neighbourhood of m contains an invariant open neighbourhood.

6.3.3 Let (M, G, Ψ) be a proper Lie group action. Show that for every $m \in M$, the isotropy representation at m admits an invariant scalar product.

6.4 The Tubular Neighbourhood Theorem

The Tubular Neighbourhood Theorem, or Slice Theorem, is one of the basic tools in the theory of proper Lie group actions. It is due to Montgomery and Yang [215] and Mostow [227] for compact Lie group actions and to Koszul [172] and Palais [233, 234] for proper Lie group actions.[17]

Recall that a G-vector bundle is a vector bundle endowed with an action of G by vector bundle automorphisms.

Definition 6.4.1 (Tubular Neighbourhood and Slice) Let (M, G, Ψ) be a Lie group action and let $O \subset M$ be an orbit.

[17] The special case of a free compact Lie group action is due to Gleason [107].

1. A tubular neighbourhood of O is an equivariant diffeomorphism $\chi : U \to E$ from an open invariant neighbourhood U of O in M onto an open neighbourhood V of the zero section s_0 in a G-vector bundle E over O such that $\chi_{\restriction O} = s_0$.
2. For $m \in O$, the G_m-invariant neighbourhood $V_m := V \cap E_m$ of the origin in E_m is called the linear slice of χ at m and the G_m-invariant embedded submanifold $U_m := \chi^{-1}(V_m)$ of M is called the slice of χ at m.

One says that the G-vector bundle E provides a local normal form for Ψ near O.

Remark 6.4.2

1. Recall from Remark 6.2.10/2 that Ψ induces an action Ψ^{N} of G on the bundle manifold of the normal bundle $\mathrm{N}O$, turning $\mathrm{N}O$ into a G-vector bundle. We will show below that one may always choose $E = \mathrm{N}O$.
2. Let $m \in O$. Since the orbit mapping $G/G_m \to O$ induced by m is an equivariant diffeomorphism, one might as well assume E to be a G-vector bundle over G/G_m. In this case, the requirement $\chi_{\restriction O} = s_0$ reduces to $\chi(m) = s_0([\mathbb{1}])$.
3. The slices U_m, $m \in M$, of a tubular neighbourhood $\chi : U \to E$ of O have the following properties.
   (a) $G \cdot U_m = U$ and U_m is closed in U.
   (b) If $a \in G$ satisfies $\Psi_a(U_m) \cap U_m \neq \varnothing$, then $a \in G_m$.
   (c) U_m intersects every orbit in U transversally.
   The proof consists in showing that the fibres E_m of the G-vector bundle E have these properties. We leave the details to the reader (Exercise 6.4.1).

Theorem 6.4.3 (Tubular Neighbourhood Theorem) *Every orbit of a proper Lie group action admits a tubular neighbourhood.*

Proof To be definite, we give the proof for a left action. As noted above, we will construct a tubular neighbourhood by means of the normal bundle $\mathrm{N}O$ and the action Ψ^{N}. Choose a point $m \in O$. The plan of the proof can be summarized in the following commutative diagram,

$$\begin{array}{ccccccc}
G \times M & \xleftarrow{\supset} & G \times U_m & \xrightarrow{\mathrm{id}_G \times \varphi} & G \times \mathrm{N}_m O & \xrightarrow{\subset} & G \times \mathrm{N}O \\
\downarrow \Psi & & \downarrow \Psi & & \downarrow \Psi^{\mathrm{N}} & & \downarrow \Psi^{\mathrm{N}} \\
M & \xleftarrow{\supset} & U & \xrightarrow{\chi} & \mathrm{N}O & \xrightarrow{=} & \mathrm{N}O
\end{array} \qquad (6.4.1)$$

whose ingredients will be constructed step by step and which will ultimately define the desired diffeomorphism χ. We start with constructing the mapping $\varphi : U_m \to \mathrm{N}_m O$. According to Remark 6.3.6/1, we may choose a G_m-invariant scalar product on $\mathrm{T}_m M$. Let $(\mathrm{T}_m O)^{\perp}$ denote the corresponding orthogonal complement[18] of $\mathrm{T}_m O$.

[18] Any G_m-invariant complement will do.

Lemma 6.4.4 *There exists a G_m-invariant embedded submanifold U_m of M containing m such that*

$$\mathrm{T}_m U_m = (\mathrm{T}_m O)^\perp$$

and a G_m-equivariant diffeomorphism φ from U_m onto a G_m-invariant open neighbourhood of the origin in $\mathrm{N}_m O$ such that $\varphi(m) = 0$.

Proof of Lemma 6.4.4 Clearly, there exists a diffeomorphism φ_1 from an open neighbourhood W_1 of m in M onto an open neighbourhood of the origin in $\mathrm{T}_m M$ satisfying $\varphi_1(m) = 0$ and

$$(\varphi_1)'_m = \mathrm{id}_{\mathrm{T}_m M} \, . \tag{6.4.2}$$

According to Exercise 6.3.2, W_1 may be chosen to be G_m-invariant. Using the unique invariant volume form da of unit volume on G_m, cf. Corollary 5.5.9, we take the average

$$\varphi_1^{G_m} : W_1 \to \mathrm{T}_m M, \qquad \varphi_1^{G_m}(p) := \int_{G_m} \big((\Psi_a)'_m \circ \varphi_1 \circ \Psi_{a^{-1}}(p)\big) \, \mathrm{d}a.$$

Using (6.4.2), for $X \in \mathrm{T}_m M$, we obtain

$$\big(\varphi_1^{G_m}\big)'_m(X) = \int_{G_m} \big((\Psi_a)'_m \circ (\varphi_1)'_m \circ (\Psi_{a^{-1}})'_m(X)\big) \, \mathrm{d}a = X. \tag{6.4.3}$$

Hence, by the Inverse Mapping Theorem 1.5.7, $\varphi_1^{G_m}$ restricts to a diffeomorphism from an open neighbourhood W_2 of m in M onto an open neighbourhood V_2 of the origin in $\mathrm{N}_m O$. According to Exercise 6.3.2 again, W_2 and hence V_2 can be chosen to be G_m-invariant. We define

$$U_m := \big(\varphi_1^{G_m}\big)^{-1}\big(V_2 \cap (\mathrm{T}_m O)^\perp\big), \qquad \varphi(\tilde{m}) := \big[\varphi_1^{G_m}(\tilde{m})\big], \quad \tilde{m} \in U_m .$$

Obviously, U_m is G_m-invariant and $\varphi(m) = 0$. Moreover, using (6.4.3), we calculate

$$\mathrm{T}_m U_m = \big(\varphi_1^{G_m}\big)'_m(\mathrm{T}_m U_m) = \mathrm{T}_0\big(\varphi_1^{G_m}(U_m)\big) = \mathrm{T}_0(\mathrm{T}_m O)^\perp = (\mathrm{T}_m O)^\perp.$$

Finally, according to Remark 6.3.6/3, the natural projection $\mathrm{T}_m M \to \mathrm{N}_m O$ restricts to a G_m-equivariant vector space isomorphism $(\mathrm{T}_m O)^\perp \to \mathrm{N}_m O$. Hence, φ is a G_m-equivariant diffeomorphism. This proves Lemma 6.4.4. □

Next, by restriction, the action Ψ induces the mapping

$$\psi : G \times U_m \to M, \qquad \psi(a, \tilde{m}) := \Psi_a(\tilde{m}).$$

Lemma 6.4.5 *U_m can be shrunk so that*

1. *ψ is a submersion,*
2. *If $a_1, a_2 \in G$ and $m_1, m_2 \in U_m$ satisfy $\psi(a_1, m_1) = \psi(a_2, m_2)$, then $a_2 a_1^{-1} \in G_m$ and $m_2 = \Psi_{a_2^{-1} a_1}(m_1)$.*

Proof of Lemma 6.4.5 Since $\psi(a,\tilde{m}) = \Psi_a(\psi(\mathbb{1},\tilde{m}))$ for all $a \in G$ and $\tilde{m} \in U_m$, to prove point 1 it suffices to show that ψ is a submersion at $(\mathbb{1},m)$. Since

$$\operatorname{im} \psi'_{(\mathbb{1},m)} = \mathrm{T}_m O + \mathrm{T}_m U_m ,$$

this follows from $\mathrm{T}_m U_m = (\mathrm{T}_m O)^{\perp}$. For point 2, we first observe that $\psi(a_1,m_1) = \psi(a_2,m_2)$ implies $m_2 = \Psi_{a_2^{-1}a_1}(m_1)$. Hence, the point to prove is that the latter implies $a_2^{-1}a_1 \in G_m$ if only m_1 and m_2 are close enough to m. Assume, on the contrary, that this is false. Then, there exists a sequence $\{m_n\}$ in U_m converging to m and a sequence $\{a_n\}$ in $G \setminus G_m$ such that $\Psi_{a_n}(m_n)$ belongs to U_m and converges to m. We will show that this implies $a_n \in G_m$ for large enough n, which is a contradiction.

By Corollary 6.3.3/2, $\{a_n\}$ contains a subsequence converging to some $a \in G$. By passing to this subsequence, we may assume that $a_n \to a$. Then, $\Psi_{a_n}(m_n)$ converges to $\Psi_a(m)$, so that $\Psi_a(m) = m$ and hence $a \in G_m$. Using a vector space complement of $\mathfrak{g}_m$ in $\mathfrak{g}$ and the exponential mapping of G, we construct a submanifold N of G of dimension $\dim G - \dim G_m$ which intersects G_m transversally in $\mathbb{1}$. The Inverse Mapping Theorem implies that the mapping $N \times G_m \to G$ induced by the group multiplication of G restricts to a diffeomorphism between open neighbourhoods of $(\mathbb{1},\mathbb{1})$ in $N \times G_m$ and of $\mathbb{1}$ in G, respectively. Thus, for large enough n, we can decompose $a_n a^{-1} = b_n c_n$ with $b_n \in N$ and $c_n \in G_m$, where both the sequences $\{b_n\}$ and $\{c_n\}$ converge to $\mathbb{1}$. On the other hand, Ψ restricts to a diffeomorphism from an open neighbourhood of $(\mathbb{1},m)$ in $N \times U_m$ onto an open neighbourhood of m in M. Indeed, since U_m is G_m-invariant, we have

$$\Psi'_{(\mathbb{1},m)}(\mathrm{T}_{\mathbb{1}} N \oplus \mathrm{T}_m U_m) = \Psi'_{(\mathbb{1},m)}(\mathrm{T}_{\mathbb{1}} G \oplus \mathrm{T}_m U_m) = \mathrm{T}_m M .$$

Hence, the restriction of $\Psi'_{(\mathbb{1},m)}$ to the subspace $\mathrm{T}_{\mathbb{1}} N \oplus \mathrm{T}_m U_m$ of $\mathrm{T}_{\mathbb{1}} G \oplus \mathrm{T}_m U_m$ is surjective. By counting dimensions one finds that it is also injective. Hence, the assertion follows from the Inverse Mapping Theorem. Now, since the sequences $\{b_n\}$ and $\{c_n\}$ both converge to $\mathbb{1}$, for large enough n, $\Psi_{a_n a^{-1}}(m_n)$ is the image of $(b_n, \Psi_{c_n}(m_n))$ under the diffeomorphism just discussed. Since $\Psi_{a_n a^{-1}}(m_n) \in U_m$, we conclude $b_n = \mathbb{1}$. It follows that $a_n a^{-1} \in G_m$ and hence $a_n \in G_m$, which is the desired contradiction. This proves Lemma 6.4.5. □

Finally, by restriction, the action Ψ^{N} induces the mapping

$$\psi^{\mathrm{N}} : G \times \mathrm{N}_m O \to \mathrm{N}O , \qquad \psi^{\mathrm{N}}\big(a,[X]\big) := \Psi_a^{\mathrm{N}}\big([X]\big). \tag{6.4.4}$$

Lemma 6.4.6

1. *The mapping ψ^{N} is a surjective submersion.*
2. *If $a_1, a_2 \in G$ and $[X_1],[X_2] \in \mathrm{N}_m O$ satisfy the relation $\psi^{\mathrm{N}}(a_1,[X_1]) = \psi^{\mathrm{N}}(a_2,[X_2])$, then $[X_2] = \Psi^{\mathrm{N}}_{a_2^{-1}a_1}([X_1])$ and $a_2^{-1}a_1 \in G_m$.*

Proof of Lemma 6.4.6 1. Surjectivity is due to the fact that Ψ^{N} covers the action of G on O, which is transitive. Since $\psi^{\mathrm{N}}(a,[X]) = \Psi_a^{\mathrm{N}}(\psi^{\mathrm{N}}(\mathbb{1},[X]))$ and since ψ^{N} is linear in the second variable, for proving that ψ^{N} is a submersion it suffices to show that it is a submersion at $(\mathbb{1}, 0)$. The latter follows by observing that with respect to the decomposition

$$\mathrm{T}_{\psi^{\mathrm{N}}(\mathbb{1},0)}\mathrm{N}O = \mathrm{T}_m O \oplus \mathrm{N}_m O,$$

provided by Remark 6.3.6/4, the tangent mapping is given by

$$\big(\psi^{\mathrm{N}}\big)'_{(\mathbb{1},0)}\big(A, [X]\big) = \big((A_*)_m, [X]\big).$$

2. Since the projection $\mathrm{N}O \to O$ is equivariant, $\psi^{\mathrm{N}}(a_1, [X_1]) = \psi^{\mathrm{N}}(a_2, [X_2])$ implies $\Psi_{a_1}(m) = \Psi_{a_2}(m)$ and hence $a_2^{-1}a_1 \in G_m$. The rest is obvious. □

Proof of Theorem 6.4.3 As a consequence of Lemmas 6.4.5/2 and 6.4.6/2 and the G_m-equivariance of φ, there exists a unique injective mapping $\chi : U \to \mathrm{N}O$ such that the central square in the diagram (6.4.1) commutes. Since ψ is a submersion, χ is smooth. Since ψ^{N} is a submersion, χ has open image V and the inverse mapping is smooth, too. Thus, χ is a diffeomorphism onto V. Since ψ intertwines the action of G on $G \times U_m$ by left translation on the first factor with Ψ and since an analogous statement holds for ψ^{N}, χ is equivariant. This completes the proof of the theorem. □

Remark 6.4.7 (General Tubular Neighbourhood Theorem) The notion of tubular neighbourhood extends to arbitrary submanifolds $P \subset M$: a tubular neighbourhood of P consists of an open neighbourhood U of P in M and a diffeomorphism $\chi : U \to \mathrm{N}P$ onto an open subset, mapping P onto the zero section. There holds a general Tubular Neighbourhood Theorem, stating the existence of a tubular neighbourhood for every embedded submanifold. The main difference in the proof is that in the general situation one does not have a group action to transport a diffeomorphism defined in a neighbourhood of some point of P to all of P. Instead, one chooses some Riemannian metric g, identifies $\mathrm{N}P$ with the orthogonal complement of $\mathrm{T}P$ in $\mathrm{T}M_{\restriction P}$ and uses the exponential mapping $\exp_{\mathrm{g}}$ associated with g to construct χ. The exponential mapping $\exp_{\mathrm{g}}$ maps an open neighbourhood of the zero section in $\mathrm{T}M$ to M. It is defined by $\exp_{\mathrm{g}}(X_m) := \gamma(1)$, where γ is the solution of the geodesic equation of g with initial conditions $\gamma(0) = m$ and $\dot{\gamma}(0) = X_m$. For a thorough discussion of the geodesic equation and the corresponding exponential mapping, we refer to [166] or to volume 2 of this book. Let us add that if one has already proved the general Tubular Neighbourhood Theorem, it is more natural to prove Theorem 6.4.3 by the same method, but using a G-invariant Riemannian metric. The latter exists due to Proposition 6.3.7/3.

Exercises

6.4.1 Verify the properties of the slices stated in Remark 6.4.2.

6.5 Free Proper Actions

In the case of a free proper Lie group action, Theorem 6.4.3 allows for constructing an atlas on the orbit space by means of local slices. This yields

Corollary 6.5.1 *The orbit space M/G of a free proper Lie group action (M, G, Ψ) admits a unique smooth structure such that the natural projection $M \to M/G$ is a submersion.*

Accordingly, if Ψ is free and proper, M/G will be referred to as the orbit manifold of Ψ.

Proof Uniqueness follows from Remark 1.5.16/4. To prove existence, let $\pi : M \to M/G$ denote the natural projection to equivalence classes. By Proposition 6.3.4/3, the orbit space M/G is Hausdorff. By Proposition 6.1.5/3, it is second countable. Let $m \in M$ be arbitrary and let O denote the orbit of m. Since the action Ψ is free, Lemma 6.4.6 implies that the mapping $\psi_m : G \times \mathrm{N}_m O \to \mathrm{N}O$, defined by Formula (6.4.4), is a diffeomorphism.[19] Taking a tubular neighbourhood $\chi : U \to \mathrm{N}O$ of O and composing χ with ψ_m^{-1}, we obtain a G-equivariant diffeomorphism $\tilde{\chi}$ from U onto an open G-invariant neighbourhood of $G \times \{0\}$, which by G-invariance must have the form $G \times V_m$ for some open neighbourhood V_m of the origin in $\mathrm{N}_m O$. By embedding V_m as the subset $\{\mathbb{1}\} \times V_m$ into $G \times V_m$, we obtain a bijective continuous mapping

$$V_m \to G \times V_m \overset{\tilde{\chi}^{-1}}{\to} U \overset{\pi}{\to} \hat{U},$$

where $\hat{U} := \pi(U)$ is an open neighbourhood of $\pi(m)$ in M/G. This mapping is open, and hence a homeomorphism, because it maps an open subset W of V_m to the open subset $\pi(\tilde{\chi}^{-1}(G \times W))$ of $\hat{U}$. Thus, the inverse of this mapping defines a local chart $(\hat{U}, \hat{\kappa})$ on M/G at $\pi(m)$. The transition mapping between two such local charts is given by

$$\hat{\kappa}_2 \circ \hat{\kappa}_1^{-1} : \hat{\kappa}_1(\hat{U}_1 \cap \hat{U}_2) \to \hat{\kappa}_2(\hat{U}_1 \cap \hat{U}_2),$$
$$\hat{\kappa}_2 \circ \hat{\kappa}_1^{-1}\big([X]\big) = \mathrm{pr}_2 \circ \tilde{\chi}_2 \circ \tilde{\chi}_1^{-1}\big(\mathbb{1}, [X]\big).$$

Hence, it is smooth. Thus, we have constructed a smooth manifold structure on M/G. In this structure, π is a submersion, because via χ and $\hat{\kappa}$, its restriction to U corresponds to pr_2. □

Remark 6.5.2 Let G be a Lie group and let $H \subset G$ be a closed subgroup. The action of H on G by left or right translation is free and proper, cf. Example 6.3.8/3. Thus, in this case, Corollary 6.5.1 reproduces Theorem 5.7.2.

Corollary 6.5.1 implies (Exercise 6.5.1)

Corollary 6.5.3 *Let (M, G, Ψ) be a free proper Lie group action.*

1. *If $H \subset G$ is a closed normal subgroup, then Ψ descends to a free proper action of G/H on M/H and id_M induces a diffeomorphism between $(M/H)/(G/H)$ and M/G.*

[19] More precisely, ψ_m is an isomorphism of G-vector bundles; in particular, $\mathrm{N}O$ is globally trivial.

2. *If (N, φ) is a G-invariant embedded submanifold, then the induced action of G on N is free and proper and $(N/G, \hat{\varphi})$, with $\hat{\varphi} : N/G \to M/G$ being induced by φ on passing to the quotients, is an embedded submanifold of M/G.*

Example 6.5.4

1. The action of $G = \mathbb{Z}$ on $M = \mathbb{R} \times (-1, 1)$ defined by $\Psi_k(x, t) := (x + k, (-1)^k t)$ is free. It is also proper:[20] let $\{k_n\}$ and $\{(x_n, t_n)\}$ be sequences in $\mathbb{Z}$ and $\mathbb{R} \times (-1, 1)$, respectively, such that $\{\Psi_{k_n}(x_n, t_n)\}$ and $\{(x_n, t_n)\}$ converge. Since
$$\begin{aligned} |k_m - k_n| &\leq \left\| \left(x_n + k_n, (-1)^{k_n} t_n\right) - \left(x_n + k_m, (-1)^{k_m} t_n\right) \right\| \\ &\leq \left\| \Psi_{k_n}(x_n, t_n) - \Psi_{k_m}(x_m, t_m) \right\| + \left\| (x_m, t_m) - (x_n, t_n) \right\|, \end{aligned}$$
there exists a number $n_0 > 0$ such that $k_n = k_{n_0}$ for all $n \geq n_0$. Hence, $\{k_n\}$ converges, and Ψ is proper. The quotient manifold coincides with the Möbius strip constructed in Example 1.1.12: the underlying topological spaces coincide by construction, and the smooth structures coincide, because the natural projection is a submersion with respect to both of them. Let us add that application of Corollary 6.5.3 to the closed normal subgroup $H = 2\mathbb{Z}$ yields that the Möbius strip is diffeomorphic to the quotient manifold of the free proper action of $G/H = \mathbb{Z}_2$ on
$$M/H = \mathbb{R}/2\mathbb{Z} \times (-1, 1) \cong \mathrm{S}^1 \times (-1, 1),$$
where on $\mathrm{S}^1 \times (-1, 1)$, the generator of $\mathbb{Z}_2$ maps (α, t) to $(-\alpha, -t)$.
2. Let $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{H}$ and let $\mathbb{K}_1$ denote the Lie group given by the unit sphere in $\mathbb{K}$, that is, $\mathbb{R}_1 = \mathrm{O}(1)$, $\mathbb{C}_1 = \mathrm{U}(1)$ and $\mathbb{H}_1 = \mathrm{Sp}(1)$. The action of $G = \mathbb{K}_1$ on $M = \mathbb{K}^n$ is proper, because $\mathbb{K}_1$ is compact. Since this action is isometric with respect to the natural scalar product, it restricts to an action of $\mathbb{K}_1$ on the unit sphere S^{dn-1}, where $d = \dim_{\mathbb{R}} \mathbb{K}$. The latter action is in addition free. The quotient manifold $\mathrm{S}^{dn-1}/\mathbb{K}_1$ coincides with the projective space $\mathbb{K}\mathrm{P}^{n-1}$, cf. Example 1.1.15; the argument is the same as under point 1.
3. Let $M = \mathrm{O}(n)$. For $k < n$, application of Corollary 6.5.3 to $G = \mathrm{O}(n-k) \times \mathrm{O}(k)$ and $H = \mathrm{O}(n-k) \times \{\mathbb{1}\}$ yields that $\mathrm{O}(k)$ acts freely and properly on the Stiefel manifold $\mathrm{S}_{\mathbb{R}}(k, n)$, cf. Example 5.7.5, and that the quotient manifold of this action is diffeomorphic to the Graßmann manifold $\mathrm{G}_{\mathbb{R}}(k, n)$, cf. Example 5.7.6. Similar statements hold in the complex and in the quaternionic case. Combining this with the diffeomorphism between $\mathrm{S}_{\mathbb{K}}(1, n)$ and S^{dn-1}, as well as point 2 above, one obtains that $\mathrm{G}_{\mathbb{K}}(1, n)$ is diffeomorphic to $\mathbb{K}\mathrm{P}^{n-1}$. This has already been asserted in Example 5.7.6, but the proof was postponed there.

Corollary 6.5.1 implies the following important bundle structures.

Definition 6.5.5 (Principal bundle) Let (P, G, Ψ) be a free Lie group action, let M be a manifold and let $\pi : P \to M$ be a smooth mapping. The tuple (P, G, M, Ψ, π)

[20] A proper action of a discrete group is called a properly discontinuous action.

is called a principal bundle, if for every $m \in M$ there exists a local trivialization at m, that is, there exist an open neighbourhood W of m and a diffeomorphism $\chi : \pi^{-1}(W) \to W \times G$ such that

1. χ intertwines Ψ with the G-action on $W \times G$ by translations[21] on the factor G,
2. $\mathrm{pr}_W \circ \chi(p) = \pi(p)$ for all $p \in \pi^{-1}(W)$.

Denoting $\kappa := \mathrm{pr}_2 \circ \chi : \pi^{-1}(W) \to G$, for a right action Ψ, property 1 can be rewritten as

$$\kappa\big(\Psi_a(p)\big) = \kappa(p)a, \quad p \in \pi^{-1}(W),\ a \in G.$$

The group G is called the structure group of P. If G is fixed, P is called a principal G-bundle. The existence of local trivializations implies that π is a submersion.[22] Hence the subsets $\pi^{-1}(m)$, $m \in M$, are submanifolds, called the fibres of P.

Now, let (P, G, Ψ) be a free proper Lie group action. Let M be the orbit space, equipped with the smooth structure provided by Corollary 6.5.1, and let $\pi : P \to M$ be the natural projection to orbits. Every tubular neighbourhood of an orbit defines a local trivialization over a neighbourhood of the corresponding point of M. Hence, the Tubular Neighbourhood Theorem 6.4.3 implies that (P, G, M, Ψ, π) is a principal bundle. Conversely, if (P, G, M, Ψ, π) is a principal bundle, (P, G, Ψ) is a free proper Lie group action, M is diffeomorphic to the orbit space P/G and π corresponds, via this diffeomorphism, to the natural projection to orbits (Exercise 6.5.2).

Now, in addition, let (F, Ξ) be a G-manifold. Without loss of generality, assume both Ψ and Ξ to be left actions. The direct product action of G on $P \times F$, given by

$$\big(a, (p, f)\big) \mapsto \big(\Psi_a(p), \Xi_a(f)\big), \tag{6.5.1}$$

is obviously free. Due to Remark 6.3.9/2, it is also proper. Hence, by Corollary 6.5.1, the orbit space

$$P \times_G F := (P \times F)/G$$

inherits a unique smooth structure.

Definition 6.5.6 (Twisted product) The manifold $P \times_G F$ is called the twisted product of (P, Ψ) with (F, Ξ).

There are two bundle structures inherent in the twisted product. First, $P \times_G F$ is the base space of the principal bundle associated with the direct product action (6.5.1) of G on $P \times F$. Second, the natural projection $P \times F \to P$ is equivariant and hence induces a smooth mapping

$$\pi_F : P \times_G F \to P/G = M.$$

[21] Left (right) translations if Ψ is a left (right) action.

[22] In general, the converse does not hold. A local trivialization is more than a local section, because it is global along the fibre.

If $s : U \to P$ is a local section of the natural projection $P \to M$, the mapping

$$U \times F \to \pi_F^{-1}(U), \qquad (m, f) \mapsto \big[\big(s(m), f\big)\big], \tag{6.5.2}$$

is a diffeomorphism projecting to the identical mapping of U (Exercise 6.5.3). Hence, $P \times_G F$ is the total space of a locally trivial fibre bundle[23] over M with typical fibre F.

Definition 6.5.7 (Associated bundle) The locally trivial fibre bundle $(P \times_G F, M, \pi_F)$ is said to be associated with the principal bundle (P, G, M, Ψ, π) and the G-manifold (F, Ξ).

In the special case where F is a finite-dimensional $\mathbb{K}$-vector space and Ξ is a representation of G on F, the associated bundle is a $\mathbb{K}$-vector bundle over M of dimension $\dim F$.

Remark 6.5.8 Let (M, G, Ψ) be a proper Lie group action, without loss of generality assumed to be a left action. Let O be an orbit and let $m \in O$. The action of G_m on G by right translation defines a right principal bundle. After having turned this into a left principal bundle, we can form the associated real vector bundle $G \times_{G_m} \mathrm{N}_m O$, where G_m acts on $\mathrm{N}_m O$ by the slice representation. The action of G on $G \times \mathrm{N}_m O$ by left translation on the first factor induces a natural G-vector bundle structure on $G \times_{G_m} \mathrm{N}_m O$. As a consequence of Lemma 6.4.6, in this structure, $G \times_{G_m} \mathrm{N}_m O$ is isomorphic to $\mathrm{N}O$ (Exercise 6.5.4). Thus, every orbit admits a tubular neighbourhood taking values in $G \times_{G_m} \mathrm{N}_m O$.

More generally, one can prove the following. Let W be a real vector space carrying a representation of G_m and let $\lambda : W \to \mathrm{T}_m M$ be an equivariant injective linear mapping onto a subspace complementary to $\mathrm{T}_m O$. Then, there exists a tubular neighbourhood $\chi : U \to G \times_{G_m} W$ of O such that, under the natural identification

$$\mathrm{T}_{[(\mathbb{1},0)]}(G \times_{G_m} W) = \mathfrak{g}/\mathfrak{g}_m \oplus W,$$

one has

$$\big(\chi^{-1}\big)'_{[(\mathbb{1},0)]}\big([A], w\big) = (A_*)_m + \lambda(w), \tag{6.5.3}$$

see Exercise 6.5.5. This will be used in Sect. 10.4.

Exercises

6.5.1 Prove Corollary 6.5.3.

6.5.2 Let (P, G, M, Ψ, π) be a principal bundle. Prove that (P, G, Ψ) is a free proper Lie group action, M is diffeomorphic to the orbit space P/G and π corresponds, via this diffeomorphism, to the natural projection to orbits.

6.5.3 Show that the mapping (6.5.2) is a diffeomorphism.

[23]The definition of locally trivial fibre bundle is obtained from that of vector bundle by replacing "vector space" by "manifold" and "linear mapping" by "smooth mapping".

6.5.4 Use Lemma 6.4.6 to show that the G-vector bundles $G \times_{G_m} \mathrm{N}_m O$ and $\mathrm{N}O$ are isomorphic.

6.5.5 Prove the existence of a tubular neighbourhood $\chi : U \to G \times_{G_m} W$ satisfying (6.5.3), see Remark 6.5.8.
Hint. Choose G_m-invariant scalar products on $\mathrm{T}_m O$ and $\lambda(W)$ and define the scalar product on $\mathrm{T}_m M$ as their direct sum. Use this scalar product in the proof of Theorem 6.4.3 to obtain a tubular neighbourhood $\chi_0 : U \to \mathrm{N}O$. Next, show that λ extends to a G-vector bundle morphism $G \times_{G_m} W \to (\mathrm{T}M)_{\restriction O}$ and that composition with the natural projection $(\mathrm{T}M)_{\restriction O} \to \mathrm{N}O$ yields a G-vector bundle isomorphism $\chi_1 : G \times_{G_m} W \to \mathrm{N}O$. Then, $\chi := \chi_1^{-1} \circ \chi_0$ has the desired properties.

6.6 The Orbit Space

The Tubular Neighbourhood Theorem 6.4.3 implies that the orbit space admits a disjoint decomposition into manifolds. Let (M, G, Ψ) be a Lie group action and let $\pi : M \to \hat{M} = M/G$ denote the natural projection. Without loss of generality, assume that Ψ is a left action. Recall that M and $\hat{M}$ decompose into orbit type subsets, that is, subsets made up by the orbits of a fixed type. The connected components of the orbit type subsets of $\hat{M}$ will be referred to as the strata of $\hat{M}$. Let $\mathbb{S}$ denote the set of strata. Elements $\sigma \in \mathbb{S}$ will be viewed as pairs consisting of a conjugacy class of subgroups of G and a label for the connected component of the respective orbit type subset. The actual stratum of $\hat{M}$ corresponding to σ will be denoted by $\hat{M}_\sigma$. For $\sigma \in \mathbb{S}$ and for a subgroup $H \subset G$ representing[24] σ, define

$$M_\sigma := \pi^{-1}(\hat{M}_\sigma), \qquad M_{\sigma,H} := M_\sigma \cap M_H .$$

The subsets M_σ and $M_{\sigma,H}$ will be referred to as the orbit type strata of M and the isotropy type strata of M, respectively. Unless G is connected, M_σ or $M_{\sigma,H}$ need not be connected. Obviously,

$$M_\sigma = G \cdot M_{\sigma,H}, \qquad M_\sigma \subset M_{[H]}, \qquad M_{\sigma,H} \subset M_H .$$

By restriction, π induces mappings

$$\pi_\sigma : M_\sigma \to \hat{M}_\sigma, \qquad \pi_{\sigma,H} : M_{\sigma,H} \to \hat{M}_\sigma .$$

Next, recall that the normalizer of a subgroup $H \subset G$ is defined to be

$$\mathrm{N}_G(H) := \{a \in G : aHa^{-1} = H\}. \tag{6.6.1}$$

This is the maximal subgroup of G containing H as a normal subgroup. If H is closed, so is $\mathrm{N}_G(H)$ (Exercise 6.6.1). Theorem 5.6.8 yields that $\mathrm{N}_G(H)$ is an embedded Lie subgroup of G. Hence, it is a Lie group and $H \subset \mathrm{N}_G(H)$ is a closed normal subgroup. Denote the quotient Lie group by

$$\Gamma_H := \mathrm{N}_G(H)/H .$$

[24]That is, representing the conjugacy class of subgroups of G corresponding to σ.

Due to $Ha = aH$ for all $a \in \mathrm{N}_G(H)$, Γ_H acts naturally from the left on G/H:

$$\Gamma_H \times G/H \to G/H, \qquad (aH, bH) \mapsto ba^{-1}H.$$

Since the natural projection $\mathrm{N}_G(H) \times G \to \Gamma_H \times G/H$ is a direct product of submersions, this action is smooth.

Proposition 6.6.1 (Structure of strata) *Let (M, G, Ψ) be a proper left Lie group action. Let $\sigma \in \mathbb{S}$ and let H be a subgroup of G representing σ.*

1. *M_σ and $M_{\sigma,H}$ are embedded submanifolds of M. For $m \in M_{\sigma,H}$, one has*
$$\mathrm{T}_m M_{\sigma,H} = (\mathrm{T}_m M)^H, \qquad \mathrm{T}_m M_\sigma = (\mathrm{T}_m M)^H + \mathrm{T}_m(G \cdot m).$$
2. *Ψ induces a free and proper left action of Γ_H on $M_{\sigma,H}$.*
3. *The twisted product $M_{\sigma,H} \times_{\Gamma_H} G/H$ carries a left G-action, given by*
$$\big(a, \big[(m, bH)\big]\big) \mapsto \big[(m, abH)\big],$$
and Ψ induces a G-equivariant diffeomorphism
$$\psi : M_{\sigma,H} \times_{\Gamma_H} G/H \to M_\sigma, \qquad \psi\big(\big[(m, aH)\big]\big) := \Psi_a(m).$$
4. *There exists a unique smooth structure on $\hat{M}_\sigma$ such that the natural projection π_σ is a submersion. With respect to this structure, the natural projection $\pi_{\sigma,H}$ is a submersion, too, and the natural inclusion mapping $M_{\sigma,H} \to M_\sigma$ descends to a diffeomorphism from $M_{\sigma,H}/\Gamma_H$ onto $\hat{M}_\sigma$.*

Proof 1. According to Remark 1.6.13/3, it suffices to prove the assertion for the intersections of the subsets $M_{\sigma,H}$ and M_σ with a tubular neighbourhood of the orbit of any of their points. Thus, according to the Tubular Neighbourhood Theorem 6.4.3 and Remark 6.5.8, it is enough to consider the case $M = G \times_H V$, where V is a finite-dimensional real vector space carrying a representation ϱ of H. To determine the stabilizers, let $a \in G$ and $v \in V$. One has $\Psi_b([(a, v)]) \equiv [(ba, v)] = [(a, v)]$ iff $ba = ac^{-1}$ and $v = \varrho(c)v$ for some $c \in H$. Hence, $G_{[(a,v)]} = aH_v a^{-1}$, where H_v denotes the stabilizer of v under ϱ. This implies

$$M_\sigma = G \times_H V^H, \qquad M_{\sigma,H} = \mathrm{N}_G(H) \times_H V^H,$$

where V^H denotes the subspace of fixed points under ϱ. That is, M_σ is a vertical vector subbundle of the vector bundle $M = G \times_H V$ and $M_{\sigma,H}$ is the restriction of this vertical subbundle to the subset $\Gamma_H \subset G/H$, which is easily seen to be an embedded submanifold. Hence, the assertion follows from Examples 2.7.2 and 2.7.3. This argument yields in addition that the mapping $G \times M_{\sigma,H} \to M_\sigma$ defined by Ψ is a submersion. This will be needed in the proof of assertion 3. To prove the formulae for the tangent spaces, we write tangent vectors of $G \times_H V$ at $[(a, v)] \in \mathrm{N}_G(H) \times_H V^H$ in the form $\mathrm{pr}'(\mathrm{L}'_a A, u)$ with $A \in \mathfrak{g}$, $u \in V$ and with $\mathrm{pr} : G \times V \to G \times_H V$ denoting the natural projection. One can show that such a tangent vector is invariant under the isotropy representation of $G_{[(a,v)]} = H$ iff

$u \in V^H$ and A belongs to the Lie algebra of $\mathrm{N}_G(H)$ (Exercise 6.6.2). This yields $\mathrm{T}_m M_{\sigma,H} = (\mathrm{T}_m M)^H$. The formula for $\mathrm{T}_m M_\sigma$ then follows.

2. Since $M_{\sigma,H}$ is invariant under $\mathrm{N}_G(H)$, Ψ induces an action of $\mathrm{N}_G(H)$ on $M_{\sigma,H}$. By Proposition 6.3.4/1, the latter is proper. Since it has kernel H, by Proposition 6.1.5/5, it induces an action of Γ_H on $M_{\sigma,H}$. Using Corollary 6.3.3/2, it is easy to see that the latter remains proper. By construction, it is also free.

3. According to point 2, we can build the twisted product $M_{\sigma,H} \times_{\Gamma_H} G/H$. Since the action of G on $M_{\sigma,H} \times G/H$ by left translations on the second factor commutes with the action of Γ_H, it descends to an action of G on $M_{\sigma,H} \times_{\Gamma_H} G/H$, given by the asserted formula. Now, consider the mapping ψ. It is well-defined: if $[(m_1, a_1 H)] = [(m_2, a_2 H)]$, then $m_2 = \Psi_b(m_1)$ and $a_2 = a_1 b^{-1}$ for some $b \in \mathrm{N}_G(H)$, and hence $\Psi_{a_2}(m_2) = \Psi_{a_1}(m_1)$. Moreover, ψ is G-equivariant. By construction, we have the commutative diagram

$$\begin{array}{ccc} M_{\sigma,H} \times G & \longrightarrow & M_\sigma \\ \downarrow & \nearrow_{\psi} & \\ M_{\sigma,H} \times_{\Gamma_H} G/H & & \end{array}$$

where the horizontal arrow is given by the action Ψ. Since the vertical arrow is a submersion, ψ is smooth. Since, as noticed under point 1, the horizontal arrow is a submersion, ψ is a submersion, too. For dimensional reasons, it is also an immersion then. Hence, it remains to show that ψ is bijective. Surjectivity is obvious. To prove injectivity, let $m_1, m_2 \in M_{\sigma,H}$ and $a_1, a_2 \in G$ be such that $\Psi_{a_1}(m_1) = \Psi_{a_2}(m_2)$. Then, $b := a_2^{-1} a_1 \in \mathrm{N}_G(H)$, because both m_1 and $\Psi_b(m_1) = m_2$ have stabilizer H. Thus, $a_2 = a_1 b^{-1}$ and hence $[(m_1, a_1 H)] = [(m_2, a_2 H)]$.

4. The natural inclusion mapping $M_{\sigma,H} \to M_\sigma$ descends to a mapping

$$\hat{\psi} : M_{\sigma,H}/\Gamma_H \to \hat{M}_\sigma .$$

Since two points of $M_{\sigma,H}$ are conjugate under G iff they are conjugate under $\mathrm{N}_G(H)$, $\hat{\psi}$ is a bijection. According to assertion 2 and Corollary 6.5.1, $M_{\sigma,H}/\Gamma_H$ carries a unique smooth structure such that the natural projection $M_{\sigma,H} \to M_{\sigma,H}/\Gamma_H$ is a submersion. We have the commutative diagram

$$\begin{array}{ccc} M_{\sigma,H} \times_{\Gamma_H} G/H & \xrightarrow{\ \psi\ } & M_\sigma \\ {\scriptstyle \pi_{G/H}}\downarrow & & \downarrow{\scriptstyle \pi_\sigma} \\ M_{\sigma,H}/\Gamma_H & \xrightarrow{\ \hat{\psi}\ } & \hat{M}_\sigma \end{array}$$

where $\pi_{G/H}$ is the natural projection in the corresponding associated fibre bundle. Thus, it is a submersion, and hence it is open. Since M_σ is an embedded submanifold, π_σ is open, too. Thus, since ψ is a homeomorphism, so is $\hat{\psi}$. We define a smooth structure on $\hat{M}_\sigma$ by requiring $\hat{\psi}$ to be a diffeomorphism. In view of the above diagram, with respect to this structure, π_σ is a smooth submersion, and so is $\pi_{\sigma,H}$. Uniqueness is then obvious. □

Remark 6.6.2

1. The finer decomposition of the orbit type subsets of $\hat{M}$ into connected components is necessary because the latter may have different dimensions, see [275] for an example.
2. Since the smooth structure of $\hat{M}_\sigma$ is unique, assertion 4 of Proposition 6.6.1 implies that different choices of the subgroup $H \subset G$ representing σ lead to diffeomorphic orbit manifolds $M_{\sigma,H}/\Gamma_H$. This assertion also generalizes Corollary 6.5.1 to the case where the action is not free, but has a single orbit type.
3. The diffeomorphism $\hat{\psi}$ turns $M_{\sigma,H}$ into a principal bundle over $\hat{M}_\sigma$ with structure group Γ_H and projection $\pi_{\sigma,H}$. Similarly, the diffeomorphisms ψ and $\hat{\psi}$ turn M_σ into a locally trivial fibre bundle over $\hat{M}_\sigma$ with typical fibre G/H and projection π_σ, associated with that principal bundle.
4. Since $\pi_\sigma : M_\sigma \to \hat{M}_\sigma$ is a submersion, it admits local sections. Tubular neighbourhoods provide a distinguished class of local sections in the following way. Let $m \in M_\sigma$, let $\chi : U \to E$ be a tubular neighbourhood of the orbit $G \cdot m$ and let U_m be the corresponding slice at m. We have $M_\sigma \cap U_m = M_{\sigma,H} \cap U_m$, where $H = G_m$. Since this subset is mapped under χ onto $V_m^H \equiv E_m^H \cap V_m$, it is an embedded submanifold of M.[25] The standard properties of slices stated in Remark 6.4.2/3 imply that any two distinct points in $M_\sigma \cap U_m$ belong to different orbits. It follows that the restriction of the surjective submersion π_σ to $M_\sigma \cap U_m$ is injective and hence a diffeomorphism onto an open neighbourhood of $\pi(m)$ in $\hat{M}_\sigma$. Then, the inverse of this mapping yields a local section of π_σ at $\pi_\sigma(m)$, taking values in U_m, and composition of this section with χ yields a local chart on $\hat{M}_\sigma$ at $\pi_\sigma(m)$, taking values in the open subset V_m^H of the vector space E_m^H.
5. The family $\{\hat{M}_\sigma : \sigma \in \mathbb{S}\}$ establishes a disjoint decomposition of $\hat{M}$ into manifolds. This decomposition has several additional properties reflecting how these manifolds fit together. In fact, it is a so-called Whitney stratification, see, for example, [238]. A similar statement is true for M and the family $\{M_\sigma : \sigma \in \mathbb{S}\}$. If the orbit space M/G is connected, there exists an orbit type which is minimal in the sense that one (and hence any) of its representatives is conjugate to a subgroup of the stabilizer of an arbitrary point. The corresponding orbit type subset of M/G is open and dense in M/G and connected [238, Thm. 4.3.2], [150, Thm. 2.1], see also [54, Thm. IV.3.1]. It is therefore referred to as the principal stratum of M/G. Correspondingly, all the other strata are referred to as the secondary strata.
6. It is sometimes[26] more convenient to work with the connected components of the submanifolds M_σ and $M_{\sigma,H}$. For a chosen point $m_0 \in M_{\sigma,H}$, let $M_\sigma^{m_0}$ and $M_{\sigma,H}^{m_0}$ denote the connected component of M_σ or $M_{\sigma,H}$, respectively, containing m_0. One can show that

$$G^{m_0} := \left\{a \in G : \Psi_a(m_0) \in M_\sigma^{m_0}\right\} \tag{6.6.2}$$

[25] This follows also from the Transversal Mapping Theorem 1.8.2.

[26] E.g. in the theory of singular symplectic reduction.

is a closed subgroup of G (Exercise 6.6.3). Since it contains H, we can form the quotient Lie group

$$\Gamma_H^{m_0} := \mathrm{N}_{G^{m_0}}(H)/H.$$

Now, points 1–3 of Proposition 6.6.1 remain true if M_σ, $M_{\sigma,H}$, G and Γ_H are replaced by, respectively, $M_\sigma^{m_0}$, $M_{\sigma,H}^{m_0}$, G^{m_0} and $\Gamma_H^{m_0}$. Point 4 can be replaced by the statement that the natural inclusion mappings $M_\sigma^{m_0} \to M_\sigma$ and $M_{\sigma,H}^{m_0} \to M_\sigma$ induce diffeomorphisms $M_\sigma^{m_0}/G^{m_0} \to \hat{M}_\sigma$ and $M_{\sigma,H}^{m_0}/\Gamma_H^{m_0} \to \hat{M}_\sigma$, respectively, where the smooth structure of $\hat{M}_\sigma$ is the one inherited from M_σ according to point 4 of the original proposition.

Next, we discuss the concept of a smooth function on the orbit space. Let (M, G, Ψ) be a free proper Lie group action. Recall that, in this case, $\hat{M}$ is a smooth manifold and $\pi : M \to \hat{M}$ is a smooth mapping. Since π is surjective, the pull-back $\pi^* : C^\infty(\hat{M}) \to C^\infty(M)$ is injective. Its image is contained in the subalgebra $C^\infty(M)^G$ of $C^\infty(M)$ of G-invariant functions. Conversely, every $f \in C^\infty(M)^G$ defines a function $\hat{f}$ on $\hat{M}$ by $\pi^*\hat{f} = f$. Since π is a submersion, $\hat{f}$ is smooth. This shows that π^* induces an algebra isomorphism from $C^\infty(\hat{M})$ onto $C^\infty(M)^G$. If (M, G, Ψ) is proper but not free, $\hat{M}$ does not inherit a smooth structure from M, hence a priori we do not have the concept of a smooth function on $\hat{M}$. However, the above observation and the fact that, due to Proposition 6.3.7/1, $C^\infty(M)^G$ separates the points of $\hat{M}$, motivate the following generalization.

Definition 6.6.3 A continuous function f on $\hat{M}$ is said to be smooth if the function $f \circ \pi$ on M is smooth. The set of smooth functions on $\hat{M}$ is denoted by $C^\infty(\hat{M})$.

Obviously, $C^\infty(\hat{M})$ is an associative algebra, with operations being defined pointwise. By construction, the pull-back π^* defines an algebra isomorphism from $C^\infty(\hat{M})$ onto $C^\infty(M)^G$.

Proposition 6.6.4 *For every $\hat{f} \in C^\infty(\hat{M})$ and every $\sigma \in \mathbb{S}$, $\hat{f}_{\restriction \hat{M}_\sigma} \in C^\infty(\hat{M}_\sigma)$.*

Proof Let $H \subset G$ be a subgroup representing σ. By assumption, $(\pi^*\hat{f})_{\restriction M_{\sigma,H}} = \pi_{\sigma,H}^*(\hat{f}_{\restriction \hat{M}_\sigma})$ is a smooth function on $M_{\sigma,H}$. Hence, the above argument for proper free actions yields that $\hat{f}_{\restriction \hat{M}_\sigma}$ is a smooth function on $\hat{M}_\sigma$. □

Example 6.6.5 (Adjoint representations of $\mathrm{U}(n)$ and $\mathrm{SU}(n)$) Let G be a compact Lie group, let $M = \mathfrak{g}$ and let Ψ be given by the adjoint representation of G. For subsets $S \subset G$ and $\mathfrak{s} \subset \mathfrak{g}$, define the centralizers

$$\mathrm{C}_\mathfrak{g}(S) := \{A \in \mathfrak{g} : \mathrm{Ad}(a)A = A \text{ for all } a \in S\},$$
$$\mathrm{C}_G(\mathfrak{s}) := \{a \in G : \mathrm{Ad}(a)_{\restriction \mathfrak{s}} = \mathrm{id}_\mathfrak{s}\}.$$

We will restrict attention to the cases $G = \mathrm{U}(n)$ and $G = \mathrm{SU}(n)$, starting with $\mathrm{U}(n)$. For a positive integer r, let S_r denote the group of permutations of r elements.

Let $\mathfrak{t}$ denote the subspace of diagonal matrices in $\mathfrak{u}(n)$. Being skew-Hermitian, the elements of $\mathfrak{t}$ have purely imaginary entries $\mathrm{i}x_1, \ldots, \mathrm{i}x_n$. Let $\mathfrak{c} \subset \mathfrak{t}$ denote the cone defined by $x_1 \leq \cdots \leq x_n$. By elementary linear algebra, every element of $\mathfrak{g} = \mathfrak{u}(n)$ is conjugate under Ψ to some $A \in \mathfrak{c}$. The stabilizer of A is given by $\mathrm{C}_{\mathrm{U}(n)}(A)$, that is, by the subgroup

$$H_{\mathbf{k}} := \left\{\mathrm{diag}(a_1, \ldots, a_r) : a_i \in \mathrm{U}(k_i)\right\}$$

of $\mathrm{U}(n)$, labelled by the sequence $\mathbf{k} = (k_1, \ldots, k_r)$ of the multiplicities of the eigenvalues of A. Let $\mathbf{K}$ denote the set of all sequences arising this way, that is, sequences $\mathbf{k} = (k_1, \ldots, k_r)$ of length $r = 1, 2, \ldots$ of positive integers satisfying $k_1 + \cdots + k_r = n$. Subgroups $H_{\mathbf{k}}$ and $H_{\mathbf{l}}$ are conjugate in $\mathrm{U}(n)$ iff the sequences $\mathbf{k}$ and $\mathbf{l}$ differ by a permutation. Hence, the orbit types of Ψ correspond bijectively to partitions $n = k_1 + \cdots + k_r$. Now, let $\mathbf{k} \in \mathbf{K}$. Obviously, $M_{H_{\mathbf{k}}}$ is given by the subset

$$\mathfrak{t}_{\mathbf{k}} := \left\{\mathrm{diag}(\mathrm{i}y_1 \mathbb{1}_{k_1}, \ldots, \mathrm{i}y_r \mathbb{1}_{k_r}) : y_i \in \mathbb{R},\ \text{pairwise distinct}\right\} \subset \mathfrak{t}.$$

It follows that $M_{[H_{\mathbf{k}}]}$ consists of the elements of $\mathfrak{u}(n)$ having r distinct eigenvalues with multiplicities $k_1, \ldots, k_r$. Next, we determine $\mathrm{N}_{\mathrm{U}(n)}(H_{\mathbf{k}})$. For $\varrho \in \mathrm{S}_r$, define $\varrho_{\mathbf{k}} \in \mathrm{S}_n$ to be obtained by dividing $(1, \ldots, n)$ into the r subsequences $(1, \ldots, k_1)$, $(k_1 + 1, \ldots, k_1 + k_2)$, $\ldots$, $(k_1 + \cdots + k_{r-1} + 1, \ldots, n)$ and permuting these subsequences according to ϱ. Since $a \in \mathrm{N}_{\mathrm{U}(n)}(H_{\mathbf{k}})$ iff $\mathrm{Ad}(a) M_{H_{\mathbf{k}}} \subset M_{H_{\mathbf{k}}}$, the normalizer $\mathrm{N}_{\mathrm{U}(n)}(H_{\mathbf{k}})$ is generated by $H_{\mathbf{k}}$ and the permutation matrices of $\varrho_{\mathbf{k}}$ for all $\varrho \in \mathrm{S}_r$ satisfying $\varrho(\mathbf{k}) = \mathbf{k}$. It follows that $\Gamma_{H_{\mathbf{k}}}$ coincides with the stabilizer $(\mathrm{S}_r)_{\mathbf{k}}$ of $\mathbf{k}$ under the action of S_r and that it acts on $M_{H_{\mathbf{k}}}$ via the corresponding permutations $\varrho_{\mathbf{k}}$ of the entries. To determine the strata, for $\varrho \in \mathrm{S}_r$, denote

$$\mathfrak{t}_{\mathbf{k}}^{\varrho} := \left\{\mathrm{diag}(\mathrm{i}y_1 \mathbb{1}_{k_1}, \ldots, \mathrm{i}y_r \mathbb{1}_{k_r}) \in \mathfrak{t}_{\mathbf{k}} : y_{\varrho(1)} < \cdots < y_{\varrho(r)}\right\}.$$

These subsets are the connected components of $\mathfrak{t}_{\mathbf{k}}$ and hence of $M_{H_{\mathbf{k}}}$. Two connected components $\mathfrak{t}_{\mathbf{k}}^{\varrho_1}$ and $\mathfrak{t}_{\mathbf{k}}^{\varrho_2}$ get identified under the action of $\Gamma_{H_{\mathbf{k}}}$ iff $\varrho_1 \varrho_2^{-1} \in (\mathrm{S}_r)_{\mathbf{k}}$. It follows that the strata correspond bijectively to the elements of $\mathbf{K}$ and that the stratum $\hat{M}_{\mathbf{k}}$, $\mathbf{k} \in \mathbf{K}$, is given by the image of $\mathfrak{t}_{\mathbf{k}}^{\mathrm{id}}$ under the natural projection π to the orbit space. Since π is obviously injective on $\mathfrak{t}_{\mathbf{k}}^{\mathrm{id}}$, it induces a diffeomorphism

$$\mathfrak{t}_{\mathbf{k}}^{\mathrm{id}} \cong \hat{M}_{\mathbf{k}}.$$

Moreover, $M_{\mathbf{k}, H_{\mathbf{k}}}$ consists of the connected components $\mathfrak{t}_{\mathbf{k}}^{\varrho}$ with $\varrho \in (\mathrm{S}_r)_{\mathbf{k}}$ and $M_{\mathbf{k}}$ consists of the elements of $\mathfrak{u}(n)$ whose eigenvalues, in ascending order, have the multiplicities $k_1, \ldots, k_r$. Finally, the submanifolds $\mathfrak{t}_{\mathbf{k}}^{\mathrm{id}}$, $\mathbf{k} \in \mathbf{K}$, of $\mathfrak{t}$ fit together to form the cone $\mathfrak{c}$. By restriction, π yields a bijection from this cone onto $\hat{M}$. It is not hard to see that this bijection is in fact a homeomorphism. Thus, via this homeomorphism, the stratum $\hat{M}_{(1,\ldots,1)}$ corresponds to the interior of $\mathfrak{c}$, given by $x_1 < \cdots < x_n$, and the other strata form the boundary of $\mathfrak{c}$, consisting of lower-dimensional cones.

For the case $G = \mathrm{SU}(n)$, we observe that restriction of the above action of $\mathrm{U}(n)$ on $\mathfrak{u}(n)$ to the subgroup $\mathrm{SU}(n)$ yields the same orbits. Since the subspace

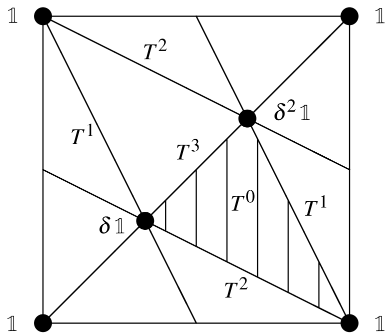


Fig. 6.1 The curves T^1, T^2, T^3 and the central elements $\mathbb{1}$, $\delta\mathbb{1}$ and $\delta^2\mathbb{1}$, where $\delta = e^{i\frac{2}{3}\pi}$, in the subgroup $T \cong U(1) \times U(1)$ of diagonal matrices in SU(3), see Example 6.6.6

$\mathfrak{su}(n) \subset \mathfrak{u}(n)$ is invariant under this action, the results for $U(n)$ carry over to $SU(n)$ as follows. The orbit types are the same. They are represented by the stabilizers

$$SH_{\mathbf{k}} := SU(n) \cap H_{\mathbf{k}}, \quad \mathbf{k} \in \mathbf{K}.$$

Since up to a phase factor, every permutation matrix may be chosen from $SU(n)$, the normalizers $N_{SU(n)}(SH_{\mathbf{k}})$ and the quotient groups $\Gamma_{SH_{\mathbf{k}}}$ can be characterized in the same way as for $G = U(n)$. We have $M_{SH_{\mathbf{k}}} = \mathfrak{t}_{\mathbf{k}} \cap \mathfrak{su}(n)$, with connected components $\mathfrak{t}_{\mathbf{k}}^{\varrho} \cap \mathfrak{su}(n)$, $\varrho \in S_r$. The strata correspond bijectively to the elements of $\mathbf{K}$ and for every $\mathbf{k} \in \mathbf{K}$, π restricts to a diffeomorphism from $\mathfrak{t}_{\mathbf{k}}^{\mathrm{id}} \cap \mathfrak{su}(n)$ onto the stratum $\hat{M}_{\mathbf{k}}$. The strata fit together to form the cone $\mathfrak{c} \cap \mathfrak{su}(n)$ which is an $(n-1)$-dimensional simplex with one face moved to infinity.[27] Note that intersection with $\mathfrak{su}(n)$ is achieved by imposing the additional condition $x_1 + \cdots + x_n = 0$, which for $\mathfrak{t}_{\mathbf{k}} \cap \mathfrak{su}(n)$ reads $k_1 y_1 + \cdots + k_r y_r = 0$. For a description of the orbit space, including the strata, in terms of discriminants we refer to [143].

Example 6.6.6 (Inner automorphisms of SU(3)) Consider the action of $G = SU(3)$ on the group manifold $M = SU(3)$ by inner automorphisms. Let T denote the subgroup of diagonal matrices. We have $T \cong U(1) \times U(1)$. Since every element of SU(3) can be diagonalized and since two elements of SU(3) are diagonal iff they have the same eigenvalues, the orbit space $\hat{M}$ of this action may be identified with the orbit space of the action of S_3 on T by permuting the matrix entries. The latter can be described as follows, see Fig. 6.1. Each of the three equations $a_{22} = a_{33}$, $a_{33} = a_{11}$ and $a_{11} = a_{22}$ defines a closed curve in T, respectively denoted by T^1, T^2 and T^3. These curves intersect in the centre Z of SU(3) and separate six open subsets in T. The closures of these subsets form 2-simplices whose edges are given by one piece of each T^i and whose vertices correspond to the elements of Z. Let T^0 be one of these open subsets. It is easy to see that the natural projection $M \to \hat{M}$ restricts to a bijection, and hence a homeomorphism, from $\overline{T^0}$ onto $\hat{M}$. Thus, we can read off the stratification of $\hat{M}$ from $\overline{T^0}$. The interior T^0 forms a single two-dimensional stratum. It consists of the points whose stabilizer under the action of

[27] In the theory of semisimple Lie algebras, this cone is called a closed Weyl chamber.

SU(3) is T or, equivalently, whose stabilizer under the action of S_3 is trivial. There are three one-dimensional strata given by $(T^i \cap \overline{T^0}) \setminus Z$, $i = 1, 2, 3$. The stabilizers under the action of SU(3) are isomorphic to U(2) and consist, respectively, of the matrices

$$\begin{bmatrix} \overline{\det a} & 0 & 0 \\ 0 & a_{11} & a_{12} \\ 0 & a_{21} & a_{22} \end{bmatrix}, \quad \begin{bmatrix} a_{11} & 0 & a_{12} \\ 0 & \overline{\det a} & 0 \\ a_{21} & 0 & a_{22} \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & \overline{\det a} \end{bmatrix}, \tag{6.6.3}$$

where $a \in \mathrm{U}(2)$. The corresponding stabilizers under the action of S_3 are generated, respectively, by the transpositions (23), (31) and (12). Since these subgroups are conjugate in SU(3) and S_3, respectively, all one-dimensional strata belong to the same orbit type. Finally, there are three point strata formed by the elements of Z. Their stabilizers are SU(3) and S_3, respectively. As a result, the stratification of $\hat{M}$ coincides with the natural stratification of the 2-simplex by its open cells.

Remark 6.6.7 The result of Example 6.6.6 generalizes to SU(n) with arbitrary n (Exercise 6.6.5). In lattice gauge theory, the orbit space of the diagonal action of G on $G \times \cdots \times G$ by inner automorphisms is relevant. We refer to [61, 62] for a detailed analysis including a description of the stratification in terms of invariants for a specific example. Correspondingly, the unreduced phase space of such models is given by $\mathrm{T}^*(G \times \cdots \times G)$ with the induced diagonal action of G. Under the identification of T^*G with $G \times \mathfrak{g}$, G acts diagonally by inner automorphisms on the factors G and by the adjoint representation on the factors $\mathfrak{g}$. It is possible to work out the reduction with respect to this action explicitly. In lattice gauge theory, however, the Gauß law yields a reduction to a level set of the momentum mapping, see Sect. 10.7.

Exercises

6.6.1 Show that the normalizer of a closed subgroup of a Lie group is closed.

6.6.2 Complete the proof of Proposition 6.6.1/1 by showing that, in the notation used there, a tangent vector $\pi'(\mathrm{L}'_a A, u)$ of $G \times_H V$ at the point $[(a, v)] \in \mathrm{N}_G(H) \times_H V^H$ is invariant under the isotropy representation of H iff $u \in V^H$ and A belongs to the Lie algebra of $\mathrm{N}_G(H)$.

6.6.3 Show that the subset G^{m_0} defined by (6.6.2) is a closed subgroup of G.

6.6.4 Determine the structure of the orbit space and its strata for the action of SO(n) on $\mathbb{R}^n$.

6.6.5 Work out Example 6.6.6 for arbitrary SU(n).

6.7 Invariant Vector Fields

Throughout this section, let (M, G, Ψ) be a proper left Lie group action and let $\pi : M \to \hat{M} = M/G$ denote the natural projection. We will discuss the basic properties of the following type of vector fields.

Definition 6.7.1 A vector field X on M is called invariant if $\Psi_{a*}X = X$ for all $a \in G$.

This is equivalent to the requirement that X be equivariant as a mapping $M \to \mathrm{T}M$. Therefore, it is also common to refer to invariant vector fields as equivariant vector fields. The invariant vector fields form a Lie subalgebra of $\mathfrak{X}(M)$, denoted by $\mathfrak{X}(M)^G$.

Now, let $X \in \mathfrak{X}(M)^G$ and let $\Phi : \mathscr{D} \to M$ be the flow of X. By Proposition 3.2.13/2, for every $a \in G$, we have $\Psi_a(\mathscr{D}_t) = \mathscr{D}_t$ for all $t \in \mathbb{R}$ and

$$\Phi_t \circ \Psi_a(m) = \Psi_a \circ \Phi_t(m) \tag{6.7.1}$$

for all $(t, m) \in \mathscr{D}$. That is, Ψ restricts to an action of G on $\mathscr{D}_t$ for all t and the diffeomorphism $\Phi_t : \mathscr{D}_t \to \mathscr{D}_{-t}$ is equivariant. This implies that Φ projects to a continuous mapping $\hat{\Phi} : \hat{\mathscr{D}} \to \hat{M}$, where $\hat{\mathscr{D}} := (\mathrm{id}_{\mathbb{R}} \times \pi)(\mathscr{D})$, defined by

$$\hat{\Phi}\big(t, \pi(m)\big) := \pi \circ \Phi(t, m). \tag{6.7.2}$$

It will be shown below that $\hat{\Phi}$ is a flow[28] on the topological space $\hat{M}$. It is, therefore, referred to as the projected flow, or the projection of Φ. Next, by equivariance,

$$G_{\Phi_t(m)} = G_m \tag{6.7.3}$$

for all $(t, m) \in \mathscr{D}$, that is, the flow leaves invariant the stabilizers and hence the orbit types. Thus, for every $\sigma \in \mathbb{S}$ and every subgroup $H \subset G$ representing σ, the submanifolds M_σ and $M_{\sigma,H}$ of M are invariant under Φ. It follows that X is tangent to M_σ, so that it restricts to a vector field X^σ on M_σ whose flow $\Phi^\sigma : \mathscr{D}^\sigma \to M$ is given by the restriction of Φ,

$$\mathscr{D}^\sigma = \mathscr{D} \cap (\mathbb{R} \times M_\sigma), \qquad \Phi^\sigma = \Phi_{\restriction \mathscr{D}^\sigma}, \tag{6.7.4}$$

cf. Proposition 2.7.16 and Remark 3.2.9/2. A similar statement holds for $M_{\sigma,H}$. Since

$$\pi'_\sigma\big(X^\sigma_{\Psi_a(m)}\big) = \pi'_\sigma\big(X^\sigma_m\big)$$

for all $a \in G$, the vector field X^σ on M_σ induces a mapping $\hat{X}^\sigma : \hat{M}_\sigma \to \mathrm{T}\hat{M}_\sigma$ by

$$\hat{X}^\sigma_{\pi_\sigma(m)} = \pi'_\sigma X^\sigma_m. \tag{6.7.5}$$

Since π_σ is a submersion, this mapping is smooth and hence a vector field on $\hat{M}_\sigma$. Denote the corresponding flow by $\hat{\Phi}^\sigma : \hat{\mathscr{D}}^\sigma \to \hat{M}_\sigma$. By construction, X^σ and $\hat{X}^\sigma$ are π_σ-related.

Proposition 6.7.2 *Let X be an invariant vector field on M. If $\gamma : (a, b) \to M_\sigma$ is an integral curve of X^σ, then $\pi_\sigma \circ \gamma : (a, b) \to \hat{M}_\sigma$ is an integral curve of $\hat{X}^\sigma$. Every integral curve of $\hat{X}^\sigma$ is of this form.*

---

[28] By an obvious modification of Definition 3.2.5, the notion of flow extends to the category of topological spaces and continuous mappings.

Proof That $\pi_\sigma \circ \gamma$ is an integral curve of $\hat{X}^\sigma$ is obvious. Conversely, let $\hat{\gamma} : (a,b) \to \hat{M}_\sigma$ be an integral curve of $\hat{X}^\sigma$. Choose a point $m \in M_\sigma$ fulfilling $\pi_\sigma(m) = \hat{\gamma}(0)$ and denote $H = G_m$. According to Remark 6.6.2/3, $M_{\sigma,H}$ is a principal bundle over $\hat{M}_\sigma$ with structure group $\Gamma_H = \mathrm{N}_G(H)/H$. By the help of local trivializations of this bundle we construct a curve $\tilde{\gamma} : (a,b) \to M_{\sigma,H}$ through m fulfilling $\pi \circ \tilde{\gamma} = \hat{\gamma}$. If Γ_H is discrete, $\tilde{\gamma}$ is uniquely determined by this condition. It is an integral curve of X, because π is a local diffeomorphism in this case.

Now, consider the case where Γ_H is not discrete. Here, it is enough to show that there exists a curve $\alpha : (a,b) \to \Gamma_H$ such that

$$\gamma : (a,b) \to M_{\sigma,H}, \qquad \gamma(t) = \Psi_{\alpha(t)}\big(\tilde{\gamma}(t)\big),$$

is an integral curve of X^σ. Here, by an abuse of notation, we have written Ψ for the action of Γ_H on $M_{\sigma,H}$. Applying $\Psi'_{\alpha(t)^{-1}}$ to the equation

$$X^\sigma_{\gamma(t)} = \dot{\gamma}(t) = \Psi'_{\tilde{\gamma}(t)}\dot{\alpha}(t) + \Psi'_{\alpha(t)}\dot{\tilde{\gamma}}(t)$$

and using the invariance of X^σ, together with $\Psi_{\alpha(t)^{-1}} \circ \Psi_{\tilde{\gamma}(t)} = \Psi_{\tilde{\gamma}(t)} \circ \mathrm{L}_{\alpha(t)^{-1}}$, we obtain

$$(\Psi_{\tilde{\gamma}(t)})'_{\mathbb{1}} \circ \mathrm{L}'_{\alpha(t)^{-1}}\big(\dot{\alpha}(t)\big) = X^\sigma_{\tilde{\gamma}(t)} - \dot{\tilde{\gamma}}(t). \tag{6.7.6}$$

For every t, the right hand side is a tangent vector at $\tilde{\gamma}(t)$ which is annihilated by $\pi'_{\sigma,H}$ and hence contained in $\mathrm{T}_{\tilde{\gamma}(t)}(\Gamma_H \cdot \tilde{\gamma}(t))$. Since the action of Γ_H on $M_{\sigma,H}$ is free, $(\Psi_{\tilde{\gamma}(t)})'_{\mathbb{1}}$ is a bijection from $\mathrm{T}_{\mathbb{1}}\Gamma_H$ onto $\mathrm{T}_{\tilde{\gamma}(t)}(\Gamma_H \cdot \tilde{\gamma}(t))$. Thus, there exists a unique curve $A_{\tilde{\gamma}} : (a,b) \to \mathrm{T}_{\mathbb{1}}\Gamma_H$ such that

$$(\Psi_{\tilde{\gamma}(t)})'_{\mathbb{1}} A_{\tilde{\gamma}}(t) = X^\sigma_{\tilde{\gamma}(t)} - \dot{\tilde{\gamma}}(t), \quad t \in (a,b),$$

and (6.7.6) implies that α must satisfy the ordinary first-order differential equation

$$\dot{\alpha}(t) = \mathrm{L}'_{\alpha(t)} A_{\tilde{\gamma}}(t), \quad t \in (a,b).$$

For the initial condition $\alpha(0) = \mathbb{1}$, the solution is given by the path-ordered exponential mapping, $\alpha(t) = \mathrm{T}\exp \int_0^t A_{\tilde{\gamma}}(s)\,\mathrm{d}s$ for $t > 0$ and $\alpha(t) = \mathrm{T}\exp(-\int_t^0 A_{\tilde{\gamma}}(s)\,\mathrm{d}s)$ for $t < 0$. □

Proposition 6.7.2 is the main ingredient in proving (Exercise 6.7.1)

Corollary 6.7.3

1. *The mapping* $\hat{\Phi} : \hat{\mathscr{D}} \to \hat{M}$ *is a flow on the topological space* $\hat{M}$.
2. *The flow* $\hat{\Phi}^\sigma$ *of* $\hat{X}^\sigma$ *is given by the restriction of* $\hat{\Phi}$,

$$\hat{\mathscr{D}}^\sigma = \hat{\mathscr{D}} \cap (\mathbb{R} \times \hat{M}_\sigma), \qquad \hat{\Phi}^\sigma = \hat{\Phi}_{\restriction \hat{\mathscr{D}}^\sigma}. \tag{6.7.7}$$

As a further consequence of Proposition 6.7.2, we note

Corollary 6.7.4 *Let* X *be an invariant vector field on* M. *For every integral curve* γ *of* X, *the subset* $G \cdot \gamma$ *is a submanifold of* M. *It is embedded iff* $\hat{\gamma} = \pi(\gamma)$ *is an embedded submanifold of the corresponding stratum of* $\hat{M}$.

Proof Let $\sigma \in \mathbb{S}$ be such that $\gamma \subset M_\sigma$. Since M_σ is an embedded submanifold of M, it is enough to prove the assertion for M_σ. By Proposition 6.7.2, the projected integral curve $\pi_\sigma \circ \gamma$ is the integral curve of a vector field on $\hat{M}_\sigma$ and thus its image $\hat{\gamma}$ is a submanifold of $\hat{M}_\sigma$. By Remark 6.6.2/3, we can cover $\hat{\gamma}$ by a system of local trivializations $\{(U_i, \chi_i)\}$ of the fibre bundle $\pi_\sigma : M_\sigma \to \hat{M}_\sigma$, mapping $\pi_\sigma^{-1}(U_i)$ to $U_i \times G/H$, where $H \subset G$ is a subgroup representing σ. Then, $\chi_i((G \cdot \gamma) \cap \pi_\sigma^{-1}(U_i)) = (\hat{\gamma} \cap U_i) \times G/H$ is a submanifold of $U_i \times G/H$. Hence, $(G \cdot \gamma) \cap \pi_\sigma^{-1}(U_i)$ is a submanifold of the open subset $\pi_\sigma^{-1}(U_i)$ of M_σ. Since $G \cdot \gamma$ is covered by the subsets $\pi_\sigma^{-1}(U_i)$, this shows that it is a submanifold of M_σ. By Remark 1.6.13/3, $\hat{\gamma}$ is embedded iff so is $G \cdot \gamma$. □

Remark 6.7.5 Let us analyse how the projected flow $\hat{\Phi}$ of an invariant vector field X can be characterized in terms of the algebra $C^\infty(\hat{M})$ of smooth functions on $\hat{M}$. Since, for every $f \in C^\infty(M)^G$,

$$\Psi_a^*\big(X(f)\big) = \big(\big(\Psi_a^{-1}\big)_* X\big)\big(\Psi_a^* f\big) = X(f),$$

X induces a derivation of the subalgebra $C^\infty(M)^G$ of $C^\infty(M)$. By the algebra isomorphism $\pi^* : C^\infty(\hat{M}) \to C^\infty(M)^G$, this derivation is mapped to a derivation $\hat{X}$ of $C^\infty(\hat{M})$. By definition,

$$\pi^*\big(\hat{X}(\hat{f})\big) = X\big(\pi^* \hat{f}\big) \tag{6.7.8}$$

for all $\hat{f} \in C^\infty(\hat{M})$. Now, the projected flow satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t} \hat{f} \circ \hat{\Phi}_p(t) = \big(\hat{X}(\hat{f}) \circ \hat{\Phi}_p\big)(t) \tag{6.7.9}$$

for all $(t, p) \in \hat{\mathscr{D}}$ and $\hat{f} \in C^\infty(\hat{M})$. Since according to Proposition 6.3.7/1, $C^\infty(\hat{M})$ separates the points of $\hat{M}$, the projected flow is uniquely determined by this equation and thus by $\hat{X}$.

In the remainder of this section, we show that every tubular neighbourhood of an orbit allows for a decomposition of invariant vector fields adapted to the bundle structure given by the slices. Let O be an orbit of Ψ, let $\chi : U \to E$ be a tubular neighbourhood of O and let S denote the regular distribution on U spanned by the tangent spaces of the slices U_m, $m \in O$. Note that S is invariant, that is, $\Psi_a' S_m = S_{\Psi_a(m)}$ for all $a \in G$ and $m \in U$. We are going to construct a complementary invariant distribution on U. Recall from Remark 2.2.10 that the tangent mappings $\Psi_m' : \mathrm{T}G \to \mathrm{T}M$ combine to a smooth mapping $\mathrm{T}G \times M \to \mathrm{T}M$. Restriction of the latter to $\mathfrak{g} \times U$ yields a smooth mapping

$$\lambda : U \times \mathfrak{g} \to \mathrm{T}U, \qquad \lambda(m, A) = \Psi_m'(A).$$

Equip $U \times \mathfrak{g}$ with the action of G on $U \times \mathfrak{g}$ by $(a, (m, A)) \mapsto (\Psi_a(m), \mathrm{Ad}(a)A)$ (direct product of G-manifolds). It is easy to see that λ is equivariant with respect to this action. Choose $m_0 \in O$. Since G_{m_0} is compact, according to Proposition 5.5.6,

$\mathfrak{g}$ admits an $\mathrm{Ad}(G_{m_0})$-invariant scalar product. Let $\mathfrak{m}$ be the corresponding orthogonal complement of $\mathfrak{g}_m$ in $\mathfrak{g}$. By construction, $\mathfrak{m}$ is $\mathrm{Ad}(G_m)$-invariant. View $U \times \mathfrak{g}$ as a trivial vector bundle and consider the invariant subset

$$\mathfrak{M} := \big\{\big(\Psi_a(m), \mathrm{Ad}(a)A\big) \in U \times \mathfrak{g} : m \in U_{m_0},\ a \in G,\ A \in \mathfrak{m}\big\}.$$

Lemma 6.7.6 $(\mathfrak{M}, \lambda_{\restriction \mathfrak{M}})$ *is an invariant vertical subbundle of* $\mathrm{T}U$ *complementary to* S *in* $\mathrm{T}U$.

Proof We start with proving that $\mathfrak{M}$ is a vertical subbundle of $U \times \mathfrak{g}$. For $m \in U$, the subset $\mathfrak{M}_m = \mathfrak{M} \cap (\{m\} \times \mathfrak{g})$ is given by the union of all subspaces $\mathrm{Ad}(a)\mathfrak{m}$ with $a \in G$ such that $m \in \Psi_a(U_{m_0})$. By Remark 6.4.2/3, any two such group elements differ by right translation by an element of G_{m_0}. Since $\mathfrak{m}$ is $\mathrm{Ad}(G_{m_0})$-invariant, all these subspaces coincide. Hence, $\mathfrak{M}_m$ is a linear subspace of $\{m\} \times \mathfrak{g}$, given by $\mathfrak{M}_m = \mathrm{Ad}(a)\mathfrak{m}$ whenever $m \in \Psi_a(U_{m_0})$. By Proposition 2.7.5, it suffices to find local r-frames, where $r = \dim \mathfrak{m}$, in $U \times \mathfrak{g}$ spanning $\mathfrak{M}$. By invariance, it suffices to find such a local r-frame in a neighbourhood of the slice U_{m_0}. Consider the mapping

$$\mathfrak{m} \times U_{m_0} \to U, \qquad (A, m) \mapsto \Psi_{\exp A}(m).$$

Since its tangent mapping at $(0, m_0)$ is bijective, it restricts to a diffeomorphism from an open neighbourhood of $(0, m_0)$ in $\mathfrak{m} \times U_{m_0}$ onto an open neighbourhood $\tilde{U}$ of U_{m_0} in U. The inverse of this diffeomorphism induces smooth mappings $p_1 : \tilde{U} \to \mathfrak{m}$ and $p_2 : \tilde{U} \to U_{m_0}$ such that

$$\Psi_{\exp \circ p_1(m)}\big(p_2(m)\big) = m, \quad m \in \tilde{U}.$$

Now, every basis $\{e_1, \dots, e_r\}$ in $\mathfrak{m}$ defines a local r-frame $\{s_1, \dots, s_r\}$ in $U \times \mathfrak{g}$ over $\tilde{U}$ by

$$s_i(m) := \Psi'_{\exp \circ p_1(m)}(e_i)$$

and this r-frame spans $\mathfrak{M}$ over $\tilde{U}$. Thus, $\mathfrak{M}$ is a vertical subbundle of $U \times \mathfrak{g}$.

Now, consider the mapping $\lambda_{\restriction \mathfrak{M}} : \mathfrak{M} \to \mathrm{T}U$. For every $a \in G$, $\mathrm{Ad}(a)\mathfrak{m}$ is a vector space complement of $\mathrm{Ad}(a)\mathfrak{g}_{m_0}$ and the latter is the Lie algebra of the subgroup $G_{\Psi_a(m_0)}$, which is the invariance group of the slice $U_{\Psi_a(m_0)}$. First, this implies that $\lambda_{\restriction \mathfrak{M}}$ is fibrewise injective and hence, by Proposition 2.7.4, $(\mathfrak{M}, \lambda_{\restriction \mathfrak{M}})$ is a vertical subbundle. Secondly, since slices are transversal to orbits, this implies

$$\lambda(\mathfrak{M}_m) + S_m = \mathrm{T}_m M$$

for all $m \in U$. Finally, since $\lambda(\mathfrak{M}_{m_0}) \cap S_{m_0} = 0$, by counting dimensions we obtain $\lambda(\mathfrak{M}_m) \cap S_m = 0$ for all $m \in U$. □

Lemma 6.7.6 yields that the image $\lambda(\mathfrak{M})$ is a regular invariant distribution on U, contained in $D^{\mathfrak{g}}$ and satisfying

$$\lambda(\mathfrak{M}) \oplus S = \mathrm{T}U. \tag{6.7.10}$$

This implies the following natural decomposition of invariant vector fields.

Proposition 6.7.7 *For every invariant vector field X on M there exists an invariant vector field $Z \in \Gamma(S)$ and a smooth equivariant mapping $A : U \to \mathfrak{g}$ such that, for all $m \in U$,*

$$X_m = \Psi'_m\big(A(m)\big) + Z_m .$$

Proof According to (6.7.10), X decomposes as $X = Y + Z$ with uniquely defined $Y \in \Gamma(\lambda(\mathfrak{M}))$ and $Z \in \Gamma(S)$. Since $\lambda(\mathfrak{M})$ and S are invariant, Y and Z are equivariant. Since vertical subbbundles are embedded, Y induces a section $\tilde{A}$ of $\mathfrak{M}$ by $Y = \lambda \circ \tilde{A}$. Since λ is equivariant, $\tilde{A}$ is invariant. Writing $\tilde{A}(m) = (m, A(m))$ with a smooth equivariant mapping $A : U \to \mathfrak{g}$ we obtain the assertion. □

The pair (A, Z) is called a slice decomposition of X on U and Z is called the slice component of X. Let Φ^Z denote the flow of Z. We show that Φ^Z projects to the restriction of $\hat{\Phi}$ to $\pi(U)$. Let $m \in U$. We seek a curve $t \mapsto \alpha(t, m)$ such that

$$\gamma(t) = \Psi_{\alpha(t,m)} \circ \Phi^Z_t(m)$$

is an integral curve of X through m. As in the proof of Proposition 6.7.2, one can show that this curve must satisfy

$$\dot{\alpha}(t, m) = \mathrm{L}'_{\alpha(t,m)} A\big(\Phi^Z_t(m)\big) \tag{6.7.11}$$

(Exercise 6.7.2). Hence, it is given by the path-ordered exponential

$$\alpha(t, m) = \mathrm{T}\exp\left\{\int_0^t A\big(\Phi^Z_s(m)\big)\mathrm{d}s\right\}$$

and we have

$$\Phi_t(m) = \Psi_{\alpha(t,m)} \circ \Phi^Z_t(m). \tag{6.7.12}$$

This shows that Φ^Z projects to $\hat{\Phi}$, indeed. The mapping $(t, m) \mapsto \alpha(t, m)$ is called the phase mapping of the slice decomposition (A, Z).

Exercises

6.7.1 Prove Corollary 6.7.3. In particular, show that the mapping $\hat{\Phi} : \hat{\mathscr{D}} \to \hat{M}$ defined by (6.7.2) fulfils the conditions of Definition 3.2.5, with C^∞ replaced by continuous.

6.7.2 Prove Formula (6.7.11).

6.8 On Relatively Critical Integral Curves

In the last section of this chapter we give a brief introduction to relative equilibria and relatively periodic integral curves, as well as the corresponding stability concepts. In parts, we will stay informal. Standard references for this material are [65, 81, 90] and, more historically, [174].

Let (M, G, Ψ) be a proper Lie group action, let X be an invariant vector field on M and let $\Phi : \mathscr{D} \to M$ be the flow of X. For an integral curve γ of X, let G_γ denote the stabilizer of the points of γ, let $\mathfrak{g}_\gamma$ be its Lie algebra and let $\hat{\gamma} = \pi \circ \gamma$ denote the projection to $\hat{M}$. Extending the terminology of Chap. 3, we say that an integral curve of $\hat{\Phi}$ is critical if it is an equilibrium, that is, it consists of a point, or if it is periodic.

Definition 6.8.1 An integral curve γ of X is called a relative equilibrium if its projection $\hat{\gamma}$ to $\hat{M}$ is an equilibrium. It is called relatively periodic if $\hat{\gamma}$ is periodic.

If γ is relatively periodic, the period of $\hat{\gamma}$ is called the relative period of γ. Relative equilibria and relative periodic integral curves are subsumed under the notion of relatively critical integral curve. According to Corollary 6.7.4, if γ is relatively critical, then $G \cdot \gamma$ is an embedded submanifold of M. Moreover, it is clear that relatively critical integral curves are defined for all times. We start with characterizing relatively critical integral curves in terms of the group action.

Proposition 6.8.2 *Let X be an invariant vector field on the G-manifold (M, Ψ).*

1. *An integral curve γ of X is a relative equilibrium iff there exist $m \in \gamma$ and $A \in \mathfrak{g}$ such that*
$$(A_*)_m = X_m. \tag{6.8.1}$$
In this case, $\Phi_t(m) = \Psi_{\exp(tA)}(m)$. The set of solutions A of (6.8.1) forms a coset with respect to $\mathfrak{g}_\gamma$ in the normalizer $\mathrm{N}_\mathfrak{g}(\mathfrak{g}_\gamma)$. It does not depend on the choice of $m \in \gamma$.
2. *An integral curve γ of X is relatively periodic iff there exist $T > 0$, $m \in \gamma$ and $a \in G$ such that $\Phi_t(m) \notin G \cdot m$ for all $0 < t < T$ and*
$$\Phi_T(m) = \Psi_a(m). \tag{6.8.2}$$
In this case, T is the relative period of γ. The set of solutions a of (6.8.2) forms a G_γ-coset in the normalizer $\mathrm{N}_G(G_\gamma)$. It does not depend on the choice of $m \in \gamma$.

The solutions of (6.8.1) and (6.8.2) are called drift velocities and relative phases, respectively.

Proof 1. First, let γ be an arbitrary integral curve of X. If γ is a relative equilibrium, then $\hat{\gamma} = \pi(m)$ and hence $\gamma \subset G \cdot m$ for all $m \in \gamma$. It follows that $X_m \in \mathrm{T}_m(G \cdot m)$, so that $X_m = (A_*)_m$ for some $A \in \mathfrak{g}$. If, conversely, $X_m = (A_*)_m$ with $A \in \mathfrak{g}$, invariance implies that X is tangent to the orbit $G \cdot m$ and hence induces a vector field on $G \cdot m$ whose integral curves coincide with those of X. We conclude that $\gamma \subset G \cdot m$ and thus $\hat{\gamma} = \pi(m)$, that is, γ is a relative equilibrium.

Now, let γ be a relative equilibrium. If A is a solution of (6.8.1), then for all $m \in \gamma$ and $t \in \mathbb{R}$ we have
$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\Psi_{\exp(tA)}(m)\big) = \Psi'_{\exp(tA)}(A_*)_m = \Psi'_{\exp(tA)} X_m = X_{\Psi_{\exp(tA)}(m)}$$

and hence $\Psi_{\exp(tA)}(m) = \Phi_t(m)$. Moreover, (6.7.1) yields

$$(A_*)_{\Phi_t(m)} = (\Phi_t)'_m (A_*)_m = (\Phi_t)'_m X_m = X_{\phi_t(m)},$$

that is, A is also a solution for $\Phi_t(m)$. We conclude that the set of solutions is independent of the choice of $m \in \gamma$. Since the left hand side of (6.8.1) is linear in the variable A, the solutions form a coset in $\mathfrak{g}$ with respect to the linear subspace of solutions of the homogeneous equation $(B_*)_m = 0$. By Proposition 6.2.2/3, this subspace is given by $\mathfrak{g}_\gamma$. It remains to show that the solutions A of (6.8.1) are contained in $\mathrm{N}_{\mathfrak{g}}(\mathfrak{g}_\gamma)$, that is, that they satisfy $[A, B] \in \mathfrak{g}_\gamma$ for all $B \in \mathfrak{g}_\gamma$. Using Proposition 6.2.2/2, for $f \in C^\infty(M)$ we find

$$\begin{aligned}\big([A, B]_*\big)_m(f) &= [B_*, A_*]_m(f) = -X_m\big(B_*(f)\big)\\ &= -\frac{\mathrm{d}}{\mathrm{d}t}\Big|_0 \frac{\mathrm{d}}{\mathrm{d}s}\Big|_0 f\big(\Psi_{\exp(sB)} \circ \Phi_t(m)\big) = 0,\end{aligned}$$

because $\Psi_{\exp(sB)} \circ \Phi_t(m) = \Phi_t(m)$. Thus, $([A, B]_*)_m = 0$ and hence $[A, B] \in \mathfrak{g}_\gamma$.

2. The first assertion is obvious. Thus, let γ be relatively periodic and let $m \in \gamma$. If a is a solution of (6.8.2) for m, it is also a solution for $\Phi_t(m)$, because

$$\Phi_T\big(\Phi_t(m)\big) = \Phi_t \circ \Phi_T(m) = \Phi_t \circ \Psi_a(m) = \Psi_a\big(\Phi_t(m)\big).$$

Hence, the set of solutions of (6.8.2) does not depend on $m \in \gamma$. It is contained in $\mathrm{N}_G(G_\gamma)$, because

$$G_\gamma = G_{\Phi_T(m)} = G_{\Psi_a(m)} = aG_m a^{-1} = aG_\gamma a^{-1}.$$

Finally, the set of solutions is obviously a (left and right) coset with respect to G_γ. □

For an analysis of the global properties of relatively critical integral curves like periodicity, quasi-periodicity or the escaping behaviour, the reader may consult Sections 7 and 8 in [81].

The concept of linearization extends in the following way from critical to relatively critical integral curves. Let γ be relatively critical and let $m \in \gamma$.

(a) If γ is a relative equilibrium and A is a drift velocity, then m is an equilibrium of the (not necessarily invariant) vector field $X - A_*$ and we can form the Hessian endomorphism

$$\mathrm{Hess}_m(X - A_*).$$

It will be referred to as the Hessian endomorphism of X at m associated with A.

(b) If γ is relatively periodic with relative period T and if a is a relative phase, then

$$\Psi_{a^{-1}} \circ \Phi_T(m) = m, \qquad (\Psi_{a^{-1}} \circ \Phi_T)'_m X_m = X_m \tag{6.8.3}$$

(Exercise 6.8.1). Hence, $(\Psi_{a^{-1}} \circ \Phi_T)'_m$ descends to an automorphism P^a_m of $\mathrm{N}_m\gamma$, called the period automorphism associated with a.

Thus, for relatively critical integral curves, instead of a single characteristic linear mapping, one has a whole family of characteristic linear mappings, labelled by the drift velocities or the relative phases. The algebraic structure of the latter, described in Proposition 6.8.2, carries over to these families.

Corollary 6.8.3 *Let γ be a relatively critical integral curve of X and let $m \in \gamma$.*

1. *If γ is a relative equilibrium, the Hessian endomorphisms* $\mathrm{Hess}_m(X - A_*)$ *form a coset in* $\mathrm{End}(\mathrm{T}_m M)$ *with respect to the image of the isotropy representation.*
2. *If γ is relatively periodic, the period automorphisms P_m^a form a coset*[29] *in* $\mathrm{Aut}(\mathrm{N}_m\gamma)$ *with respect to the image of the isotropy representation induced on* $\mathrm{N}_m\gamma$.[30]

Proof 1. If A_1, A_2 are drift velocities, then $A_2 - A_1 \in \mathfrak{g}_\gamma$ and

$$\mathrm{Hess}_m(X - A_{1*}) - \mathrm{Hess}_m(X - A_{2*}) = \mathrm{Hess}_m\big((A_2 - A_1)_*\big) \qquad (6.8.4)$$

(Exercise 6.8.2). Hence, the Hessian endomorphisms form a coset with respect to the image of $\mathfrak{g}_\gamma$ under the mapping $B \mapsto \mathrm{Hess}_m(B_*)$. By Remark 6.2.3, this is the isotropy representation.

2. If a_1, a_2 are relative phases, then $a_1^{-1}a_2 \in G_\gamma$ and the automorphism $P_m^{a_1} \circ (P_m^{a_2})^{-1}$ of $\mathrm{N}_m\gamma$ is induced by the automorphism $(\Psi_{a_1^{-1}a_2})'_m$ of $\mathrm{T}_m M$ on passing to $\mathrm{N}_m\gamma$. □

It remains to extract those properties of the spectra of the characteristic linear mappings which do not depend on the particular choice of the latter. For point 2 of the following proposition, we have to assume that G is a subgroup of $\mathrm{GL}(n, \mathbb{R})$ or $\mathrm{GL}(n, \mathbb{C})$ defined by algebraic relations.

Proposition 6.8.4 *Let X be an invariant vector field on the G-manifold (M, Ψ).*

1. *For a relative equilibrium γ, the real parts and the multiplicities of the elements of* $\mathrm{spec}(\mathrm{Hess}_m(X - A_*))$ *are independent of the point $m \in \gamma$ and of the drift velocity A.*
2. *For a relatively periodic integral curve γ of X, the absolute values and the multiplicities of the elements of* $\mathrm{spec}(P_m^a)$ *are independent of the point $m \in \gamma$ and the relative phase a.*

Proof 1. Since the set of drift velocities is independent of $m \in \gamma$ and since

$$\mathrm{Hess}_{\Phi_t(m)}(X - A_*) = (\Phi_t)'_m \circ \mathrm{Hess}_m(X - A_*) \circ \big((\Phi_t)'_m\big)^{-1}$$

for all $t \in \mathbb{R}$, the spectrum of $\mathrm{Hess}_m(X - A_*)$ is independent of m. To prove the assertion we need the following standard relation, valid for compact subgroups (Exercise 6.8.3)

$$\mathrm{N}_\mathfrak{g}(\mathfrak{g}_\gamma) = \mathfrak{g}_\gamma + \mathrm{C}_\mathfrak{g}(\mathfrak{g}_\gamma). \qquad (6.8.5)$$

Since, by Proposition 6.8.2/1, the drift velocities form an affine subspace of $\mathrm{N}_\mathfrak{g}(\mathfrak{g}_\gamma)$ with underlying vector subspace $\mathfrak{g}_\gamma$, (6.8.5) implies that there exists a drift velocity

[29] Since the relative phases lie in $\mathrm{N}_G(G_\gamma)$, this is both a left and a right coset.

[30] Which exists due to $(\Psi_a)'_m X_m = X_{\Psi_a(m)} = X_m$ for all $a \in G_\gamma$.

B commuting with all elements of $\mathfrak{g}_\gamma$. For given drift velocity A, denote $C := B - A$ and write $H_A \equiv \mathrm{Hess}_m(X - A_*)$ and $H_B \equiv \mathrm{Hess}_m(X - B_*)$. By (6.8.4), we have

$$H_A = H_B + \mathrm{Hess}_m(C_*). \tag{6.8.6}$$

Since $C \in \mathfrak{g}_\gamma$ and hence $[B, C] = 0$, Proposition 6.2.2/2 implies $[C_*, B_*] = [B, C]_* = 0$. Moreover, by the invariance of X and by Proposition 3.2.15, $[C_*, X] = 0$. Hence, this proposition implies that the flows of C_* and $X - B_*$ commute. Then, $\mathrm{Hess}_m(C_*)$ commutes with H_B and hence, by (6.8.6), it also commutes with H_A. For the argument to follow, view H_A, H_B and $\mathrm{Hess}_m(C_*)$ as endomorphisms of the complexification of $\mathrm{T}_m M$. Then, since G_γ is compact, Proposition 5.5.7 yields that $\mathrm{Hess}_m(C_*)$ is diagonalizable and possesses purely imaginary eigenvalues. Since H_B and H_A commute with $\mathrm{Hess}_m(C_*)$, they map every eigenspace of $\mathrm{Hess}_m(C_*)$ to itself. Thus, according to (6.8.6), on a given eigenspace, H_B and H_A differ by the identical mapping, multiplied by the corresponding eigenvalue of $\mathrm{Hess}_m(C_*)$. This yields the assertion.

2. It suffices to prove the assertion for the automorphism $(\Psi_{a^{-1}} \circ \Phi_T)'$ of $\mathrm{T}_m M$ which induces P_m^a on passing to $\mathrm{N}_m\gamma$. By an abuse of notation, we denote this automorphism by the same symbol, P_m^a. Since the set of relative phases is independent of $m \in \gamma$ and since

$$P^a_{\Phi_t(m)} = (\Phi_t)'_m \circ P^a_m \circ \big((\Phi_t)'_m\big)^{-1},$$

the spectrum of P_m^a does not depend on m. To prove the assertion, we need that there exists an integer $n > 0$ fulfilling $a^n = c\,h$ with $c \in \mathrm{C}_G(G_\gamma)$ and $h \in G_\gamma$. Under the assumption made on G, this follows from Proposition 1.2 in [316]. Then, for every relative phase b, we have $b = ak_0$ for some $k_0 \in G_\gamma$ and hence

$$b^n = (ak_0)^n = ch\big(a^{-(n-1)}k_0 a^{n-1}\big)\cdots\big(a^{-1}k_0 a\big)k_0 \equiv ck$$

with $k \in G_\gamma$. Consider the automorphism $\tau := (\Psi_{c^{-1}} \circ \Phi_{nT})'_m$ of $\mathrm{T}_m M$. Since it commutes with $(\Psi_{h^{-1}})'_m$ and $(\Psi_{k^{-1}})'_m$, it maps the eigenspaces of these automorphisms to themselves. According to Proposition 5.5.7, $(\Psi_{h^{-1}})'_m$ and $(\Psi_{k^{-1}})'_m$ are diagonalisable and their eigenvalues lie on the unit circle. Since

$$\big(P_m^a\big)^n = (\Psi_{h^{-1}})'_m \circ \tau, \qquad \big(P_m^b\big)^n = (\Psi_{k^{-1}})'_m \circ \tau,$$

the eigenvalues of $(P_m^a)^n$ and of $(P_m^b)^n$ differ from those of τ, and hence from one another, by phase factors only. Since this carries over to the n-th roots, the assertion follows. □

As a result of Proposition 6.8.4, the notions of characteristic exponent and characteristic multiplier extend in the following way to relatively critical integral curves.

(a) Let γ be a relative equilibrium of X. A real number λ is called a reduced characteristic exponent of γ if there exists a point $m \in \gamma$, a drift velocity A and an element $\tilde\lambda$ of $\mathrm{spec}(\mathrm{Hess}_m(X - A_*))$ such that $\lambda = \mathrm{Re}(\tilde\lambda)$. In this case, the multiplicity of λ is defined as the sum of the multiplicities of all $\tilde\lambda \in \mathrm{spec}(\mathrm{Hess}_m(X - A_*))$ with $\lambda = \mathrm{Re}(\tilde\lambda)$.

(b) Let γ be a relatively periodic integral curve of X. A positive number λ is called a reduced characteristic multiplier of γ if there exists a point $m \in \gamma$, a relative phase a and an element $\tilde{\lambda}$ of $\operatorname{spec}(P_m^a)$ such that $\lambda = |\tilde{\lambda}|$. In this case, the multiplicity of λ is defined to be the sum of the multiplicities of all $\tilde{\lambda} \in \operatorname{spec}(P_m^a)$ with $\lambda = |\tilde{\lambda}|$.

Remark 6.8.5

1. The existence and uniqueness results for Poincaré mappings along periodic integral curves have an analogue for relatively periodic integral curves, see e.g. [65], providing equivariant Poincaré mappings. Using a tubular neighbourhood, the corresponding slice decomposition of X and a relative phase, from an equivariant Poincaré mapping one can construct a relative Poincaré mapping. The latter can be interpreted as a Poincaré mapping along the periodic integral curve $\hat{\gamma}$ with respect to the projected flow $\hat{\Phi}$. It can be used to analyse the flow $\hat{\Phi}$ near $\hat{\gamma}$.
2. The reduced characteristic exponents and the reduced characteristic multipliers may be defined as the elements of the so-called reduced spectrum, see [90]. The reduced spectrum of $\operatorname{Hess}_m(X - A_*)$ is defined as the quotient of $\operatorname{spec}(\operatorname{Hess}_m(X - A_*))$ with respect to the action of the additive group $\mathrm{i}\mathbb{R}$. The reduced spectrum of P_m^a is defined as the quotient of $\operatorname{spec}(P_m^a)$ with respect to the action of $\mathrm{U}(1)$. In both cases, the multiplicity of an equivalence class is the sum of the multiplicities of its elements.
3. The analysis of the characteristic linear mappings of a relatively critical integral curve γ can be refined as follows. Choose $m \in \gamma$. Using a tubular neighbourhood $\chi : U \to E$ of the orbit $G \cdot m$ and a G_γ-invariant scalar product on $\mathrm{T}_m M$, one can decompose
$$\mathrm{T}_m M = \mathrm{T}_m(G \cdot m) \oplus \mathrm{T}_m U_m^{G_\gamma} \oplus \left(\mathrm{T}_m U_m^{G_\gamma}\right)^\perp. \tag{6.8.7}$$
Here, U_m denotes the slice of the tubular neighbourhood χ through m, $U_m^{G_\gamma}$ denotes the submanifold of G_γ-invariant points of U_m and $(\mathrm{T}_m U_m^{G_\gamma})^\perp$ denotes the orthogonal complement of $\mathrm{T}_m U_m^{G_\gamma}$ in $\mathrm{T}_m U_m$. Under π, $U_m^{G_\gamma}$ projects to an open neighbourhood of $\pi(m)$ in its stratum. One can prove that, with respect to the decomposition (6.8.7), the characteristic linear mappings are given by upper triangular block matrices. Accordingly, their spectra decompose into a symmetry part (corresponding to the subspace $\mathrm{T}_m(G \cdot m)$), an isotypic part (corresponding to the subspace $\mathrm{T}_m U_m^{G_\gamma}$) and an aliotypic (i.e. type-transversal) part (corresponding to the subspace $(\mathrm{T}_m U_m^{G_\gamma})^\perp$). This partition does not depend on the choice of E and χ nor on the scalar product on $\mathrm{T}_m M$. It turns out that the isotypic part coincides with the spectrum of the ordinary characteristic linear mapping associated with the ordinary critical integral curve $\hat{\gamma}$ inside the corresponding stratum. In particular, this part of the spectrum does not depend on the particular characteristic linear mapping chosen. The partition so constructed allows for an analysis of how the flow near γ behaves in orbit direction (symmetry part), along the stratum (isotypic part) and transversal to the stratum (aliotypic part). On the other

hand, the latter two parts characterize the projected flow $\hat{\Phi}$ near $\hat{\gamma}$. Therefore, the corresponding characteristic exponents or multipliers may be interpreted as the characteristic exponents or multipliers of $\hat{\gamma}$ with respect to the topological flow $\hat{\Phi}$.

4. The stability concept adapted to relatively critical integral curves is the following. For a closed subgroup H of G, a Φ-invariant subset S of M is called H-stable if for every H-invariant neighbourhood U of S in M there exists a neighbourhood V of S in M such that $\Phi_t(m)$ is defined and contained in U for all $m \in V$ and $t \geq 0$. It is called asymptotically H-stable if there exists a neighbourhood V of S in M with $V \times \mathbb{R}_+ \subset \mathscr{D}$ such that for every $m \in V$ and every H-invariant neighbourhood U of S in M there exists $t_0 \in \mathbb{R}$ fulfilling $\Phi_t(m) \in U$ for all $t \geq t_0$. The special case $H = G$ recovers the usual stability concepts of Definition 3.8.1 for the projected flow[31] $\hat{\Phi}$: S is G-stable (asymptotically G-stable) iff $\pi(S)$ is stable (asymptotically stable) under the projected flow $\hat{\Phi}$ (Exercise 6.8.4). As a consequence of this and of the observation that the proof of Proposition 3.8.5 carries over word by word to the topological flow $\hat{\Phi}$ and the periodic integral curve $\hat{\gamma}$, one finds that a relatively periodic integral curve is G-stable iff one of the following equivalent conditions holds:
   (a) it is G-stable under the G-equivariant local diffeomorphism Φ_T of M,
   (b) it is G-stable under an equivariant Poincaré mapping,
   (c) it is stable under a relative Poincaré mapping.

   A similar result holds for asymptotic G-stability. Finally, the concept of Lyapunov function can be adapted to H-stability as follows. An H-Lyapunov function is a continuous H-invariant function $f : U \to \mathbb{R}$ on an H-invariant neighbourhood U of γ in M with the following properties:
   (a) $f(m) = 0$ for all $m \in \gamma$ and $f(m) > 0$ for all $m \in U \setminus (H \cdot \gamma)$,
   (b) $f(\Phi_t(m)) \leq f(m)$ for all $m \in U \setminus (H \cdot \gamma)$ and $t \geq 0$ satisfying $\Phi_s(m) \in U$ for all $s \in [0, t]$.

   If the second condition holds with $f(\Phi_t(m)) < f(m)$, the function f is called an H-Lyapunov function in the strong sense. If f is differentiable, the second condition is equivalent to $X_m(f) \leq 0$ or $X_m(f) < 0$, respectively. Under the assumption that the Lie algebra of H contains all the drift velocities of γ (in case γ is a relative equilibrium) or that H contains all relative phases (in case γ is relatively periodic), one can show the following. If there exists an H-Lyapunov function for γ, then γ is H-stable. If there exists an H-Lyapunov function in the strong sense, then γ is asymptotically H-stable, see Exercise 6.8.5.

Exercises

6.8.1 Prove the relations in (6.8.3).

6.8.2 Prove Formula (6.8.4).

6.8.3 Prove Formula (6.8.5).

[31] These concepts are of topological nature and thus carry over word by word to topological flows.

Hint. Use an $\mathrm{Ad}(G_\gamma)$-invariant scalar product on $\mathfrak{g}$ to decompose $\mathrm{N}_{\mathfrak{g}}(\mathfrak{g}_\gamma) = \mathfrak{g}_\gamma \oplus (\mathfrak{g}_\gamma^\perp \cap \mathrm{N}_{\mathfrak{g}}(\mathfrak{g}_\gamma))$ and show that $\mathfrak{g}_\gamma^\perp \cap \mathrm{N}_{\mathfrak{g}}(\mathfrak{g}_\gamma) \subset \mathrm{C}_{\mathfrak{g}}(\mathfrak{g}_\gamma)$.

6.8.4 Prove that a subset S of M is G-stable (asymptotically G-stable) under the flow Φ of an invariant vector field iff $\pi(S)$ is stable (asymptotically stable) under the projected flow on the orbit space $\hat{M}$.

Hint. Use that the domain $\hat{\mathscr{D}}$ of $\hat{\Phi}$ is given by the orbit space of the action of G on $\mathscr{D}$ and that, via π, G-invariant neighbourhoods of $S \subset M$ in M bijectively correspond to neighbourhoods of $\pi(S)$ in $\hat{M}$.

6.8.5 Prove the criterion for H-stability in terms of an H-Lyapunov function stated in Remark 6.8.5/4. Here is an outline: the assumptions imply that γ is relatively critical with respect to the action of H. Hence, it is enough to prove the statement for the case $H = G$. Since the G-stability of γ is equivalent to the stability of $\hat{\gamma}$ under $\hat{\Phi}$, it suffices to observe that Definition 3.8.15 and Theorem 3.8.16 carry over literally from differentiable to continuous flows and to check that $\hat{f}$ is a Lyapunov function for $\hat{\gamma}$ in the topological sense.

Chapter 7
Linear Symplectic Algebra

In this chapter, we present linear symplectic algebra, starting with a discussion of the elementary properties of subspaces of a symplectic vector space and of the symplectic group. We also present linear symplectic reduction. In the second part of this chapter, we come to some more advanced topics, all related to the study of the space of Lagrangian subspaces of a given symplectic vector space. In particular, the Maslov index and the Kashiwara index will be presented in some detail. These topological invariants will play an essential role in Chap. 12, in the context of geometric asymptotics.

7.1 Symplectic Vector Spaces

Let V be a finite-dimensional vector space over $\mathbb{R}$ and let $\omega \in \bigwedge^2 V^*$. The bilinear form ω induces a linear mapping

$$\omega^\flat : V \to V^*, \qquad \langle \omega^\flat(v), u \rangle := \omega(v, u), \quad v, u \in V.$$

We define the kernel and the rank of ω to be the kernel and the rank of $\omega^\flat$, respectively.[1] The form ω is called non-degenerate if $\omega^\flat$ is an isomorphism. This is equivalent to $\operatorname{rank}\omega = \dim V$ or $\ker\omega = 0$, that is, vanishing of $\omega(u, v)$ for all $v \in V$ implies $u = 0$. If $\omega^\flat$ is an isomorphism, we denote its inverse by $\omega^\sharp$. If there is no danger of confusion, we often write $v^\flat \equiv \omega^\flat(v)$ and $\rho^\sharp \equiv \omega^\sharp(\rho)$.

Remark 7.1.1 Let $\{e_i\}$ be a basis in V and let $\{e^{*i}\}$ be the dual basis in V^*. Then,

$$\omega = \frac{1}{2}\omega_{ij} e^{*i} \wedge e^{*j}, \qquad \omega^\flat(e_i) = \omega_{ij} e^{*j}, \qquad \operatorname{rank}(\omega) = \operatorname{rank}(\omega_{ij}), \qquad (7.1.1)$$

[1] This is consistent with the definition of the kernel of a multilinear form given in (4.2.12).

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7_7, © Springer Science+Business Media Dordrecht 2013

where $\omega_{ij} := \omega(e_i, e_j)$. Let $\tilde{V} = V/\ker\omega^\flat$ and denote the induced bilinear form $\tilde{V} \times \tilde{V} \to \mathbb{R}$ by $\tilde{\omega}$. Obviously, $\tilde{\omega}$ is antisymmetric and non-degenerate. Let $\{\tilde{e}_1, \ldots, \tilde{e}_r\}$ be a basis in $\tilde{V}$ and let $\tilde{\omega}_{ij} = \tilde{\omega}(\tilde{e}_i, \tilde{e}_j)$. Since

$$\det(\tilde{\omega}_{ij}) = \det(-\tilde{\omega}_{ji}) = (-1)^r \det(\tilde{\omega}_{ji}) = (-1)^r \det(\tilde{\omega}_{ij}),$$

we have $(-1)^r = 1$, that is, $r = \dim\tilde{V} = \operatorname{rank}\omega$ is even.

For a subspace $W \subset V$, the ω-orthogonal subspace is defined by

$$W^\omega := \big\{v \in V : \omega(v, u) = 0 \text{ for all } u \in W\big\}. \tag{7.1.2}$$

Proposition 7.1.2 *For every antisymmetric bilinear form ω on a finite-dimensional real vector space V, there exists an ordered basis $\{e_i\}$ in V such that*

$$\omega = \sum_{i=1}^{n} e^{*i} \wedge e^{*(i+n)}, \tag{7.1.3}$$

that is, ω_{ij} has the form

$$\omega_{ij} = \begin{bmatrix} J_n & 0 \\ 0 & 0 \end{bmatrix}$$

with

$$J_n = \begin{bmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{bmatrix}. \tag{7.1.4}$$

Proof We carry out the following iterative procedure, starting with $V_0 := V$. If $\omega = 0$ on V_0, we can choose any basis in V_0 and thus we are done. Otherwise, there exist $v_1, u_1 \in V_0$ such that $\omega(v_1, u_1) = 1$. By bilinearity, v_1 and u_1 are nonzero. By antisymmetry, they cannot be parallel. Hence, they are linearly independent. Let E_1 be the subspace spanned by v_1 and u_1. We show

$$E_1 \cap E_1^\omega = 0, \qquad E_1 + E_1^\omega = V_0.$$

For the first equation, we decompose $v \in E_1 \cap E_1^\omega$ as $v = \alpha v_1 + \beta u_1$ and calculate $\beta = -\omega(v, v_1) = 0$ and $\alpha = \omega(v, u_1) = 0$. For the second equation, let $v \in V_0$. Write

$$v = \big(\omega(v, u_1)v_1 - \omega(v, v_1)u_1\big) + \big(v - \omega(v, u_1)v_1 + \omega(v, v_1)u_1\big)$$

and check that the second term is contained in E_1^ω.

Next, we replace V_0 by $V_1 := E_1^\omega$ and iterate the argument until we arrive at a subspace V_n (possibly trivial) on which ω vanishes. This yields $2n$ linearly independent vectors $e_i := v_i$, $e_{n+i} := u_i$, $i = 1, \ldots, n$, satisfying

$$\omega(e_i, e_j) = \omega(e_{n+i}, e_{n+j}) = 0, \qquad \omega(e_i, e_{n+j}) = \delta_{ij}.$$

In case $V_n \neq 0$, we choose a basis $\{e_{2n+1}, \ldots, e_{\dim V}\}$ in V_n. Then, $\{e_1, \ldots, e_{\dim V}\}$ is a basis in V with the desired properties. □

Definition 7.1.3 (Symplectic vector space)

1. A symplectic vector space is a pair (V, ω) consisting of a real vector space V and a non-degenerate bilinear form ω, called the symplectic form. A basis in V for which ω has the canonical form (7.1.3) is called symplectic or canonical.
2. Let (V_1, ω_1) and (V_2, ω_2) be symplectic vector spaces. A linear mapping $f : V_1 \to V_2$ is called symplectic if $f^*\omega_2 = \omega_1$. A bijective symplectic mapping is called a symplectomorphism.

Note that the inverse of a symplectomorphism is symplectic. Hence, if one takes the symplectic mappings as morphisms of symplectic vector spaces, the symplectomorphisms are the corresponding isomorphisms.

Example 7.1.4 Let $V = \mathbb{R}^{2n}$. The bilinear form $\omega_0(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T J_n \mathbf{y}$ obviously defines a symplectic vector space structure and the standard basis is symplectic. We call J_n the standard symplectic matrix of $\mathbb{R}^{2n}$ and the pair $(\mathbb{R}^{2n}, \omega_0)$ the canonical symplectic vector space structure in dimension $2n$.

Proposition 7.1.2 yields

Corollary 7.1.5 *Let (V, ω) be a symplectic vector space of dimension $2n$. Every symplectic basis of V defines a symplectomorphism onto $(\mathbb{R}^{2n}, \omega_0)$.*

Example 7.1.6 Let W be a vector space and let W^* be its dual space. Then, $V = W \oplus W^*$ endowed with the bilinear form

$$\omega_{W \oplus W^*}(v \oplus \rho, u \oplus \sigma) := \rho(u) - \sigma(v) \tag{7.1.5}$$

is a symplectic vector space (Exercise 7.1.1). This form is referred to as the canonical symplectic form on $W \oplus W^*$.

The following proposition yields another criterion for the non-degeneracy of antisymmetric bilinear forms.

Proposition 7.1.7 *An antisymmetric bilinear form ω on a real vector space V is non-degenerate iff $\dim V = 2n$ and $\omega^n := \omega \wedge \cdots \wedge \omega \neq 0$.*

Proof If ω is non-degenerate, by Proposition 7.1.2, there exists a basis $\{e_k\}$ in V such that $\omega = \sum_{i=1}^n e^{*i} \wedge e^{*(i+n)}$. Then,

$$\omega^n = n! \cdot (-1)^{\frac{n(n-1)}{2}} \cdot e^{*1} \wedge e^{*2} \wedge \cdots \wedge e^{*(2n-1)} \wedge e^{*2n} \neq 0.$$

Conversely, assume that $\dim V = 2n$ and $\omega^n \neq 0$ and let ω be degenerate. Then, there exists a vector $0 \neq e_1 \in V$ with $\omega^\flat(e_1) = 0$. If we extend e_1 to a basis $\{e_1, \dots, e_{2n}\}$, then $\omega^n(e_1, \dots, e_{2n}) = 0$, in contradiction to the first assumption. $\square$

Definition 7.1.8 (Canonical volume form) Let (V, ω) be a symplectic vector space. The $2n$-form

$$\Omega_\omega := \frac{1}{n!}(-1)^{\frac{n(n-1)}{2}}\omega^n \tag{7.1.6}$$

is called induced canonical volume form on V.

In a symplectic basis $\{e_i\}$ of V, one has $\Omega_\omega = e^{*1} \wedge \cdots \wedge e^{*2n}$.

Exercises

7.1.1 Let W be a vector space and let W^* be its dual. Prove that the bilinear form $\omega_{W \oplus W^*}$ defined by (7.1.5) is symplectic.

7.2 Subspaces of a Symplectic Vector Space

Let (V, ω) be a symplectic vector space. To start with, we present a number of important relations characterizing subspaces of a symplectic vector space, their annihilators and their symplectic orthogonal complements. For that purpose, recall that the annihilator $W^0 \subset V^*$ of a subspace W of V is defined by

$$W^0 := \{\eta \in V^* : \langle \eta, u \rangle = 0 \text{ for all } u \in W\}.$$

From linear algebra we recall (Exercise 7.2.1)

$$\dim W + \dim W^0 = \dim V, \qquad (W^0)^0 = W, \qquad (V/W)^* \cong W^0. \tag{7.2.1}$$

Proposition 7.2.1 *For subspaces W, W_1 and W_2 of a symplectic vector space (V, ω), one has*

1. $(W^\omega)^\flat = W^0$,
2. $\dim W + \dim W^\omega = \dim V$,
3. $(W^\omega)^\omega = W$,
4. *If* $W_1 \subset W_2$, *then* $W_2^\omega \subset W_1^\omega$,
5. $W_1^\omega \cap W_2^\omega = (W_1 + W_2)^\omega$ *and* $(W_1 \cap W_2)^\omega = W_1^\omega + W_2^\omega$.

Proof 1. Obviously, $(W^\omega)^\flat \subset W^0$ and $(W^0)^\sharp \subset W^\omega$.

2. This follows from point 1 and from (7.2.1).

3. The inclusion relation $W \subset (W^\omega)^\omega$ is immediate. Since application of point 2 to both W and W^ω yields $\dim(W^\omega)^\omega = \dim W$, the assertion follows.

4. This is obvious.

5. It suffices to prove the first equation, because the second one follows by replacing W_i by W_i^ω and by using point 3. Due to $W_i \subset W_1 + W_2$ and $W_i^\omega \subset W_1^\omega + W_2^\omega$, point 4 implies

$$(W_1 + W_2)^\omega \subset W_1^\omega \cap W_2^\omega, \qquad \big(W_1^\omega + W_2^\omega\big)^\omega \subset W_1 \cap W_2.$$

Hence, to prove the assertion it suffices to show that

$$\dim(W_1 + W_2)^\omega + \dim\big(W_1^\omega + W_2^\omega\big)^\omega = \dim(W_1 \cap W_2) + \dim\big(W_1^\omega \cap W_2^\omega\big).$$

Using point 2 and $\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$, we can write

$$\dim(W_1 + W_2)^\omega = \dim V - \dim W_1 - \dim W_2 + \dim(W_1 \cap W_2),$$
$$\dim\big(W_1^\omega + W_2^\omega\big)^\omega = \dim V - \dim W_1^\omega - \dim W_2^\omega + \dim\big(W_1^\omega \cap W_2^\omega\big).$$

Addition of these two equations yields the desired equality. □

The following types of subspaces of a symplectic vector space are important.

Definition 7.2.2 Let (V, ω) be a symplectic vector space and let $W \subset V$ be a subspace. W is called

1. isotropic if $W \subset W^\omega$,
2. coisotropic if $W^\omega \subset W$,
3. Lagrange if $W^\omega = W$,
4. symplectic if $W \cap W^\omega = \{0\}$.

Let W be a subspace of a symplectic vector space (V, ω) and let ω_W be the bilinear form on W induced by restriction of ω. In terms of ω_W, the several types of subspaces can be characterized as follows.

Proposition 7.2.3 *A subspace W of a symplectic vector space (V, ω) is*

1. *isotropic iff $\omega_W = 0$,*
2. *coisotropic iff W^ω is isotropic, that is, iff $\omega_{W^\omega} = 0$,*
3. *Lagrangian iff $\omega_W = \omega_{W^\omega} = 0$,*
4. *symplectic iff ω_W or ω_{W^ω} is a symplectic form.*

The key for the analysis of the algebraic properties of W is the study of the rank of ω_W which is called the symplectic rank of W. By Proposition 7.1.2, this is an even number. Since

$$\ker \omega_W \equiv \ker(\omega_W)^\flat = W \cap W^\omega, \tag{7.2.2}$$

we have

$$\operatorname{rank} \omega_W = \dim W - \dim \ker \omega_W = \dim W - \dim\big(W \cap W^\omega\big), \tag{7.2.3}$$

and a similar equation for ω_{W^ω}. Subtracting these two equations, we get

$$\operatorname{rank}\omega_W - \operatorname{rank}\omega_{W^\omega} = \dim W - \dim W^\omega = 2\dim W - \dim V. \tag{7.2.4}$$

This implies

$$\max(0, 2\dim W - \dim V) \leq \operatorname{rank}\omega_W \leq \dim W. \tag{7.2.5}$$

From (7.2.5) and (7.2.4) we read off that W and W^ω reach their minimal or maximal rank always simultaneously. Moreover, Eq. (7.2.4) yields the following

Proposition 7.2.4 *Let (V, ω) be a symplectic vector space of dimension $2n$.*

1. *The dimension of an isotropic (coisotropic) subspace is at most (at least) equal to n. The dimension of a Lagrangian subspace is equal to n.*
2. *Every subspace of dimension 1 (of codimension 1) is isotropic (coisotropic).*
3. *A subspace of dimension $m \geq n$ is coisotropic iff its symplectic rank is $2(m-n)$.*
4. *Every subspace of an isotropic subspace is isotropic and every subspace which contains a coisotropic subspace is coisotropic.*

Example 7.2.5 Let $V = \mathbb{R}^{2n}$, $\omega_0(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T J_n \mathbf{y}$ and let $\{\mathbf{e}_i\}$, $i = 1, \ldots, 2n$, be the canonical basis.

1. Subspaces which are spanned by vectors $\mathbf{e}_i$ with $i \leq n$ or by vectors $\mathbf{e}_i$ with $i > n$ are isotropic.
2. Subspaces which contain all vectors $\mathbf{e}_i$ with $i \leq n$ or all vectors $\mathbf{e}_i$ with $i > n$ are coisotropic.
3. The subspace spanned by $\mathbf{e}_1, \ldots, \mathbf{e}_n$ and the subspace spanned by $\mathbf{e}_{n+1}, \ldots, \mathbf{e}_{2n}$ are Lagrangian.
4. Subspaces which are spanned by pairs $\mathbf{e}_i, \mathbf{e}_{n+i}$ with $i \leq n$ are symplectic.

The following two propositions characterize Lagrangian subspaces.

Proposition 7.2.6 *Let W be a subspace of a symplectic vector space (V, ω). The following statements are equivalent.*

1. *W is a Lagrangian subspace.*
2. *W is isotropic and* $\dim W = \frac{1}{2}\dim V$.
3. *W is isotropic and possesses an isotropic complement in V.*

Proof 1 $\Rightarrow$ 3: Of course, W is isotropic. An isotropic complement can be constructed by the following procedure. Choose a nonzero $v_1 \in V \setminus W$ and let $V_1 := \mathbb{R}v_1$. We check

(a) V_1 is isotropic: this follows from Proposition 7.2.4/2.
(b) $W + V_1^\omega = V$: by Proposition 7.2.1/5, we have $W + V_1^\omega = (W^\omega \cap V_1)^\omega$. Hence, the assertion follows from $W^\omega = W$ and $W \cap V_1 = 0$.

If now $W + V_1 = V$, we are done. Otherwise, by point (b), there exists a nonzero $v_2 \in V_1^\omega \setminus (W + V_1)$. Let $V_2 := V_1 + \mathbb{R}v_2$. We check that (a) and (b) hold for V_2. For (a), this follows from the isotropy of V_1 and from the fact that $v_2 \in V_1^\omega$. The argument for (b) is the same as before. Now replace V_1 by V_2 and iterate the argument to finally arrive at an isotropic subspace V_n satisfying $W + V_n = V$. By construction, V_n is an isotropic complement of W in V.

$3 \Rightarrow 2$: This follows from Proposition 7.2.4/1.

$2 \Rightarrow 1$: Since $W \subset W^\omega$ and $\dim W^\omega = \dim V - \dim W = \dim W$, we conclude $W = W^\omega$. □

Points 2 and 3 of Proposition 7.2.6 imply that every Lagrangian subspace has a Lagrangian complement.

Remark 7.2.7 The above result can be generalized as follows (Exercise 7.2.3). For every tuple $(L_1, \dots, L_r)$ of Lagrangian subspaces there exists a Lagrangian subspace L_0 which is transversal to all elements of the tuple, that is,

$$L_0 \cap L_i = \{0\}, \quad i = 1, \dots, r. \tag{7.2.6}$$

Example 7.2.8 For the canonical symplectic vector space $W \oplus W^*$, the subspaces $W \oplus \{0\}$ and $\{0\} \oplus W^*$ are obviously complementary Lagrangian subspaces.

Proposition 7.2.9 *Let (V, ω) be a symplectic vector space and let $V = W \oplus W'$ be a decomposition into complementary Lagrangian subspaces. Then, this decomposition induces a symplectomorphism onto the canonical symplectic vector space $W \oplus W^*$.*

Proof Let $\chi : W' \to W^*$ be the natural mapping induced by ω,

$$\langle \chi(u), v \rangle := \omega(u, v). \tag{7.2.7}$$

Note that this is an isomorphism of vector spaces. Thus, $\mathbb{1}_W \oplus \chi : W \oplus W' \to W \oplus W^*$ is an isomorphism, too. Finally, for $u, v \in W$ and $u', v' \in W'$, we have

$$\omega(u + u', v + v') = \omega(u', v) - \omega(v', u) = \langle \chi(u'), v \rangle - \langle \chi(v'), u \rangle,$$

that is, $\mathbb{1}_W \oplus \chi$ is symplectic. □

Exercises

7.2.1 Prove the relations stated in (7.2.1).

7.2.2 Let (V, ω) be a symplectic vector space of dimension $2n$. Prove the following statements.

(a) Every basis of an isotropic subspace can be extended to a symplectic basis of V.

(b) Every isotropic subspace is contained in a Lagrangian subspace. Correspondingly, every coisotropic subspace contains a Lagrangian subspace.

(c) If (W_1, W_2) and (U_1, U_2) are pairs of complementary Lagrangian subspaces, there exists a symplectomorphism which transforms W_i into U_i.

7.2.3 Prove the statement of Remark 7.2.7.

7.3 Linear Symplectic Reduction

Let W be a subspace of the symplectic vector space (V, ω). Since $\ker \omega_W = W \cap W^\omega$, the quotient

$$\hat{W} := W/\big(W \cap W^\omega\big) \tag{7.3.1}$$

carries a natural symplectic structure $\hat{\omega}$ defined by

$$\hat{\omega}\big([u], [v]\big) := \omega(u, v)$$

for all $u, v \in W$. Let $p \colon W \to \hat{W}$ be the canonical projection. Then,

$$\omega_W = p^* \hat{\omega} \tag{7.3.2}$$

and ω_W and $\hat{\omega}$ have the same rank. If W is coisotropic, then

$$\hat{W} = W/W^\omega. \tag{7.3.3}$$

Proposition 7.3.1 *Let (V, ω) be a symplectic vector space. Let $W \subset V$ be coisotropic and let $L \subset V$ be Lagrange. Then, the image of $W \cap L$ under the canonical projection $p \colon W \to \hat{W}$ is a Lagrangian subspace of $(\hat{W}, \hat{\omega})$.*

Proof Since L is isotropic, so is $W \cap L$. Hence, (7.3.2) implies that $p(W \cap L)$ is isotropic, too. Let us calculate the dimension. Since $W^\omega \subset W$, we have $p(W \cap L) \cong (W \cap L)/(W^\omega \cap L)$ as vector spaces. Using points 2 and 5 of Proposition 7.2.1, we get

$$\dim(W \cap L) = \dim V - \dim(W \cap L)^\omega = \dim V - \dim W^\omega - \dim L + \dim\big(W^\omega \cap L\big)$$

and a similar equation for $W^\omega \cap L$. Subtracting these equations and using (7.3.3), we obtain

$$\dim(W \cap L) - \dim\big(W^\omega \cap L\big) = \frac{1}{2}\big(\dim W - \dim W^\omega\big) = \frac{1}{2} \dim \hat{W}.$$

Thus, Proposition 7.2.6 yields the assertion. □

Lemma 7.3.2 *Let W be a subspace of the symplectic vector space (V, ω). If $E \subset W$ is a subspace complementary to $W \cap W^\omega$, that is, $W = E \oplus (W \cap W^\omega)$, then E is maximally symplectic in W.*

Proof Since $p_{\restriction E}$ is injective and since $\hat{\omega}$ is non-degenerate, (7.3.2) implies that E is symplectic. Suppose there exists a symplectic subspace E' such that $E \subset E' \subset W$ and $\dim E' > \dim E$. Then, for dimensional reasons, $E' \cap (W \cap W^\omega) \neq \{0\}$. Choose a nonzero $v \in E' \cap (W \cap W^\omega)$. Then, $v \in W^\omega \subset E'^\omega$ and therefore $\omega(v, v') = 0$ for all $v' \in E'$. This contradicts the assumption on E' to be symplectic. □

As an important conclusion we obtain that every subspace of a symplectic vector space (V, ω) induces a decomposition of V into a direct sum of ω-orthogonal symplectic subspaces.

Theorem 7.3.3 (Witt-Artin decomposition) *Let (V, ω) be a symplectic vector space and let W be a subspace of V. Let E and F be subspaces such that*

$$W = E \oplus \big(W \cap W^\omega\big) \quad \textit{and} \quad W^\omega = F \oplus \big(W \cap W^\omega\big).$$

Then, E, F and $(E \oplus F)^\omega$ are symplectic and V decomposes into the ω-orthogonal direct sum

$$V = E \oplus F \oplus (E \oplus F)^\omega. \tag{7.3.4}$$

Moreover, $W \cap W^\omega$ is a Lagrangian subspace of $(E \oplus F)^\omega$.

We will see that the above decomposition plays an important role in symplectic reduction of systems with symmetries, cf. Chap. 10.

Proof The subspaces E and F are symplectic due to Lemma 7.3.2. Since $E \cap F \subset W \cap W^\omega$ and $E \cap (W \cap W^\omega) = \{0\}$, we obtain $E \cap F = \{0\}$. Since $F \subset W^\omega$, we have $W \subset F^\omega$ and thus $E \subset F^\omega$. Consequently, $F \subset E^\omega$. Therefore, E and F are ω-orthogonal.

Let us denote $Z \equiv (E \oplus F)^\omega$. Since $E \oplus F$ is symplectic, Z is symplectic, too. By Proposition 7.2.1/5, we have $Z = F^\omega \cap E^\omega \supset W \cap W^\omega$. Moreover, by Proposition 7.2.3/1, $W \cap W^\omega$ is isotropic. Finally, due to

$$\dim V = \dim W + \dim W^\omega = \dim E + \dim F + 2\dim\big(W \cap W^\omega\big),$$

we obtain $\dim(W \cap W^\omega) = \frac{1}{2}\dim Z$. Therefore, $W \cap W^\omega$ is a Lagrangian subspace in Z. □

Example 7.3.4 Consider $\mathbb{R}^6$ with the standard symplectic structure. Denote the standard basis elements by $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{f}_1 \equiv \mathbf{e}_4, \mathbf{f}_2 \equiv \mathbf{e}_5, \mathbf{f}_3 \equiv \mathbf{e}_6$ and choose the subspace

$$W = \mathbb{R}\mathbf{e}_1 \oplus \mathbb{R}\mathbf{e}_2 \oplus \mathbb{R}\mathbf{f}_1.$$

Then,

$$W^\omega = \mathbb{R}\mathbf{e}_2 \oplus \mathbb{R}\mathbf{e}_3 \oplus \mathbb{R}\mathbf{f}_3, \qquad W \cap W^\omega = \mathbb{R}\mathbf{e}_2.$$

We choose $E = \mathbb{R}\mathbf{e}_1 \oplus \mathbb{R}\mathbf{f}_1$ and $F = \mathbb{R}\mathbf{e}_3 \oplus \mathbb{R}\mathbf{f}_3$. Then, $(E \oplus F)^\omega = \mathbb{R}\mathbf{e}_2 \oplus \mathbb{R}\mathbf{f}_2$ and (7.3.4) holds, indeed.

Remark 7.3.5

1. The Witt-Artin decomposition (7.3.4) induces the following decomposition of the symplectic form:

$$\omega = \omega^E + \omega^F + \omega^{(E\oplus F)^\omega}. \tag{7.3.5}$$

2. Applying Proposition 7.2.9 to the Lagrangian subspace $W \cap W^\omega$ of $(E \oplus F)^\omega$, we obtain

$$(E \oplus F)^\omega \cong W \cap W^\omega \oplus \big(W \cap W^\omega\big)^*. \tag{7.3.6}$$

3. In particular, let W be a Lagrangian subspace of the symplectic vector space (V, ω). Then, we have $E = \{0\} = F$ and therefore $(E \oplus F)^\omega = V$, thus, in this case the Witt-Artin decomposition induced by W is trivial.

7.4 The Symplectic Group

In this section, we discuss symplectic mappings. The following proposition collects their elementary properties.

Proposition 7.4.1 *Let (V, ω) and (W, ρ) be symplectic vector spaces and let $f: V \to W$ be a symplectic mapping.*

1. *The mapping f is injective. If* $\dim V = \dim W$, *then f is a symplectomorphism.*
2. *The image of f is a symplectic subspace of W.*
3. *Every symplectomorphism preserves the canonical symplectic volume form.*
4. *If $(V, \omega) = (W, \rho)$, then* $\det f = 1$.

Proof 1. Let $v \in V$ such that $f(v) = 0$. Then,

$$\omega(v, u) = f^*\rho(v, u) = \rho\big(f(v), f(u)\big) = 0$$

for all $u \in V$. Since ω is non-degenerate, we conclude $v = 0$.

2. Let $w \in \operatorname{im} f \cap (\operatorname{im} f)^\rho$ and let $v \in V$ such that $w = f(v)$. Then, $\rho(f(v), f(u)) = \omega(v, u) = 0$ for all $u \in V$. It follows that $v = 0$ and hence $w = 0$.

3. By $f^*\rho = \omega$, we have

$$f^*\Omega_\rho = \frac{1}{n!}(-1)^{\frac{n(n-1)}{2}} f^*\rho \wedge \cdots \wedge f^*\rho = \Omega_\omega.$$

4. By definition of the determinant, $f^*\Omega_\omega = (\det f)\Omega_\omega$. Hence, point 3 implies $\det f = 1$. □

Remark 7.4.2

1. As a generalization of point 4 of Proposition 7.4.1, let V and W be vector spaces of the same dimension, $\dim V = \dim W = k$, with volume forms Ω_V and Ω_W.

Let $f \in L(V, W)$ be a linear mapping. Then, there exists a unique constant $\det(f)$ such that

$$f^* \Omega_W = \det(f) \cdot \Omega_V ,$$

because $\bigwedge^k V$ is one-dimensional. Thus, f is volume preserving iff $\det(f) = 1$.

2. Since on a 2-dimensional symplectic vector space the canonical volume form coincides with the symplectic form, every volume preserving mapping is automatically symplectic. However, in higher dimensions this is not the case as the following examples shows. Let $V = \mathbb{R}^4$, let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ denote the canonical basis and consider the canonical symplectic form $\omega = \mathbf{e}^{*1} \wedge \mathbf{e}^{*3} + \mathbf{e}^{*2} \wedge \mathbf{e}^{*4}$. The linear mapping $f : V \to V$ defined by

$$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4) \mapsto (-\mathbf{e}_1, -\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$$

preserves the volume $\Omega_\omega = \mathbf{e}^{*1} \wedge \mathbf{e}^{*2} \wedge \mathbf{e}^{*3} \wedge \mathbf{e}^{*4}$, but since $f^*\omega = -\omega$, it is not symplectic.

The automorphisms of a symplectic vector space (V, ω) form a closed subgroup of $\mathrm{GL}(V)$, called the symplectic group of (V, ω) and denoted by $\mathrm{Sp}(V, \omega)$. According to Theorem 5.6.8, $\mathrm{Sp}(V, \omega)$ is a Lie group and a Lie subgroup of $\mathrm{GL}(V)$. Let $\mathfrak{sp}(V, \omega)$ denote the Lie algebra of $\mathrm{Sp}(V, \omega)$. Recall that in a symplectic basis $\{e_i\}$ of V we have $\omega = J_n$, cf. Example 7.1.4. Let $f \in \mathrm{Sp}(V, \omega)$ and let a be the matrix representing f in this basis, $f(e_i) = a^j{}_i e_j$. The condition $f^*\omega = \omega$ takes the form

$$a^{\mathrm{T}} J_n a = J_n . \tag{7.4.1}$$

Thus, every symplectic basis induces an isomorphism of the Lie groups $\mathrm{Sp}(V, \omega)$ and $\mathrm{Sp}(n, \mathbb{R})$, cf. Example 1.2.6.

For the analysis of the stability of equilibria or periodic integral curves of Hamiltonian systems, the properties of the spectrum of symplectomorphisms are relevant. The following proposition puts strong limits on how eigenvalues of symplectomorphisms are located in the complex plane.

Proposition 7.4.3 (Symplectic Eigenvalue Theorem) *Let (V, ω) be a symplectic vector space.*

1. *If λ is an eigenvalue of $a \in \mathrm{Sp}(V, \omega)$ with multiplicity k, then both the complex conjugate $\overline{\lambda}$ and λ^{-1} are eigenvalues of a with multiplicity k.*
2. *If λ is an eigenvalue of $A \in \mathfrak{sp}(V, \omega)$ with multiplicity k, then both $\overline{\lambda}$ and $-\lambda$ are eigenvalues of A with multiplicity k.*

Proof It suffices to prove the assertion for $V = \mathbb{R}^{2n}$ with the canonical symplectic structure. The assertions about $\overline{\lambda}$ are due to the fact that a and A are real matrices.

1. First, note that $\lambda \neq 0$, because a is invertible. Let $\chi_a(z) := \det(a - z\mathbb{1})$ be the characteristic polynomial of a. We rewrite (7.4.1) in the form $(a^{\mathrm{T}})^{-1} = J_n a J_n^{-1}$ to

obtain

$$\chi_a(z) = \det\big(J_n(a - z\mathbb{1})J_n^{-1}\big) = \det\big((a^{\mathrm{T}})^{-1} - z\mathbb{1}\big) = \det\big(a^{-1} - z\mathbb{1}\big).$$

Writing $a^{-1} - z\mathbb{1} = (-za^{-1})(a - z^{-1}\mathbb{1})$ and using $\det(a) = 1$ we arrive at

$$\chi_a(z) = z^{2n}\chi_a\big(z^{-1}\big).$$

Thus, if λ is a zero of χ_a of multiplicity k, so is λ^{-1}.

2. Similarly, due to $J_n^2 = -\mathbb{1}$, from (7.4.1) we conclude $A^{\mathrm{T}} = J_n A J_n$. Using this and $\det J_n = 1$, we obtain

$$\chi_A(-z) = \det(A + z\mathbb{1}) = \det\big(J_n(A + z\mathbb{1})J_n\big) = \det\big(A^{\mathrm{T}} - z\mathbb{1}\big) = \chi_A(z).$$

Thus, if λ is a zero of χ_A of multiplicity k, so is $-\lambda$.

□

Corollary 7.4.4 *Let (V, ω) be a symplectic vector space.*

1. *If ± 1 is an eigenvalue of $a \in \mathrm{Sp}(V, \omega)$, it has even multiplicity.*
2. *If 0 is an eigenvalue of $A \in \mathfrak{sp}(V, \omega)$, it has even multiplicity.*

Proof Let $a \in \mathrm{Sp}(V, \omega)$. Unless $\lambda = \pm 1$, the sum of the multiplicities of the eigenvalues λ, $\overline{\lambda}$, λ^{-1} and $\overline{\lambda}^{-1}$ is even. It follows that the sum of the multiplicities of the eigenvalues $+1$ and -1 (if present) is even, too. A similar argument for A yields that, there, the eigenvalue 0 can occur with even multiplicity only. Moreover, unless $\lambda = -1$, the family of eigenvalues λ, $\overline{\lambda}$, λ^{-1} and $\overline{\lambda}^{-1}$ contributes a unit factor to $\det a$. Since $\det a = 1$, the eigenvalue -1 can occur with even multiplicity only. Hence, this holds for the eigenvalue $+1$, too. □

It follows that the eigenvalues of a symplectomorphism must be located symmetrically with respect to reflections about the unit circle and the real axis. For a given λ, the following cases can occur.

1. If $|\lambda| \neq 1$ and $\operatorname{Im}\lambda \neq 0$, then $\lambda, \overline{\lambda}, \lambda^{-1}$ and $\overline{\lambda}^{-1}$ are all distinct (4 different eigenvalues).
2. If $|\lambda| \neq 1$, $\operatorname{Im}\lambda = 0$, then $\lambda = \overline{\lambda}$ and $\lambda^{-1} = \overline{\lambda}^{-1}$ (pairs of eigenvalues on the real axis).
3. If $|\lambda| = 1$, $\operatorname{Im}\lambda \neq 0$, then $\lambda = \overline{\lambda}^{-1}$ and $\overline{\lambda} = \lambda^{-1}$ (pairs of eigenvalues on the unit circle).
4. If $|\lambda| = 1$, $\operatorname{Im}\lambda = 0$, then $\lambda = \overline{\lambda} = \lambda^{-1} = \overline{\lambda}^{-1} = \pm 1$ (one eigenvalue, $+1$ or -1).

Figure 7.1 illustrates these statements for the case $V = \mathbb{R}^4$ endowed with the canonical symplectic structure.

It turns out that every linear symplectomorphism can be brought to a normal form, see Sect. 3 in [225]. The following special case will be useful in the sequel. The proof is left to the reader (Exercise 7.4.2).

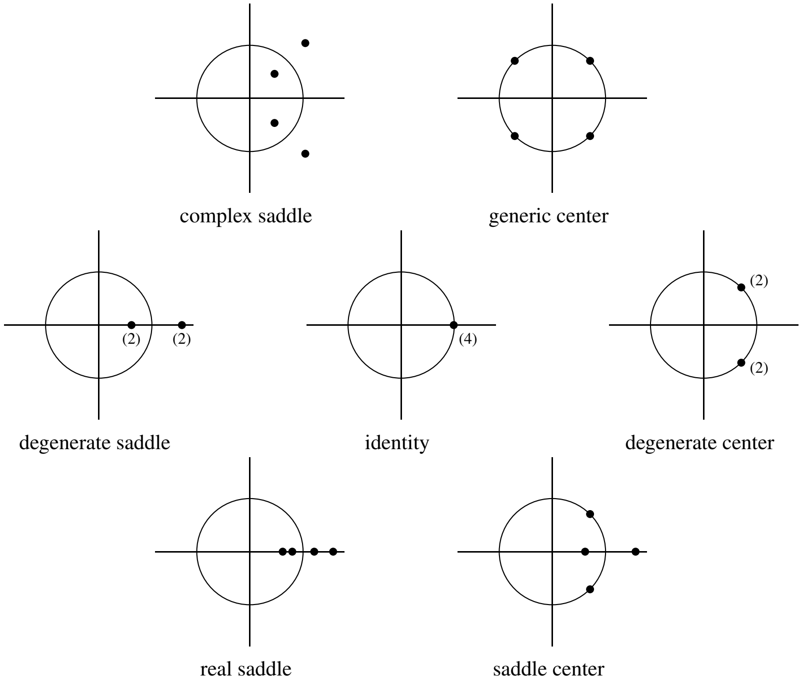


Fig. 7.1 Possible eigenvalue configurations for symplectomorphisms of $\mathbb{R}^4$

Proposition 7.4.5 *Let $f \in \mathrm{Sp}(V, \omega)$, $\dim V = 2n$. If the eigenvalues λ_k of f are all distinct and lie on the unit circle, $\lambda_k = \mathrm{e}^{\mathrm{i}\alpha_k}$, there exists a symplectic basis in V such that f is given by the matrix $\mathrm{e}^{-J_n A}$, where $A = \mathrm{diag}(\alpha_1, \ldots, \alpha_n, \alpha_1, \ldots, \alpha_n)$.*

Finally, we prove that for a compact subgroup $H \subset \mathrm{Sp}(V, \omega)$, the subspace V^H of H-invariant elements is symplectic. Later on, this will be applied to the stabilizers of proper Lie group actions.

Lemma 7.4.6 *Let (V, ω) be a symplectic vector space and let $H \subset \mathrm{Sp}(V, \omega)$ be a compact subgroup. Then, V^H is a symplectic subspace.*

Proof We have to show that $V^H \cap (V^H)^\omega = 0$. Since H is compact, Proposition 5.5.6 implies that V admits an H-invariant scalar product g. Clearly, it suffices to show that

$$\left(V^H\right)^\omega = \left(V^H\right)^\perp, \tag{7.4.2}$$

where $(V^H)^\perp$ denotes the g-orthogonal complement of V^H in V. By Proposition 7.2.1/1, $(V^H)^\omega = ((V^H)^\flat)^0$. Note that H acts on V^* from the left via $(h^{-1})^{\mathrm{T}}$, $h \in H$, and that $H \subset \mathrm{Sp}(V, \omega)$ implies that the mappings $\sharp$ and $\flat$ intertwine this action with the original action of H on V. It follows that $(V^H)^\flat = (V^*)^H$ and hence $(V^H)^\omega = ((V^*)^H)^0$. On the other hand, since g is H-invariant, the induced isomorphism $\mathrm{g} : V \to V^*$ maps V^H onto $(V^*)^H$. This implies

$$\big((V^*)^H\big)^0 = \big(\mathrm{g}\big(V^H\big)\big)^0 = \big(V^H\big)^\perp$$

and thus (7.4.2). □

Exercises

7.4.1 Show that $a = \left[\begin{smallmatrix} B & C \\ D & E \end{smallmatrix}\right]$ belongs to $\mathrm{Sp}(n, \mathbb{R})$ iff $B^\mathrm{T} D$ and $C^\mathrm{T} E$ are symmetric and $B^T E - D^T C = \mathbf{1}$. Here, B, C, D, E are quadratic real matrices.

7.4.2 Prove Proposition 7.4.5.

7.5 Compatible Complex Structures

In symplectic algebra and geometry, complex structures play an important auxiliary role.

Definition 7.5.1 A complex structure on a real vector space V is an endomorphism J with the property $J^2 = -\mathrm{id}$.

If a complex structure exists, V has even dimension and acquires the structure of a complex vector space of dimension $\frac{1}{2} \dim V$ with scalar multiplication by complex numbers defined by

$$(x + \mathrm{i}y)v := xv + yJv, \quad v \in V, \ x, y \in \mathbb{R}.$$

In what follows, we do not distinguish in notation between the real and the complex vector space structure on V. Note that a real linear mapping of V is complex linear iff it commutes with J. Denote by $\mathrm{End}_\mathbb{C}(V)$ the vector space of complex linear mappings and by $\mathrm{GL}_\mathbb{C}(V)$ the group of complex linear automorphisms of V.

Now, let (V, ω) be a symplectic vector space.

Definition 7.5.2 A complex structure J on V is said to be compatible with ω if the bilinear form

$$\mathrm{g}_J : V \times V \to \mathbb{R}, \qquad \mathrm{g}_J(u, v) = \omega(u, Jv) \tag{7.5.1}$$

is a (real) scalar product.

We will prove below that compatible complex structures exist on every symplectic vector space. The isometry group of g_J will be denoted by $\mathrm{O}(V, \mathrm{g}_J)$. Furthermore, we denote the space of ω-compatible complex structures on (V, ω)

by $\mathscr{J}(V,\omega)$ and endow it with the topology induced from $\mathrm{End}(V)$. If J is ω-compatible, it is both symplectic and isometric with respect to g_J:

$$\omega(Ju, Jv) = \mathrm{g}_J(Ju, v) = \mathrm{g}_J(v, Ju) = -\omega(v, u) = \omega(u, v),$$
$$\mathrm{g}_J(Ju, Jv) = -\omega(Ju, v) = \omega(v, Ju) = \mathrm{g}_J(v, u) = \mathrm{g}_J(u, v).$$

Due to $J^2 = -\mathrm{id}$, the first equation implies $\mathscr{J}(V,\omega) \subset \mathfrak{sp}(V,\omega)$. The proof of the following proposition is left to the reader (Exercise 7.5.1).

Proposition 7.5.3 *Let $J \in \mathscr{J}(V,\omega)$.*

1. *J defines a scalar product[2] on the complex vector space V by*

$$\mathrm{h}_J(u, v) := \mathrm{g}_J(u, v) + \mathrm{i}\omega(u, v). \tag{7.5.2}$$

   For the corresponding isometry group $\mathrm{U}(V, \mathrm{h}_J)$, one has

$$\mathrm{U}(V,\mathrm{h}_J) = \mathrm{Sp}(V,\omega) \cap \mathrm{GL}_{\mathbb{C}}(V) = \mathrm{O}(V,\mathrm{g}_J) \cap \mathrm{GL}_{\mathbb{C}}(V) = \mathrm{O}(V,\mathrm{g}_J) \cap \mathrm{Sp}(V,\omega). \tag{7.5.3}$$

2. *For every Lagrangian subspace L of (V,ω), the image of L under J is a complementary Lagrangian subspace, that is, $V = L \oplus JL$. It coincides with the g_J-orthogonal complement $L^\perp$ of L in V.*
3. *Every endomorphism A of a Lagrangian subspace L extends uniquely to a complex linear endomorphism of V by $A(u + Jv) := Au + JAv$, where $u, v \in L$. Every complex linear endomorphism of V is uniquely determined by its values on L.*
4. *Every g_J-orthonormal basis $\{e_i\}$ of a Lagrangian subspace L constitutes an h_J-orthonormal basis in V, viewed as a complex vector space. Moreover, the elements e_i and $f_i := Je_i$ form a g_J-orthonormal symplectic basis in (V,ω).*

Remark 7.5.4

1. Point 2 of Proposition 7.5.3 has the following converse. Let L_1 and L_2 be complementary Lagrangian subspaces. According to Proposition 7.2.9, ω induces an isomorphism $L_1^* \to L_2$. Using this isomorphism and a basis in L_1, one can construct an ω-compatible complex structure J such that $L_2 = JL_1$ (Exercise 7.5.2).
2. Due to $J^2 = -\mathbb{1}$, the eigenvalues of J are $\pm\mathrm{i}$, both with multiplicity $\frac{1}{2}\dim V$. Hence, for any two complex structures J and J', there exists $a \in \mathrm{GL}(V)$ such that $J' = aJa^{-1}$. If both J and J' are ω-compatible, a can be chosen from $\mathrm{Sp}(V,\omega)$.

Example 7.5.5 The matrices $\pm J_n$ are complex structures on $V = \mathbb{R}^{2n}$. The matrix $J = -J_n$ is compatible with the standard symplectic structure of $\mathbb{R}^{2n}$ defined by J_n. The corresponding scalar product g_J coincides with the standard scalar product on

[2] That is, a positive definite Hermitian form.

$\mathbb{R}^{2n}$. With respect to the complex vector space structure induced on $\mathbb{R}^{2n}$ by J, the mapping $\mathbb{R}^{2n} \equiv \mathbb{R}^n \oplus \mathbb{R}^n \to \mathbb{C}^n$, given by $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} + \mathrm{i}\mathbf{y}$, is an isomorphism of complex vector spaces. Via this isomorphism, h_J corresponds to the standard scalar product on $\mathbb{C}^n$. The induced mapping $\mathrm{M}_n(\mathbb{C}) \to \mathrm{M}_{2n}(\mathbb{R})$, defined by

$$A + iB \mapsto \begin{bmatrix} A & -B \\ B & A \end{bmatrix}, \quad A, B \in \mathrm{M}_n(\mathbb{R}), \tag{7.5.4}$$

yields an isomorphism of real algebras from $\mathrm{M}_n(\mathbb{C})$ onto the subalgebra of $\mathrm{M}_{2n}(\mathbb{R})$ of elements commuting with $-J_n$. This isomorphism maps $\mathrm{U}(n)$ onto $\mathrm{U}(V, \mathsf{h}_J)$. Under the identification of $\mathrm{M}_n(\mathbb{C})$ with its image in $\mathrm{M}_{2n}(\mathbb{R})$, the identities (7.5.3) read

$$\mathrm{U}(n) = \mathrm{Sp}(n, \mathbb{R}) \cap \mathrm{GL}(n, \mathbb{C}) = \mathrm{O}(2n) \cap \mathrm{GL}(n, \mathbb{C}) = \mathrm{O}(2n) \cap \mathrm{Sp}(n, \mathbb{R}).$$

The Lagrangian subspace of $\mathbb{R}^{2n}$ defined by the first n standard basis vectors is mapped under J to that spanned by the last n basis vectors. The counterparts of these Lagrangian subspaces in $\mathbb{C}^n$ are given by the real subspaces $\mathbb{R}^n$ and $\mathrm{i}\mathbb{R}^n$, respectively.

Denote the space of symmetric bilinear forms on V by $\mathrm{S}^2 V^*$ and the convex open subset of positive definite elements by $\mathrm{S}^2_+ V^*$.

Proposition 7.5.6 *Every symplectic vector space (V, ω) admits a compatible complex structure. More precisely, there exists a surjective and continuous mapping $F \colon \mathrm{S}^2_+ V^* \to \mathscr{J}(V, \omega)$ satisfying $F(\mathrm{g}_J) = J$.*

Proof Since ω is non-degenerate, every $\mathrm{g} \in \mathrm{S}^2_+ V^*$ defines an endomorphism a of V by

$$\mathrm{g}(u, v) = \omega(u, av).$$

Since g is non-degenerate, a is invertible. Let a^* denote the adjoint of a relative to g, defined by $\mathrm{g}(u, a^* v) = \mathrm{g}(au, v)$ for all $u, v \in V$, and let $a = J|a|$ be the polar decomposition of a relative to g, that is, $|a| = \sqrt{a^* a}$ and $J = a|a|^{-1}$. The anti-symmetry of ω implies $a = -a^*$. It follows that $J = a|a|^{-1} = |a|^{-1} a$ and hence $J^2 = a^2(|a|^2)^{-1} = a^2(-a^2)^{-1} = -\mathrm{id}$. Consequently, J is a complex structure on (V, ω). Due to

$$\omega(u, Jv) = \omega\big(u, a|a|^{-1} v\big) = \mathrm{g}\big(u, |a|^{-1} v\big) = \mathrm{g}\big(|a|^{-\frac{1}{2}} u, |a|^{-\frac{1}{2}} v\big),$$

it is ω-compatible. Thus, $\mathscr{J}(V, \omega) \neq \varnothing$ and we can define a mapping

$$F \colon \mathrm{S}^2_+ V^* \to \mathscr{J}(V, \omega), \qquad F(\mathrm{g}) := J.$$

Since $J^2 = -\mathrm{id}$, we have $F(\mathrm{g}_J) = J$. In particular, F is surjective.

It remains to show that F is continuous. Since $J = a|a|^{-1}$ and since multiplication and inversion in $\mathrm{GL}(V)$, as well as the assignment $\mathrm{g} \mapsto a$, are obviously continuous, it suffices to show that for arbitrary $A \in \mathrm{End}(V)$, the assignment $\mathrm{g} \mapsto |A|$ is continuous. To see this, fix a reference scalar product $\mathrm{g}_0 \in \mathrm{S}^2_+ V^*$. By means of orthonormal bases, one with respect to g and another one with respect to g_0, one can construct $b \in \mathrm{GL}(V)$ such that $\mathrm{g}(u, v) = \mathrm{g}_0(bu, bv)$ for all $u, v \in V$. Consider the algebra homomorphism

$$\mathrm{C}_b : \mathrm{End}(V) \to \mathrm{End}(V), \qquad \mathrm{C}_b(A) := bAb^{-1}.$$

It satisfies $\mathrm{C}_b(A^*) = \mathrm{C}_b(A)^{*_0}$, where A^{*_0} denotes the adjoint of A with respect to g_0 (Exercise 7.5.3). This implies that C_b maps endomorphisms which are positive relative to g to endomorphisms which are positive relative to g_0. Thus, $\mathrm{C}_b(\sqrt{A^*A})$ is positive relative to g_0. Since it satisfies $\mathrm{C}_b(\sqrt{A^*A})^2 = \mathrm{C}_b(A)^{*_0}\mathrm{C}_b(A)$, it is the square root of $\mathrm{C}_b(A)^{*_0}\mathrm{C}_b(A)$. Hence,

$$|A| = \mathrm{C}_b^{-1}\left(\sqrt{\mathrm{C}_b(A)^{*_0}\mathrm{C}_b(A)}\right).$$

Since $\mathrm{C}_b(A)$ depends continuously on b and b can be chosen[3] to depend continuously on g, the assertion follows. □

Proposition 7.5.7 *The space $\mathscr{J}(V, \omega)$ is contractible.*

Proof By Corollary 7.1.5, we may assume that $(V, \omega) = (\mathbb{R}^{2n}, \omega_0)$ with $\omega_0(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot (J_n\mathbf{y})$. According to Example 7.5.5, $-J_n \in \mathscr{J}(\mathbb{R}^{2n}, \omega_0)$. We are going to construct a homotopy between the constant mapping $J \mapsto -J_n$ and the identical mapping of $\mathscr{J}(\mathbb{R}^{2n}, \omega_0)$. By ω_0-compatibility, for every $J \in \mathscr{J}(\mathbb{R}^{2n}, \omega_0)$, the matrix $a_J := J_n J$ is positive and symmetric. Hence, we can take a_J^s for every $s \in \mathbb{R}$ and define a mapping

$$H : [0, 1] \times \mathscr{J}\left(\mathbb{R}^{2n}, \omega_0\right) \to \mathrm{M}_{2n}(\mathbb{R}), \qquad H(s, J) := -J_n a_J^s.$$

Clearly, H is continuous and satisfies $H(0, J) = -J_n$ and $H(1, J) = J$. Thus, in order to prove that it yields the desired homotopy, we have to check that $H(s, J) \in \mathscr{J}(\mathbb{R}^{2n}, \omega_0)$ for all $0 < s < 1$. Compatibility with ω_0 follows from $\omega_0(\mathbf{x}, H(s, J)\mathbf{y}) = \mathbf{x} \cdot (a_J^s\mathbf{y})$ and the fact that a_J^s is positive symmetric for all s. Since $H(s, J)^2 = J_n(a_J^s)^{\mathrm{T}} J_n a_J^s$, to see that $H(s, J)^2 = -\mathbb{1}$, it suffices to show that a_J^s is symplectic. Since a_J is positive symmetric, it possesses a basis $\{\mathbf{e}_i\}$ of eigenvectors with eigenvalues $\alpha_i > 0$. Since $-J_n$ and J are ω_0-compatible, they are symplectic. Hence, so is a_J. Therefore,

$$\omega_0(\mathbf{e}_i, \mathbf{e}_j) = \omega_0(a_J\mathbf{e}_i, a_J\mathbf{e}_j) = \alpha_i\alpha_j\omega_0(\mathbf{e}_i, \mathbf{e}_j)$$

[3] By means of a local section of the submersion $\mathrm{GL}(V) \to \mathrm{GL}(V)/\mathrm{O}(V, \mathrm{g}_0)$ (right cosets).

for all i, j. It follows that $\alpha_i \alpha_j = 1$ or $\omega_0(\mathbf{e}_i, \mathbf{e}_j) = 0$. Then,

$$\omega_0\big(a_J^s \mathbf{e}_i, a_J^s \mathbf{e}_j\big) = (\alpha_i \alpha_j)^s \omega_0(\mathbf{e}_i, \mathbf{e}_j) = \omega_0(\mathbf{e}_i, \mathbf{e}_j).$$

Thus, a_J^s is symplectic for all s. This proves the proposition. □

Exercises

7.5.1 Prove Proposition 7.5.3.

7.5.2 Prove the statement of Remark 7.5.4/1.

7.5.3 Let V be a real vector space. Let $\mathrm{g}_1, \mathrm{g}_2 \in \mathrm{S}^2_+ V^*$ and let $b \in \mathrm{GL}(V)$ such that $\mathrm{g}_1(u, v) = \mathrm{g}_2(bu, bv)$ for all $u, v \in V$. Show that for $A \in \mathrm{End}(V)$, the adjoints A^{*_i} of A with respect to g_i are related by $b(A^{*_1})b^{-1} = (bAb^{-1})^{*_2}$.

7.6 The Lagrange-Graßmann Manifold

Let (V, ω) be a symplectic vector space of dimension $2n$. In this section, we study the structure of the space $\mathscr{L}(V, \omega)$ of Lagrangian subspaces of (V, ω). For that purpose, we choose an ω-compatible complex structure J on (V, ω). For simplicity we write $\mathrm{U}(V) = \mathrm{U}(V, \mathsf{h}_J)$ and $\mathrm{O}(V) = \mathrm{O}(V, \mathrm{g}_J)$. For $L \in \mathscr{L}(V, \omega)$, let $\mathrm{O}(L)$ denote the isometry group of L with respect to the scalar product on L induced by g_J. Proposition 7.5.3/3 implies that $\mathrm{O}(L)$ may be naturally viewed as a Lie subgroup of $\mathrm{U}(V)$ via the identification

$$\mathrm{O}(L) = \big\{a \in \mathrm{U}(V) : aL = L\big\}. \tag{7.6.1}$$

Let $\mathrm{G}(k, V)$ denote the Graßmann manifold of k-dimensional (real) linear subspaces of V, cf. Example 5.7.6.

Proposition 7.6.1 *$\mathscr{L}(V, \omega)$ is an embedded submanifold of $\mathrm{G}(n, V)$ of dimension $\frac{n(n+1)}{2}$. Every $L \in \mathscr{L}(V, \omega)$ defines a diffeomorphism*

$$\mathrm{U}(V)/\mathrm{O}(L) \to \mathscr{L}(V, \omega), \quad [a] \mapsto aL. \tag{7.6.2}$$

Proof Let $L \in \mathscr{L}(V, \omega)$ be given. It defines a diffeomorphism

$$\varphi : \mathrm{O}(V)/\big(\mathrm{O}(L) \times \mathrm{O}\big(L^\perp\big)\big) \to \mathrm{G}(n, V), \quad [a] \mapsto aL,$$

where $L^\perp$ denotes the g_J-orthogonal complement of L, which is Lagrange by Proposition 7.5.3/2. For every $a \in \mathrm{U}(V) \subset \mathrm{O}(V)$, the subspace aL is Lagrange, because, by (7.5.3), a is symplectic. Conversely, for every Lagrangian subspace L' we find $a \in \mathrm{U}(V)$ such that $L' = aL$: choose orthonormal bases in L and L'. According to Proposition 7.5.3/4, these bases induce orthonormal symplectic bases in V. There exists $a \in \mathrm{O}(V)$ transforming the latter bases into one another. Since these bases are symplectic, so is a. Hence, (7.5.3) implies $a \in \mathrm{U}(V)$. We conclude that aL is Lagrange iff $a \in U(V)$.

On the other hand, by (7.6.1), the stabilizer of L under $\mathrm{U}(V)$ is given by the Lie subgroup $\mathrm{O}(L) \subset \mathrm{U}(V)$. Thus, $\mathscr{L}(V,\omega)$ is the image of the homogeneous space $\mathrm{U}(V)/\mathrm{O}(L)$ under φ. According to Theorem 5.7.2, this endows $\mathscr{L}(V,\omega)$ with a natural smooth structure. Applying points 1 and 2 of Corollary 6.5.3 to the mapping

$$\mathrm{U}(V)/\mathrm{O}(L) \to \mathrm{O}(V)/\mathrm{O}(L) \to \mathrm{O}(V)/\big(\mathrm{O}(L) \times \mathrm{O}\big(L^{\perp}\big)\big),$$

we obtain that $\mathscr{L}(V,\omega)$ is an embedded submanifold of $\mathrm{G}(n,V)$. Finally, using $U(V) \cong U(n)$ and $\mathrm{O}(L) \cong \mathrm{O}(n)$, we find $\dim \mathscr{L}(V,\omega) = \frac{n(n+1)}{2}$. □

Remark 7.6.2

1. Under the identification (7.6.1), with respect to the decomposition $V = L \oplus L^{\perp}$, the subgroup $\mathrm{O}(L)$ consists of the block diagonal endomorphisms $a \oplus (-JaJ)$ with $a \in \mathrm{O}(L)$.
2. By Proposition 7.6.1 and (7.5.3), $\mathrm{Sp}(V,\omega)$ acts transitively on $\mathscr{L}(V,\omega)$.

Definition 7.6.3 $\mathscr{L}(V,\omega)$ is called the Lagrange-Graßmann manifold of (V,ω).

Example 7.6.4 Consider $V = \mathbb{R}^{2n}$ with the standard symplectic structure defined by J_n. Take $J = -J_n$ and $L = \mathbb{R}^n \times \{0\}$. Then, under the identification of $\mathbb{R}^{2n}$ with $\mathbb{C}^n$, cf. Example 7.5.5, $\mathrm{O}(L)$ corresponds to $\mathrm{O}(n) \subset \mathrm{U}(n)$.

Later on, the intersection properties of Lagrangian subspaces will be relevant. To study them, one exploits the partition of $\mathscr{L}(V,\omega)$ relative to a chosen Lagrangian subspace L, given by the subsets

$$\mathscr{L}_k(L) := \big\{L' \in \mathscr{L}(V,\omega) : \dim\big(L' \cap L\big) = k\big\}, \quad k = 0,\ldots,n.$$

Definition 7.6.5 The subset

$$\hat{\mathscr{L}}(L) = \mathscr{L}(V,\omega) \setminus \mathscr{L}_0(L) \equiv \bigcup_{k=1}^{n} \mathscr{L}_k(L) \tag{7.6.3}$$

of $\mathscr{L}(V,\omega)$ is called the Maslov cycle of L.

First, we study $\mathscr{L}_0(L)$. Let $L' \in \mathscr{L}_0(L)$ and let $P_{L'}: V \to L$ be the projection relative to the decomposition $V = L \oplus L'$. Then, $\mathbb{1} - P_{L'}$ yields a vector space isomorphism between the Lagrangian subspaces $L^{\perp} \subset V$ and L', because the kernel of $\mathbb{1} - P_{L'}$ is L. Moreover, by Proposition 7.5.3/2, we have $L^{\perp} = JL$. Thus, we obtain

$$L' = (\mathbb{1} - P_{L'})\big(L^{\perp}\big) = \big((\mathbb{1} - P_{L'})J\big)(L). \tag{7.6.4}$$

Next, we extend the endomorphism $P_{L'}J$ of L to a complex linear endomorphism $Q_{L'}$ of V, cf. Proposition 7.5.3/3:

$$Q_{L'}(u + Jv) := P_{L'}Ju + JP_{L'}Jv, \quad u,v \in L. \tag{7.6.5}$$

By construction,

$$Q_{L'}(L) = L, \qquad (J - Q_{L'})(L) = L', \qquad Q_{L'}J = JQ_{L'}, \tag{7.6.6}$$

and for any $u, v \in L$ we have

$$\begin{aligned} &\mathsf{h}_J(Q_{L'}u, v) - \mathsf{h}_J(u, Q_{L'}v) \\ &\quad = \omega(Q_{L'}u, Jv) - \omega(u, JQ_{L'}v) = -\omega\big((J - Q_{L'})u, (J - Q_{L'})v\big), \end{aligned} \tag{7.6.7}$$

because L and $L^{\perp}$ are Lagrange. Since $L' = (J - Q_{L'})(L)$ is also Lagrange, $Q_{L'}$ is Hermitian with respect to h_J. Using this, together with $Q_{L'}J = JQ_{L'}$, we obtain the polar decomposition of $J - Q_{L'}$ with respect to h_J:

$$J - Q_{L'} = \frac{J - Q_{L'}}{\sqrt{1 + Q_{L'}^2}} \sqrt{1 + Q_{L'}^2}.$$

Since $Q_{L'}(L) = L$, we have $\sqrt{1 + Q_{L'}^2}(L) = L$. To summarize, the polar decomposition yields a certain element a of $\mathrm{U}(V)$ such that

$$L' = aL, \qquad a = \frac{J - Q_{L'}}{\sqrt{1 + Q_{L'}^2}}. \tag{7.6.8}$$

Note that a is a square root of the Cayley transform of $Q_{L'}$ and that $Q_{L'}$ can be reconstructed from a via

$$Q_{L'} = J \frac{\mathbb{1} + a^2}{\mathbb{1} - a^2} \tag{7.6.9}$$

(Exercise 7.6.1).

Proposition 7.6.6 *Let* $L \in \mathscr{L}(V, \omega)$.

1. $\mathscr{L}_0(L)$ *is open and dense in* $\mathscr{L}(V, \omega)$.
2. *There exists a natural diffeomorphism*[4] $\mathscr{L}_0(L) \to \mathrm{S}^2 L^*$, $L' \mapsto S_{L'}$, *defined by*

$$S_{L'}(u, v) := \mathsf{g}_J(u, Q_{L'}v), \quad u, v \in L. \tag{7.6.10}$$

The kernel of $S_{L'}$ *is given by* $JL' \cap L$.

As a consequence of point 2, $\mathscr{L}_0(L)$ is contractible.

Proof 1. In view of Proposition 7.6.1 and the fact that the natural projection from $\mathrm{U}(V)$ to $\mathrm{U}(V)/\mathrm{O}(L)$ is a submersion and hence open, it suffices to show that the

[4] Recall that $\mathrm{S}^2 L^*$ denotes the vector space of symmetric bilinear forms on L.

subset of $\mathrm{U}(V)$ of elements a satisfying $aL \cap L = \{0\}$ is open and dense. The latter follows from the fact that the condition $aL \cap L = \{0\}$ is equivalent to the condition that, for some basis $\{e_i\}$ in L, the system $\{e_i\} \cup \{ae_i\}$ is linearly independent.

2. Since L is Lagrange, the restrictions of g_J and h_J to L coincide. Thus, symmetry of $S_{L'}$ follows from hermiticity of $Q_{L'}$. We prove bijectivity. To show injectivity, assume that there are two Lagrangian subspaces L' and L'' such that $S_{L'} = S_{L''}$. This means

$$\mathsf{g}_J\big(u, (Q_{L'} - Q_{L''})v\big) = 0$$

for all $u, v \in L$, that is, $Q_{L'}$ and $Q_{L''}$ coincide on L. Now, the second relation in (7.6.6) implies $L' = L''$. To prove surjectivity, let $S \in \mathrm{S}^2 L^*$. Via (7.6.10), to S there corresponds a g_J-symmetric endomorphism Q of L, which according to Proposition 7.5.3/3 extends uniquely to a Hermitian endomorphism of V. Consider the subspace $L' = (J - Q)(L)$. Since $J(L) = L^\perp$, we have $L' \cap L = \{0\}$ and since Q is Hermitian, (7.6.7) implies that L' is Lagrange. Hence, $L' \in \mathscr{L}_0(L)$. Using that Q leaves L invariant and that

$$(P_{L'})_{\restriction L} = \mathrm{id}_L, \qquad (P_{L'})_{\restriction L'} = 0,$$

we find

$$(Q_{L'} - Q)_{\restriction L} = P_{L'} \circ (J - Q)_{\restriction L} = 0.$$

This implies $S_{L'} = S$. Thus, the assignment $L' \mapsto S_{L'}$ is bijective, indeed. Differentiability in both directions follows from the facts that the natural projection $\mathrm{U}(V) \to \mathscr{L}(V, \omega)$ is a submersion and that the assignments $Q \mapsto a$ and $a \mapsto Q$ given by (7.6.8) and (7.6.9), respectively, are analytic. Finally, to determine the kernel of $S_{L'}$, let $u \in L$ such that $Q_{L'}(u) = P_{L'} J(u) = 0$. Then, $J(u) \in L'$ and hence $u \in JL' \cap L$. □

Remark 7.6.7

1. By definition, we have

$$S_{L'}(u, v) = \omega(u, JQ_{L'} v), \quad u, v \in L.$$

2. Let $\{e_i\}$ be an orthonormal basis in L and let A be the matrix of $-Q_{L'}$ in this basis. The relation $L' = (J - Q_{L'})(L)$ reads

$$L' = \left\{ \sum_{i=1}^{n} (q_i f_i + p_i e_i) : p_i = \sum_{j=1}^{n} A_{ij} q_j \right\}, \tag{7.6.11}$$

where $f_i = Je_i$, and the diffeomorphism $\mathscr{L}_0(L) \to \mathrm{S}^2 L^*$ yields a diffeomorphism

$$\varphi : \mathscr{L}_0(L) \to \mathrm{S}^2 \mathbb{R}^n, \qquad L' \mapsto \varphi\big(L'\big) = A. \tag{7.6.12}$$

With respect to the basis $\{e_i\}$ of the complex vector space V, the element $a \in \mathrm{U}(V)$ transforming L to L', found in (7.6.8), is represented by the matrix

$$a = \frac{\mathrm{i} + A}{\sqrt{\mathbb{1} + A^2}}. \tag{7.6.13}$$

A can be reconstructed from a via

$$A = -\mathrm{i}\frac{\mathbb{1} + a^2}{\mathbb{1} - a^2} \tag{7.6.14}$$

(Exercise 7.6.1).

Next, we study the Maslov cycle $\hat{\mathscr{L}}(L)$ of L. For that purpose, we choose an orthonormal basis $\{e_i\}$ in L and take the diffeomorphism (7.6.12) as a local chart on the open subset $\mathscr{L}_0(L)$ of $\mathscr{L}(V,\omega)$. To construct local charts covering the remaining part of $\mathscr{L}(V,\omega)$, that is, the Maslov cycle $\hat{\mathscr{L}}(L)$, for every subset $K \subset \{1,\dots,n\}$, we define

(a) L_K to be the real subspace of V spanned by all f_i with $i \in K$ and all e_j with $j \notin K$,
(b) a_K to be the linear mapping $V \to V$ defined by

$$a_K e_i = \begin{cases} f_i & i \in K, \\ e_i & i \notin K, \end{cases} \qquad a_K f_i = \begin{cases} -e_i & i \in K, \\ f_i & i \notin K. \end{cases} \tag{7.6.15}$$

Obviously, $L_\emptyset = L$. One has $L_K = a_K L$ and $a_K \in \mathrm{U}(V)$, because it is both orthogonal and symplectic. It follows that L_K is Lagrange and that $\mathscr{L}_0(L_K) = a_K \mathscr{L}_0(L)$. Hence, for every K we can define a local chart φ_K on $\mathscr{L}_0(L_K)$ by transporting φ by the help of a_K:

$$\varphi_K : \mathscr{L}_0(L_K) \to \mathrm{S}^2\mathbb{R}^n, \qquad \varphi_K\big(L'\big) := \varphi\big(a_K^{-1} L'\big). \tag{7.6.16}$$

Proposition 7.6.8 *For every $L' \in \mathscr{L}_k(L)$, there exists a subset K of $\{1,\dots,n\}$ with k elements such that $L' \in \mathscr{L}_0(L_K)$. In particular, the family*

$$\big\{\big(\mathscr{L}_0(L_K), \varphi_K\big) : K \subset \{1,\dots,n\}\big\}$$

yields an atlas on $\mathscr{L}(V,\omega)$.

Proof Let $W_1 := L' \cap L$. If $\dim W_1 \equiv k < n$, there exists $i_1 \in \{1,\dots,n\}$ such that $e_{i_1} \notin W_1$. Now, let W_2 be the subspace of L spanned by W_1 and e_{i_1}. Iterating the argument, after $n-k$ steps we arrive at a sequence $i_1,\dots,i_{n-k}$ with the property that the subspace of L spanned by $e_{i_1},\dots,e_{i_{n-k}}$ is complementary to $L' \cap L$. Let

$$K := \{1,\dots,n\} \setminus \{i_1,\dots,i_{n-k}\}.$$

Then, $L \cap L_K$ is spanned by $e_{i_1}, \dots, e_{i_{n-k}}$, so that $L = (L \cap L') \oplus (L \cap L_K)$. Since both L' and L_K are Lagrange, this implies $\omega(L' \cap L_K, L) = 0$ and hence, since L is Lagrange, $L' \cap L_K \subset L$. It follows that

$$L' \cap L_K = L' \cap L_K \cap L = \big(L \cap L'\big) \cap (L \cap L_K) = \{0\}$$

and hence $L' \in \mathscr{L}_0(L_K)$. It is obvious that the family $\{\mathscr{L}_0(L_K) : K \subset \{1, \dots, n\}\}$ defines a covering of $\mathscr{L}(V, \omega)$. The proof of the compatibility of the chart mappings with the smooth structure provided by Proposition 7.6.1 is left to the reader. □

Remark 7.6.9 Let $L' \in \mathscr{L}_0(L_K)$ and denote $A = \varphi_K(L')$. Then, $\varphi^{-1}(A) = a_K^{-1} L'$ and thus $L' = a_K a L$, where the matrix of a in the basis $\{e_i\}$ is given by (7.6.13). Moreover, since A is the symmetric matrix assigned to $a_K^{-1} L'$ by φ, (7.6.11) yields

$$a_K^{-1} L' = \left\{ \sum_{i=1}^n (q_i f_i + p_i e_i) : p_i = \sum_{j=1}^n A_{ij} q_j \right\}. \tag{7.6.17}$$

Using (7.6.15), we conclude that L' consists of the vectors $\sum_{i=1}^n (q_i f_i + p_i e_i)$ which satisfy the n equations

$$q_k = \sum_{j \notin K} A_{kj} q_j - \sum_{l \in K} A_{kl} p_l, \quad k \in K, \tag{7.6.18}$$

$$p_i = \sum_{j \notin K} A_{ij} q_j - \sum_{l \in K} A_{il} p_l, \quad i \notin K. \tag{7.6.19}$$

Proposition 7.6.10 *For every $k = 1, \dots, n$, the local charts $(\mathscr{L}_0(L_K), \varphi_K)$ with K consisting of k elements cover $\mathscr{L}_k(L)$. For every such K, the subset*

$$\mathscr{L}_k(L) \cap \mathscr{L}_0(L_K)$$

is mapped under φ_K onto the subspace of $\mathrm{S}^2\mathbb{R}^n$ defined by the $\frac{k(k+1)}{2}$ relations

$$A_{ij} = 0, \quad i, j \in K.$$

Proof The first assertion is a direct consequence of Proposition 7.6.8. To prove the second assertion, let K be given, let $L' \in \mathscr{L}_k(L) \cap \mathscr{L}_0(L_K)$ and let $A = \varphi_K(L')$. Then, $a_K^{-1} L'$ is given by (7.6.17). By definition of a_K, we have $a_K^{-1} L' = a_K L'$ and thus

$$a_K^{-1} L' \cap L_K = a_K L' \cap L_K = a_K \big(L' \cap L\big).$$

Hence the subspace $a_K^{-1} L' \cap L_K$ has dimension k. Writing

$$L_K = \left\{ \sum_{i=1}^n (q_i f_i + p_i e_i) : q_i = 0 \text{ for all } i \notin K,\ p_j = 0 \text{ for all } j \in K \right\}$$

we see that in the intersection $a_K^{-1}L' \cap L_K$, p_i is determined by q_i for all $i = 1, \ldots, n$ and q_i can only be nonzero for $i \in K$. Since $a_K^{-1}L' \cap L_K$ has dimension k, the equations

$$0 = p_i = \sum_{j=1}^{n} A_{ij} q_j \equiv \sum_{j \in K} A_{ij} q_j, \quad i \in K,$$

must, therefore, be satisfied for all $q_i \in \mathbb{R}$. Hence, $A_{ij} = 0$ for all $i, j \in K$. □

Combining Proposition 7.6.10 with Proposition 1.7.3, we obtain

Corollary 7.6.11 *For every $k = 0, \ldots, n$, the subset $\mathscr{L}_k(L)$ of $\mathscr{L}(V, \omega)$ is an embedded submanifold of codimension $\frac{k(k+1)}{2}$.*

Remark 7.6.12 Corollary 7.6.11 implies that the Maslov cycle $\hat{\mathscr{L}}(L)$ is a stratified subset of $\mathscr{L}(V, \omega)$, with the stratum $\mathscr{L}_1(L)$ having codimension 1 in $\mathscr{L}(V, \omega)$ and the other strata having codimension at least 3. This implies that $\mathscr{L}_1(L)$ is open and dense in $\hat{\mathscr{L}}(L)$.

Exercises

7.6.1 Prove the reconstruction formulae (7.6.9) and (7.6.14).

7.7 The Universal Maslov Class

In this section, we use some elementary facts from topology, see e.g. [55], [124] or part III of [76]. Our presentation follows the line of reasoning of a classical paper by Arnold [13], see also [79]. As in Sect. 7.6, let there be chosen an ω-compatible complex structure J. Let us choose $L \in \mathscr{L}(V, \omega)$ and let us consider the mapping

$$\det\colon U(V) \to \mathrm{S}^1 \subset \mathbb{C}, \quad a \mapsto \det a.$$

Since on the subgroup $\mathrm{O}(L)$ it takes the values ± 1, via the isomorphism (7.6.2) defined by L, it induces a well-defined mapping

$$\det{}_L^2 : \mathscr{L}(V, \omega) \to \mathrm{S}^1 \subset \mathbb{C}. \tag{7.7.1}$$

For $a \in \mathrm{U}(V)$, the mappings $\det_L^2$ and $\det_{aL}^2$ differ by a constant phase factor:

$$\det{}_{aL}^2 = \det{}^{-2}(a) \det{}_L^2. \tag{7.7.2}$$

Definition 7.7.1 (Maslov index) The Maslov index $\mu(\gamma)$ of a closed curve $\gamma : \mathrm{S}^1 = \mathbb{R}/2\pi\mathbb{Z} \to \mathscr{L}(V, \omega)$ is defined to be the degree of the mapping[5] $\det_L^2 \circ \gamma\colon \mathrm{S}^1 \to \mathrm{S}^1$.

[5]Cf. Remark 4.3.6/4. Both $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ and $S^1 \subset \mathbb{C}$ are endowed with the natural orientations.

Proposition 7.7.2 *The Maslov index is homotopy invariant, that is, it is constant on homotopy classes of closed curves. It does not depend on the choice of L or J.*

Proof Homotopy invariance follows from the continuity of the mapping (7.7.1) and from the fact that the mapping degree is homotopy invariant. By (7.7.2), μ does not depend on L. Using this, as well as homotopy invariance of μ and point 2 of Remark 7.5.4, one obtains independence of J. □

According to Proposition 7.7.2, the Maslov index defines a homomorphism

$$\mu\colon \pi_1\big(\mathscr{L}(V,\omega)\big) \to \mathbb{Z}, \tag{7.7.3}$$

denoted by the same symbol.

Proposition 7.7.3 *The mapping (7.7.3) is an isomorphism.*

Accordingly, $\pi_1(\mathscr{L}(V,\omega))$ is isomorphic to $\mathbb{Z}$ and μ is a generator of the Abelian group $\mathrm{Hom}(\pi_1(\mathscr{L}(V,\omega)),\mathbb{Z}) \cong \mathbb{Z}$.

Proof Let $L \in \mathscr{L}(V,\omega)$. We choose an orthonormal basis in L and use the corresponding matrix representation to identify $\mathrm{U}(V)$ with $\mathrm{U}(n)$. Then, $\mathrm{O}(L) = \{a \in \mathrm{U}(V) : aL = L\}$ is given by the subgroup $\mathrm{O}(n) \subset \mathrm{U}(n)$. We have to show that the homomorphism $\pi_1(\mathrm{U}(n)/\mathrm{O}(n)) \to \pi_1(\mathrm{S}^1) = \mathbb{Z}$ induced by $\gamma \mapsto \det_L^2 \circ\gamma$, where γ is a closed curve in $\mathrm{U}(n)$, is bijective. Thus, for every $k \in \mathbb{Z}$, define a curve

$$\gamma_k : [0,1] \to \mathrm{U}(n), \qquad \gamma_k(t) := \mathrm{diag}\big(\mathrm{e}^{\mathrm{i}k\pi t}, 1, \dots, 1\big).$$

Since $\gamma_k(0)$ and $\gamma_k(1)$ lie in $\mathrm{O}(n)$, the curves γ_k project to closed curves $\hat{\gamma}_k$ in $\mathrm{U}(n)/\mathrm{O}(n)$, starting and ending at $[\mathbb{1}]$. Note that $\det_L^2 \circ\gamma_k$ winds k times around S^1. This proves surjectivity. To prove that it is injective, we show that every closed curve $\hat{\gamma}$ in $\mathrm{U}(n)/\mathrm{O}(n)$, starting and ending at $[\mathbb{1}]$, is homotopic to one of the curves $\hat{\gamma}_k$. Since the projection $\mathrm{U}(n) \to \mathrm{U}(n)/\mathrm{O}(n)$ admits local sections, $\hat{\gamma}$ has a lift $\gamma : [0,1] \to \mathrm{U}(n)$ starting at $\mathbb{1}$ and ending at some $a \in \mathrm{O}(n)$. Depending on whether $\det(\gamma(1))$ equals $+1$ or -1, by composing γ with a curve in $\mathrm{O}(n)$, we can make it end at $\mathbb{1}$ or at $(-1, 1, \dots, 1)$. In the first case, γ is a closed curve in $\mathrm{U}(n)$. It is therefore homotopic to γ_k for some even k, see Exercise 7.7.1. Then, $\hat{\gamma}$ is homotopic to $\hat{\gamma}_k$. In the second case, we use that $\gamma \cdot \gamma_{-1}$ (pointwise multiplication in $\mathrm{U}(n)$) is closed in $\mathrm{U}(n)$ and hence is homotopic to γ_k for some even k. Then, γ is homotopic to $\gamma_k \cdot \gamma_1 = \gamma_{k+1}$ and hence $\hat{\gamma}$ is homotopic to $\hat{\gamma}_{k+1}$. □

Now, let ϕ denote the standard angle coordinate on $\mathrm{S}^1 \subset \mathbb{C}$ and consider the 1-form

$$\mu := \frac{1}{2\pi}\big(\det_L^2\big)^* \mathrm{d}\phi \tag{7.7.4}$$

on $\mathscr{L}(V,\omega)$. By (7.7.2), for $a \in \mathrm{U}(V)$, we have

$$\big(\det^2_{aL}\big)^*\mathrm{d}\phi = \mathrm{d}\big(\phi \circ \det^2_{aL}\big) = \mathrm{d}\big(\phi \circ \det^{-2}(a) + \phi \circ \det^2_L\big) = \big(\det^2_L\big)^*\mathrm{d}\phi,$$

because $\phi \circ \det^{-2}(a)$ is a constant. Hence, the 1-form μ does not depend on the choice of L. The following observation justifies the notation μ.

Proposition 7.7.4 *For every closed curve* $\gamma : \mathrm{S}^1 \to \mathscr{L}(V,\omega)$,

$$\mu(\gamma) = \int_\gamma \mu.$$

Proof By definition of the integral over a submanifold,

$$\int_\gamma \mu = \int_{\mathrm{S}^1} \gamma^*\mu = \frac{1}{2\pi}\int_{\mathrm{S}^1} \big(\det^2_L \circ \gamma\big)^*\,\mathrm{d}\phi.$$

By definition of the mapping degree, cf. Remark 4.3.6/4, this equals

$$\frac{1}{2\pi}\deg\big(\det^2_L \circ \gamma\big)\int_{\mathrm{S}^1} \mathrm{d}\phi = \mu(\gamma).$$

□

Definition 7.7.5 (Universal Maslov class) The 1-form (7.7.4) is called the universal Maslov class of $\mathscr{L}(V,\omega)$.

Remark 7.7.6

1. Integration of 1-forms over closed curves yields a homomorphism from the de-Rham cohomology group $H^1(\mathscr{L}(V,\omega))$ to $\mathrm{Hom}(\pi_1(\mathscr{L}(V,\omega)),\mathbb{R})$. Using methods of algebraic topology, one can show that this homomorphism is in fact an isomorphism.[6] Since $\pi_1(\mathscr{L}(V,\omega)) \cong \mathbb{Z}$, we have $\mathrm{Hom}(\pi_1(\mathscr{L}(V,\omega)),\mathbb{R}) \cong \mathbb{R}$. Hence, Proposition 7.7.4 states that under the above isomorphism, the 1-form $\mu \in H^1(\mathscr{L}(V,\omega))$ defined by (7.7.4) corresponds to the homomorphism $\mu \in \mathrm{Hom}(\pi_1(\mathscr{L}(V,\omega)),\mathbb{Z}) \subset \mathrm{Hom}(\pi_1(\mathscr{L}(V,\omega)),\mathbb{R})$ defined by (7.7.3). In particular, the first one spans $H^1(\mathscr{L}(V,\omega))$. This explains the name universal class.
2. We express the mapping (7.7.1) in terms of the local charts φ_K on $\mathscr{L}(V,\omega)$, induced by an orthonormal basis $\{e_i\}$ in L, cf. Proposition 7.6.8. Let $L' \in \mathscr{L}_0(L_K)$ and let $A = \varphi_K(L')$. Due to Remark 7.6.9, $L' = a_K a L$, where the matrix of a

[6] It decomposes into the isomorphisms

$$H^1\big(\mathscr{L}(V,\omega)\big) \to H^1\big(\mathscr{L}(V,\omega),\mathbb{R}\big) \to \mathrm{Hom}\big(H_1\big(\mathscr{L}(V,\omega)\big),\mathbb{R}\big) \to \mathrm{Hom}\big(\pi_1\big(\mathscr{L}(V,\omega)\big),\mathbb{R}\big)$$

provided by, respectively, the de Rham Theorem, the Universal Coefficient Theorem and the Hurewicz Theorem. Here, $H^1(\cdot,\mathbb{R})$ and $H_1(\cdot)$ denote the first integer-valued singular cohomology group and the first singular homology group, respectively.

with respect to the basis $\{e_i\}$ in the complex vector space V is given by $\frac{\mathrm{i}+A}{\sqrt{\mathbb{1}+A^2}}$. Using that $\det a_K = 1$, cf. Proposition 7.4.1/4, we obtain

$$\det{}_L^2(L') = \det(a^2) = \det\left(\frac{A+\mathrm{i}}{A-\mathrm{i}}\right).$$

Next, we show that the Maslov index can be viewed as an intersection index. It was this way this index occurred first in the work of Maslov [197]. For that purpose, we consider the action of the center $\mathrm{U}(1) \subset \mathrm{U}(V)$ on $\mathscr{L}(V,\omega)$, given by

$$\mathrm{U}(1) \times \mathscr{L}(V,\omega) \to \mathscr{L}(V,\omega), \qquad (\mathrm{e}^{\mathrm{i}t}, L) \mapsto \mathrm{e}^{\mathrm{i}t} L \in \mathscr{L}(V,\omega).$$

This action is generated by the vector field[7]

$$(J_*)_L = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction_0} (\mathrm{e}^{\mathrm{i}t} L).$$

Proposition 7.7.7 *The vector field J_* is transversal to $\mathscr{L}_1(L)$ and thus defines a coorientation*[8] *on $\mathscr{L}_1(L)$ for all $L \in \mathscr{L}(V,\omega)$. This coorientation does not depend on J.*

Proof Let $L \in \mathscr{L}(V,\omega)$ and consider the local charts $(\mathscr{L}_0(L_K), \varphi_K)$ defined by an orthonormal basis in L. According to Proposition 7.6.10, $\mathscr{L}_1(L)$ is covered by the charts $(\mathscr{L}_0(L_K), \varphi_K)$ with $K = \{m\}$, $m = 1, \ldots, n$, and the subset $\mathscr{L}_1(L) \cap \mathscr{L}_0(L_K)$ is mapped under φ_K to the subspace of $\mathrm{S}^2\mathbb{R}^n$ defined by the single equation $A_{mm} = 0$. Let $L' \in \mathscr{L}_1(L) \cap \mathscr{L}_0(L_K)$ and let $A = \varphi_K(L')$. According to Remark 7.6.9, we have $L' = a_K a L$, where the matrix of a with respect to the basis $\{e_i\}$ is given by (7.6.13). Hence, the local representative of $(J_*)_{L'}$ in the chart φ_K is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction_0} \varphi_K(\mathrm{e}^{\mathrm{i}t} L') = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction_0} \varphi_K(\mathrm{e}^{\mathrm{i}t} a_K a L) = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction_0} \varphi(\mathrm{e}^{\mathrm{i}t} a L)$$

see (7.6.16). For sufficiently small t, the matrix of $\mathrm{e}^{2it} a^2$ does not have an eigenvalue 1, hence we can use (7.6.14) to calculate

$$\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction_0} \varphi(\mathrm{e}^{\mathrm{i}t} a L) = -\mathrm{i} \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction_0} \left(\frac{\mathbb{1} + \mathrm{e}^{2\mathrm{i}t} a^2}{\mathbb{1} - \mathrm{e}^{2\mathrm{i}t} a^2}\right) = 4a^2(\mathbb{1} - a^2)^{-2} = -(\mathbb{1} + A^2).$$

Thus, in the chart φ_K, $(J_*)_{L'}$ is represented by $-(\mathbb{1} + A^2)$. Since A is symmetric, A^2 has nonnegative diagonal entries. This proves that J_* is transversal to $\mathscr{L}_1(L)$. That the orientation of the normal bundle so defined does not depend on J follows

[7] This is the Killing vector field generated by J under the action of $\mathrm{Sp}(V,\omega)$ on $\mathscr{L}(V,\omega)$.

[8] That is, an orientation of the normal bundle of the submanifold $\mathscr{L}_1(L)$, cf. Remark 2.7.18/2.

from the fact that, according to Proposition 7.5.7, the subset $\mathscr{J}(V,\omega)$ of $\mathfrak{sp}(V,\omega)$ is arcwise connected. Indeed, since the mapping

$$\mathfrak{sp}(V,\omega) \to \mathrm{T}_{L'}\mathscr{L}(V,\omega) \to \mathrm{N}_{L'}\mathscr{L}_1(L) \cong \mathbb{R},$$

defined by the Killing vector fields of the action of $\mathrm{Sp}(V,\omega)$ on $\mathscr{L}(V,\omega)$, is smooth and since the image of an element of $\mathscr{J}(V,\omega)$ under this mapping is nonzero, the orientation cannot change along a curve in $\mathscr{J}(V,\omega)$. □

Remark 7.7.8 As a corollary of the proof, we note that the local representative of the Killing vector field J_* with respect to the local chart $\varphi_{\{m\}}$ points from the side where A_{mm} is positive to the side where A_{mm} is negative.

Now let $L \in \mathscr{L}(V,\omega)$, let $a,b \in \mathbb{R}$ and let $\gamma : [a,b] \to \mathscr{L}(V,\omega)$ be a curve with endpoints in $\mathscr{L}_0(L)$. A number $t \in (a,b)$ such that $\gamma(t) \in \hat{\mathscr{L}}(L)$ is called a crossing of γ with $\hat{\mathscr{L}}(L)$. A crossing t is said to be simple if $\gamma(t) \in \mathscr{L}_1(L)$. A simple crossing is said to be transversal if $\dot\gamma(t) \notin \mathrm{T}_{\gamma(t)}\mathscr{L}_1(L)$. Depending on whether $\dot\gamma(t)$ is positively or negatively oriented with respect to the canonical coorientation of $\mathscr{L}_1(L)$, a simple transversal crossing is said to be positive or negative. That is, a simple transversal crossing is positive iff $\dot\gamma(t)$ and $(J_*)_{\gamma(t)}$ point to the same side of $\mathscr{L}_1(L)$. In the following, when we say that two curves with the same end points are homotopic with fixed end points we always mean that there is a homotopy preserving the endpoints.

Proposition 7.7.9

1. *Every curve in $\mathscr{L}(V,\omega)$ with end points in $\mathscr{L}_0(L)$ is homotopic with fixed end points to a curve which has only simple transversal crossings with $\hat{\mathscr{L}}(L)$.*
2. *If two curves in $\mathscr{L}(V,\omega)$ which have the same end points in $\mathscr{L}_0(L)$ and which have only simple and transversal crossings with $\hat{\mathscr{L}}(L)$ are homotopic with fixed end points, their differences between the numbers of positive and negative crossings coincide.*

Proof Below, by a homotopy we mean a homotopy with fixed end points. Denote $m = \frac{n(n+1)}{2}$.

1. Let $\gamma : [a,b] \to \mathscr{L}(V,\omega)$ with $\gamma(a),\gamma(b) \in \mathscr{L}_0(L)$ be given. Since the complement $\hat{\mathscr{L}}(L) \setminus \mathscr{L}_1(L)$ is the closure of the embedded submanifold $\mathscr{L}_2(L)$ which has codimension 3, γ is obviously homotopic to a curve, denoted by the same symbol, which has the same endpoints but only simple crossings, that is, it is contained in $\mathscr{L}_0(L) \cup \mathscr{L}_1(L)$. Since $\mathscr{L}_1(L)$ is an embedded submanifold and since $[a,b]$ is compact, there exist real numbers $a = t_0 < t_1 < \cdots < t_{r-1} < t_r = b$ and local charts (U_i,κ_i), $i = 1,\ldots,r$, on $\mathscr{L}(V,\omega)$ mapping $\mathscr{L}_1(L)$ to open subsets of hyperplanes in $\mathbb{R}^m$ such that

$$\gamma\big([t_{i-1},t_i]\big) \subset U_i, \qquad \gamma\big((t_i - 2\varepsilon, t_i + 2\varepsilon)\big) \subset U_{i-1} \cap U_i, \quad i = 1,\ldots,r.$$

Using κ_1, we can construct a smooth homotopy H_1 from γ to a curve γ_1 with $\gamma_1(t_1) \in \mathscr{L}_0(L)$, which has only transversal simple crossings between a and t_1 and which coincides with γ for $t \geq t_1 + \varepsilon$. Next, using κ_2, in a similar way we can construct a smooth homotopy H_2 from γ_1 to a curve γ_2 with $\gamma_2(t_1), \gamma_2(t_2) \in \mathscr{L}_0(L)$, which has only transversal simple crossings between $t_1 - \varepsilon$ and t_2 and which coincides with γ_1 outside $(t_1 - \varepsilon, t_2 + \varepsilon)$. We iterate this procedure until we arrive at a smooth homotopy H_r from γ_{r-1} to a curve γ_r with $\gamma_r(t_{r-1}) \in \mathscr{L}_0(L)$, which has only transversal simple crossings between $t_{r-1} - \varepsilon$ and b, and which coincides with γ_{r-1} for $t \leq t_{r-1} - \varepsilon$. Then, γ_r is homotopic to γ and has only transversal simple crossings.

2. Denote the homotopy by $H : [a, b] \times [0, 1] \to \mathscr{L}(V, \omega)$. For $s \in [0, 1]$, we define $t \mapsto \gamma_s(t) := H(t, s)$. By the codimension argument of point 1, H can be chosen so that it stays in $\mathscr{L}_0(L) \cup \mathscr{L}_1(L)$. Choose an orthonormal basis in L and consider the corresponding local charts $(\mathscr{L}_0(L_K), \varphi_K)$. Due to Proposition 7.6.10, $\mathscr{L}_0(L) \cup \mathscr{L}_1(L)$ is covered by the charts φ_K with K being empty or having a single element. Using this fact and compactness of $[a, b] \times [0, 1]$, we find $a = t_0 < t_1 < \cdots < t_r = b$ and $0 = s_0 < s_1 < \cdots < s_p = 1$ such that each of the squares $[t_i, t_{i+1}] \times [s_j, s_{j+1}]$ is mapped under H to the domain of a single chart $\varphi_{K_{ij}}$. We may furthermore assume that the crossings of γ_0 and γ_1 are distinct from $t_1, \dots, t_{r-1}$. For each $j = 1, \dots, p - 1$, by applying the procedure of point 1, the family of charts $\{\varphi_{K_{0j}}, \dots, \varphi_{K_{r-1j}}\}$ can be used to construct a homotopy from γ_{s_j} to some curve, still covered by these charts, whose crossings are simple and transversal and are distinct from $t_1, \dots, t_{r-1}$. Thus, by plugging in such a homotopy forth and back at $s = s_j$ if necessary, we may assume that the crossings of γ_{s_j} are simple transversal and distinct from $t_1, \dots, t_{r-1}$. This shows that it suffices to prove the assertion under the assumption that the homotopy H stays in the domain of a single chart φ_K with K having at most one element. Now, if H stays in $\mathscr{L}_0(L)$, the assertion obviously holds. If not, then K has one element and φ_K maps the relevant subset of $\mathscr{L}_1(L)$ onto a whole subspace of $\mathbb{R}^m$ of codimension 1. Here, the assertion is obvious as well. □

As a consequence of Proposition 7.7.9, we can define

Definition 7.7.10 (Maslov intersection index) The Maslov intersection index relative to $L \in \mathscr{L}(V, \omega)$ of a curve γ in $\mathscr{L}(V, \omega)$ with end points in $\mathscr{L}_0(L)$ is defined as

$$\mathrm{Ind}_L(\gamma) := \nu_+ - \nu_-,$$

where ν_+ is the number of positive crossings and ν_- is the number of negative crossings with $\hat{\mathscr{L}}(L)$ of a curve which is homotopic with fixed end points to γ and whose crossings are all simple and transversal.

According to Proposition 7.7.9, the Maslov intersection index is invariant under homotopies with fixed end points. Moreover, by construction, it is additive with respect to the composition of curves. We now show that, for closed curves, it coincides with the Maslov index defined before.

Theorem 7.7.11 *Let $\gamma : [0,1] \to \mathscr{L}(V,\omega)$ be a closed curve and let $L \in \mathscr{L}(V,\omega)$ be transversal to the Lagrangian subspace $\gamma(0) = \gamma(1)$. Then,* $\mathrm{Ind}_L(\gamma) = \mu(\gamma)$.

Proof Let us take $L' := \gamma(0) = \gamma(1)$ as the base point of $\pi_1(\mathscr{L}(V,\omega))$. By additivity, Ind_L defines a homomorphism $\pi_1(\mathscr{L}(V,\omega)) \to \mathbb{Z}$. Since, by Proposition 7.7.2, $\pi_1(\mathscr{L}(V,\omega)) \cong \mathbb{Z}$ and since two homomorphisms $\mathbb{Z} \to \mathbb{Z}$ coincide if there exists $k \in \mathbb{Z}$ where they take the same nonzero value, it suffices to find a closed curve $\gamma_1 : [0,1] \to \mathscr{L}(V,\omega)$ with $\mathrm{Ind}_L(\gamma_1) = \mu(\gamma_1) \neq 0$. Such a curve will be constructed now. Since L and L' are complementary, according to Remark 7.5.4/1, there exists $J \in \mathscr{J}(V,\omega)$ such that $L' = JL$. Choose a g_J-orthonormal basis $\{e_i\}$ in L and let $f_i := Je_i$. Define the curve γ_1 by requiring $\gamma_1(t)$ to be the real subspace of V spanned by $\mathrm{e}^{\pi \mathrm{i} t} f_1 = \cos(\pi t) e_1 - \sin(\pi t) f_1$ and $f_2, \dots, f_n$. Obviously, $\gamma_1(1) = \gamma_1(0) = L'$. Under the identification of $\mathscr{L}(V,\omega)$ with $\mathrm{U}(n)/\mathrm{O}(n)$ induced by L', J and the orthonormal basis $\{f_i\}$ in L', this curve corresponds to the curve γ_k with $k = 1$ used in the proof of Proposition 7.7.3. Either by this observation or by direct inspection, we find

$$\det{}^2\big(\gamma_1(t)\big) = \det{}^2 \mathrm{diag}\big(\mathrm{e}^{\pi \mathrm{i} t}, 1, \dots, 1\big) = \mathrm{e}^{2\pi \mathrm{i} t},$$

hence $\mu(\gamma_1) = 1$. To determine $\mathrm{Ind}_L(\gamma_1)$, we observe that $\gamma_1(t) \cap L = \{0\}$ unless $s = \frac{1}{2}$ and that $\gamma_1(\frac{1}{2}) \cap L$ is spanned over $\mathbb{R}$ by e_1. Hence, γ_1 has a single crossing with $\hat{\mathscr{L}}(L)$ and this crossing is simple and transversal. To find out whether it is positive or negative, we use the local charts $(\mathscr{L}_0(L_K), \varphi_K)$ defined by the basis $\{e_i\}$ in L. For $t \in (0,1)$, $\gamma_1(t) \in \mathscr{L}_0(L_K)$ with $K = \{1\}$, because $L_{\{1\}}$ is spanned by $f_1, e_2, \dots, e_n$. Since $a_{\{1\}}^{-1}(\gamma_1(t))$ is spanned by $f_2, \dots, f_n$ and the vector

$$a_{\{1\}}^{-1}\big(\mathrm{e}^{\pi \mathrm{i} t} f_1\big) = \mathrm{e}^{\pi \mathrm{i} t} e_1 = \cos(\pi t) e_1 + \sin(\pi t) f_1,$$

from (7.6.17) we read off $\varphi_{\{1\}}(\gamma_1(t)) = \mathrm{diag}(\cot(\pi t), 0, \dots, 0)$. Since $\cot(\pi t)$ is positive for $t < \frac{1}{2}$ and negative for $t > \frac{1}{2}$, according to Remark 7.7.8, the crossing is positive. This proves the assertion. □

Corollary 7.7.12 *If two curves $\gamma_0, \gamma_1 : [0,1] \to \mathscr{L}(V,\omega)$ with the same end points satisfy* $\mathrm{Ind}_L(\gamma_1) = \mathrm{Ind}_L(\gamma_2)$ *for some $L \in \mathscr{L}(V,\omega)$ which is transversal to both end points, they are homotopic with fixed end points.*

Proof By additivity, the closed curve $\gamma_1^{-1} \cdot \gamma_0$ obtained by composing γ_0 with γ_1 running backwards has Maslov intersection index zero. By Proposition 7.7.3, it is, therefore, homotopic to the constant curve $t \mapsto \gamma_0(0) = \gamma_1(0)$. We can certainly find a homotopy $H : [0,1] \times [0,1] \to \mathscr{L}(V,\omega)$ such that $H(1,t) = \gamma_0(2t)$ for $t \leq \frac{1}{2}$ and $H(1,t) = \gamma_1(2-2t)$ for $t \geq \frac{1}{2}$. It is enough to show that γ_0 and γ_1 are homotopic with fixed endpoints to the curve $\gamma_{\frac{1}{2}} : [0,1] \to \mathscr{L}(V,\omega)$ defined by

$\gamma_{\frac{1}{2}}(t) := H(t, \frac{1}{2}t)$. For γ_0, a homotopy is given by

$$\tilde{H} : [0,1] \times [0,1] \to \mathscr{L}(V,\omega), \qquad \tilde{H}(s,t) := \begin{cases} H(s, \frac{1}{2}t) & t \leq s \\ H(t, \frac{1}{2}t) & t > s. \end{cases}$$

For γ_1, there is an analogous formula. □

Remark 7.7.13 In [252], Robbin and Salamon have shown that the definition of the Maslov index relative to a chosen Lagrangian subspace for non-closed curves can be extended in a natural way to the case of arbitrary positions of the end points. We add a few comments on that important paper. In the next section, we will present a systematic treatment of this situation in terms of the Kashiwara index. The basic tool of the analysis of Robbin and Salamon is the following quadratic form associated with a tangent vector of $\mathscr{L}(V,\omega)$. Let $L' \in \mathscr{L}(V,\omega)$ and let $X \in \mathrm{T}_{L'}\mathscr{L}(V,\omega)$ be represented by a curve γ. Without loss of generality, we may assume that γ stays in $\mathscr{L}_0(L'^{\perp})$, where $L'^{\perp} = J(L')$ for some chosen $J \in \mathscr{J}(V,\omega)$. Then, Proposition 7.6.6/2 assigns to γ a curve S_γ in $\mathrm{S}^2 L'^*$. Define $\mathsf{Q}_X \in \mathrm{S}^2 L'^*$ by

$$\mathsf{Q}_X := -\dot{S}_\gamma(0).$$

The assignment $X \mapsto \mathsf{Q}_X$ is a vertical vector bundle morphism from $\mathrm{T}\mathscr{L}(V,\omega)$ to the vector bundle over $\mathscr{L}(V,\omega)$ with the fiber over L' being $\mathrm{S}^2 L'^*$. The argument of the proof of Theorem 1.1(1) in [252] shows that Q_X does not depend on J, which entered the definition presented here through $L'^{\perp}$. Now, let $\gamma : [a,b] \to \mathscr{L}(V,\omega)$ be a curve and let $L \in \mathscr{L}(V,\omega)$ be given. For every crossing t of γ with $\hat{\mathscr{L}}(L)$, define a crossing form by

$$\Gamma_{\gamma,L,t} := (\mathsf{Q}_{\dot{\gamma}(t)})_{\restriction \gamma(t) \cap L}. \tag{7.7.5}$$

A crossing is said to be regular if $\Gamma_{\gamma,L,t}$ is non-degenerate. This generalizes the notion of transversality for crossings with $\mathscr{L}_1(L)$ to arbitrary crossings. Regular crossings are isolated. For curves γ which have only regular crossings, the Maslov index relative to L is defined by

$$\mathrm{Ind}_L(\gamma) = \frac{1}{2}\,\mathrm{sign}\,\Gamma_{\gamma,L,a} + \sum_{a<t<b} \mathrm{sign}\,\Gamma_{\gamma,L,t} + \frac{1}{2}\,\mathrm{sign}\,\Gamma_{\gamma,L,b}, \tag{7.7.6}$$

where the sum runs over all inner crossings and sign denotes the signature.[9] One can show that every curve is homotopic with fixed end points to a curve with only regular crossings and that Ind_L is invariant under homotopies with fixed end points. Thus, Ind_L naturally extends to arbitrary curves. The Maslov index so defined has

[9] The number of positive minus the number of negative eigenvalues.

a number of natural properties, listed in Theorem 2.3 of [252]. In particular, it is additive with respect to the composition of curves and satisfies

$$\operatorname{Ind}_L(\gamma) = \operatorname{Ind}_{aL}(a\gamma), \quad a \in \operatorname{Sp}(V, \omega). \tag{7.7.7}$$

The latter is a consequence of the invariance of the crossing form (7.7.5) under symplectomorphisms (Exercise 7.7.3). Finally, to express $\Gamma_{\gamma,L,t}$ for a crossing t with $\gamma(t) \in \mathscr{L}_k(L)$ in local charts, choose an orthonormal basis $\{e_i\}$ in L. For an appropriately chosen index set K consisting of k elements, let $A(s) := \varphi_K(\gamma(t+s))$. Then,

$$\Gamma_{\gamma,L,t}(v, w) = \sum_{i,j=1}^{n} v^i \dot{A}_{ij}(0) w^j, \quad v, w \in \gamma(t) \cap L, \tag{7.7.8}$$

where v^i and w^j denote the coefficients with respect to the basis $\{e_i\}$ (Exercise 7.7.4).

Exercises

7.7.1 Prove that the mapping

$$\mathrm{U}(1) \times \mathrm{SU}(n) \to \mathrm{U}(n), \qquad (\alpha, a) \mapsto \operatorname{diag}(\alpha, 1, \ldots, 1)a,$$

is a diffeomorphism. Use this and the fact that $\mathrm{SU}(n)$ is simply connected to show that every closed curve in $\mathrm{U}(n)$ is homotopic to one of the closed curves

$$\gamma_k : [0, 1] \to \mathrm{U}(n), \qquad \gamma_k(t) := \operatorname{diag}\big(\mathrm{e}^{\mathrm{i}2\pi kt}, 1, \ldots, 1\big).$$

This completes the proof of Proposition 7.7.3.

7.7.2 Determine $\mathscr{L}(V, \omega)$ for $V = \mathbb{R}^2$ and analyze the Maslov indices in this example.

7.7.3 Prove that the crossing form (7.7.5) is invariant under symplectic transformations, that is,

$$\Gamma(a\Lambda, aL, t) \circ a = \Gamma(\Lambda, L, t), \quad a \in \operatorname{Sp}(n, \mathbb{R}).$$

7.7.4 Verify Formula (7.7.8). Analyze this formula by bringing $A(s)$ for small s to an appropriate block form and showing that the corresponding $(k \times k)$-block of $\dot{A}$ spans the normal bundle of $\mathscr{L}_k(L)$.

7.8 The Kashiwara Index

In this section, we discuss another very useful index, which in the literature is usually called the Kashiwara index.[10]

[10] Some authors call it Wall-Kashiwara index, others Hörmander-Kashiwara index. We refer to the textbooks by de Gosson [72], Libermann and Marle [181], as well as to the Lecture Notes of

By Remark 7.6.2/2, the symplectic group acts transitively on the space of Lagrangian subspaces of the symplectic vector space (V, ω), that is, any two Lagrangian subspaces can be transformed into one another via a symplectomorphism. A pair (L_1, L_2) of Lagrangian subspaces can be transformed via a symplectomorphism into another pair (L'_1, L'_2) of Lagrangian subspaces iff $\dim(L_1 \cap L_2) = \dim(L'_1 \cap L'_2) = k$, because in this case there exists a symplectic basis $\{e_i, f_i\}$, such that L_1 and L_2 are spanned by $\{e_1, \dots, e_n\}$ and $\{e_1, \dots, e_k, f_{k+1}, \dots, f_n\}$, respectively (Exercise 7.8.1). That is, the action of $\mathrm{Sp}(V, \omega)$ on $\mathscr{L}(V, \omega) \times \mathscr{L}(V, \omega)$ has $n + 1$ orbits, labeled by $\dim(L_1 \cap L_2)$. It turns out that the symplectic group does not act transitively on triples (L_1, L_2, L_3) of Lagrangian subspaces with prescribed dimensions of

$$L_1 \cap L_2, \qquad L_1 \cap L_3, \qquad L_2 \cap L_3, \qquad L_1 \cap L_2 \cap L_3.$$

This is related to the fact that for triples there is one further symplectic invariant.

Definition 7.8.1 (Kashiwara index) Let (V, ω) be a symplectic vector space. The Kashiwara index $s(L_1, L_2, L_3)$ of a triple (L_1, L_2, L_3) of Lagrangian subspaces of (V, ω) is defined as the signature of the quadratic form

$$\mathsf{Q}(L_1, L_2, L_3)(v_1, v_2, v_3) := \omega(v_1, v_2) + \omega(v_2, v_3) + \omega(v_3, v_1) \tag{7.8.1}$$

on the vector space $L_1 \oplus L_2 \oplus L_3$.

Example 7.8.2 Let us consider the canonical symplectic vector space $(\mathbb{R}^2, -J_1)$, cf. Example 7.1.4. We denote $\mathbf{v}_i = (q_i, p_i)$, $i = 1, 2, 3$, and choose

$$L_1 = \mathbb{R} \times \{0\}, \qquad L_2 := \big\{(q_2, p_2) : p_2 = a q_2\big\}, \qquad L_3 = \{0\} \times \mathbb{R},$$

where $a \in \mathbb{R}$ is fixed. This yields

$$\begin{aligned} \mathsf{Q}(L_1, L_2, L_3)(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) &= -\mathbf{v}_1^{\mathrm{T}} J_1 \mathbf{v}_2 - \mathbf{v}_2^{\mathrm{T}} J_1 \mathbf{v}_3 - \mathbf{v}_3^{\mathrm{T}} J_1 \mathbf{v}_1 \\ &= -a q_1 q_2 - p_3 q_2 + p_3 q_1. \end{aligned}$$

Diagonalizing this quadratic form in the variables q_1, q_2 and p_3, one gets (Exercise 7.8.2)

$$\mathsf{Q}(L_1, L_2, L_3) = x^2 - y^2 - \operatorname{sign}(a) z^2. \tag{7.8.2}$$

From this formula we read off

$$s(L_1, L_2, L_3) = \begin{cases} -1 & \text{for } a > 0 \\ 0 & \text{for } a = 0 \\ 1 & \text{for } a < 0. \end{cases} \tag{7.8.3}$$

Meinrenken [208], where the reader can find an exhaustive treatment including historical remarks. An axiomatic approach can be found in [58].

We note that the index s measures the relative position of the three subspaces with respect to a chosen orientation: if $\mathbb{R}^2$ is oriented clockwise, s vanishes if two of the three subspaces coincide, its value is $+1$ if L_2 is located between L_1 and L_3 (relative to the orientation) and it is equal to -1 otherwise. For the counter-clockwise orientation, the signs are reverted. In the sequel, we will see that an analogous result holds for any n.

Proposition 7.8.3 *The Kashiwara index s has the following properties.*

1. *Symplectic invariance*: $s(aL_1, aL_2, aL_3) = s(L_1, L_2, L_3)$ *for every* $a \in \mathrm{Sp}(V, \omega)$.
2. *Antisymmetry*: $s(L_{\pi(1)}, L_{\pi(2)}, L_{\pi(3)}) = (-1)^{\mathrm{sign}(\pi)} s(L_1, L_2, L_3)$ *for every permutation* π.
3. *Additivity: for all triples* (L_1', L_2', L_3') *and* (L_1'', L_2'', L_3'') *of Lagrangian subspaces in* (V', ω') *and* (V'', ω''), *respectively,* $(L_1' \oplus L_1'', L_2' \oplus L_2'', L_3' \oplus L_3'')$ *is a triple of Lagrangian subspaces in* $(V', \omega') \oplus (V'', \omega'')$ *and one has*

$$s\big(L_1' \oplus L_1'', L_2' \oplus L_2'', L_3' \oplus L_3''\big) = s'\big(L_1', L_2', L_3'\big) + s''\big(L_1'', L_2'', L_3''\big).$$

4. *Normalization*: *the triple* $(\mathbb{R}, (1+\mathrm{i})\mathbb{R}, \mathrm{i}\mathbb{R})$ *in* $\mathbb{R}^2$ *has Kashiwara index* 1.
5. *Cocycle property*:

$$s(L_1, L_2, L_3) = s(L_2, L_3, L_4) - s(L_1, L_3, L_4) + s(L_1, L_2, L_4)$$

for arbitrary Lagrangian subspaces L_1, L_2, L_3 *and* L_4 *in* (V, ω).

Proof Points 1–3 are obvious.

4. This is Example 7.8.2 with $a = 1$ and $-J_1$ replaced by J_1. Hence, $\mathrm{sign}(\mathrm{Q}) = +1$.

5. According to Remark 7.2.7, there exists a Lagrangian subspace L which is transversal to each L_i. Choose an orthonormal basis $\{e_j\}$ in L, let $\varphi : \mathscr{L}_0(L) \to \mathrm{S}^2\mathbb{R}^n$ be the diffeomorphism (7.6.12) and let $A_i := \varphi(L_i)$. Then, for each i, the vectors $f_j + \sum_{l=1}^n (A_i)_{jl} e_l$ form a basis in L_i. With respect to these bases, $\mathrm{Q}(L_i, L_j, L_k)$ is given by the matrix

$$\hat{\mathrm{Q}} = \begin{bmatrix} 0 & A_i - A_j & A_k - A_i \\ A_i - A_j & 0 & A_j - A_k \\ A_k - A_i & A_j - A_k & 0 \end{bmatrix}. \tag{7.8.4}$$

The reader easily checks that $\hat{\mathrm{Q}}$ is block-diagonalized by the help of

$$M = \begin{bmatrix} 0 & \mathbb{1} & \mathbb{1} \\ \mathbb{1} & 0 & \mathbb{1} \\ \mathbb{1} & \mathbb{1} & 0 \end{bmatrix}.$$

One obtains

$$\frac{1}{2}M^T\hat{Q}M = \begin{bmatrix} A_j - A_k & 0 & 0 \\ 0 & A_k - A_i & 0 \\ 0 & 0 & A_i - A_j \end{bmatrix}. \tag{7.8.5}$$

Thus,

$$s(L_i, L_j, L_k) = \operatorname{sign}(A_i - A_j) + \operatorname{sign}(A_j - A_k) + \operatorname{sign}(A_k - A_i).$$

This yields the assertion. □

Remark 7.8.4

1. One can show that properties 1–4 define the Kashiwara index uniquely, see [58]. Thus, these properties can be used for an axiomatic definition of s. If a system $\tilde{s}(L_1, L_2, L_3)$ has only properties 1–3, it is proportional to the Kashiwara index s.
2. For the special case $L_1 \cap L_3 = 0$ we obtain

$$s(L_1, L_2, L_3) = \operatorname{sign}(q), \tag{7.8.6}$$

where q is a quadratic form on L_2, given by $q(v) := \omega(v_1, v_3)$, where v_1 and v_3 denote the components of $v \in L_2$ with respect to the decomposition $V = L_1 \oplus L_3$ (Exercise 7.8.2). This is the situation of Example 7.8.2. Thus, in this case, (7.8.6) can be taken as the definition of s. If, in addition, $L_1 \cap L_2 = 0$ and $L_3 = L_1^{\perp}$, Formula (7.8.6) yields

$$s\big(L_1, L_2, L_1^{\perp}\big) = \operatorname{sign}(A), \tag{7.8.7}$$

where $A = \varphi(L_2)$ and φ is the diffeomorphism (7.6.12) (Exercise 7.8.2). This generalizes (7.8.3) to the case $n > 1$.
3. Point 5 of Proposition 7.8.3 means that the Kashiwara index defines a 2-cocycle on $\mathscr{L}(V, \omega)$ with respect to the natural coboundary operator

$$\partial : \operatorname{Map}\big(\mathscr{L}(V,\omega)^3, \mathbb{R}\big) \to \operatorname{Map}\big(\mathscr{L}(V,\omega)^4, \mathbb{R}\big),$$

given by

$$\partial f(L_1, L_2, L_3, L_4) = f(L_1, L_2, L_3) - f(L_2, L_3, L_4) + f(L_1, L_3, L_4) - f(L_1, L_2, L_4),$$

see [72, §1.4.2].

The Kashiwara index has a number of topological properties.

Proposition 7.8.5 *Let $L_i : [a, b] \to \mathscr{L}(V, \omega)$ be curves, $i = 1, 2, 3$.*

1. *If the dimensions of $L_1(t) \cap L_2(t)$, $L_2(t) \cap L_3(t)$ and $L_3(t) \cap L_1(t)$ are constant in t, so is the Kashiwara index $s(L_1(t), L_2(t), L_3(t))$.*

2. *Let L be a Lagrangian subspace transversal to both $L_1(t)$ and $L_2(t)$ for every t. Then, the difference*

$$s\big(L_1(a), L_2(a), L\big) - s\big(L_1(b), L_2(b), L\big) \tag{7.8.8}$$

does not depend on the choice of L.

Proof 1. It suffices to prove that $s(L_1(t), L_2(t), L_3(t))$ is constant in a neighbourhood of t_0 for all $a \le t_0 \le b$. Choose $L \in \mathscr{L}(V,\omega)$ transversal to $L_i(t_0)$. There exists $\varepsilon > 0$ such that L remains transversal to $L_i(t)$ for $|t - t_0| < \varepsilon$. As in the proof of Proposition 7.8.3/5, we use an orthonormal basis in L to represent $\mathsf{Q}(L_1(t), L_2(t), L_3(t))$ for every such t by a matrix $\hat{\mathsf{Q}}(t)$, given by (7.8.4). Since $\hat{\mathsf{Q}}(t)$ depends continuously on t, if the rank is constant, the signature must be constant, too. Diagonalizing $\hat{\mathsf{Q}}(t)$, we obtain (7.8.5) and hence

$$\dim\ker\hat{\mathsf{Q}} = \dim\big(L_1(t)\cap L_2(t)\big) + \dim\big(L_2(t)\cap L_3(t)\big) + \dim\big(L_1(t)\cap L_3(t)\big).$$

Since $\operatorname{rank}\hat{\mathsf{Q}}(t) = 3n - \dim\ker\hat{\mathsf{Q}}(t)$, the assertion follows.

2. Let L and L' be two different Lagrangian subspaces with the required property. Then, the cocycle property implies

$$s\big(L_1(a), L_2(a), L\big) - s\big(L_1(a), L_2(a), L'\big) = s\big(L_2(a), L, L'\big) - s\big(L_1(a), L, L'\big)$$

and analogously

$$s\big(L_1(b), L_2(b), L\big) - s\big(L_1(b), L_2(b), L'\big) = s\big(L_2(b), L, L'\big) - s\big(L_1(b), L, L'\big).$$

By point 1, $s(L_1(t), L, L')$ and $s(L_2(t), L, L')$ are constant in t. This yields the assertion. □

Remark 7.8.6

1. Point 1 of Proposition 7.8.5 implies the following. Let L_1 and L_2 be Lagrangian subspaces and let $t \mapsto L(t)$ be a curve of Lagrangian subspaces, such that for every t, the subspace $L(t)$ is transversal both to L_1 and to L_2. Then, the Kashiwara index $s(L_1, L_2, L(t))$ is constant in t.
2. One can show that, up to a symplectomorphism, every triple of Lagrangian subspaces is determined by the following five numbers [208]:

$$\dim(L_1\cap L_2),\qquad \dim(L_1\cap L_3),\qquad \dim(L_2\cap L_3),$$
$$\dim(L_1\cap L_2\cap L_3),\qquad s(L_1, L_2, L_3).$$

Now we can define the Maslov intersection index for an arbitrary pair of curves of Lagrangian subspaces.

Definition 7.8.7 (Maslov intersection index for pairs) For a given pair of curves $L_1, L_2 : [a, b] \to \mathscr{L}(V, \omega)$, choose $a = t_0 < t_1 < \cdots < t_r = b$ such that for every $j = 1, \ldots, r$ there exists a Lagrangian subspace L^j which is transversal to both $L_1(t)$ and $L_2(t)$ for all $t \in [t_{j-1}, t_j]$. Then, the Maslov intersection index of L_1 and L_2 is defined by

$$[L_1 : L_2] := \frac{1}{2} \sum_{j=1}^{r} \Big(s\big(L_1(t_{j-1}), L_2(t_{j-1}), L^j\big) - s\big(L_1(t_j), L_2(t_j), L^j\big)\Big). \tag{7.8.9}$$

By Remark 7.2.7, a partition of the interval $[a, b]$ having the desired properties exists. By Proposition 7.8.5/2, $[L_1 : L_2]$ neither depends on the chosen partition of $[a, b]$ nor on the choice of the Lagrangian subspaces L^j. Note that the curves L_1 and L_2 need not be transversal at the endpoints, cf. Remark 7.7.13. The intersection index of a curve, introduced in Definition 7.7.10, is a special case of the Maslov intersection index for pairs of curves.

Proposition 7.8.8 *Let $L \in \mathscr{L}(V, \omega)$ and let γ be a curve in $\mathscr{L}(V, \omega)$ with end points in $\mathscr{L}_0(L)$. Then, $[\gamma : L] = \mathrm{Ind}_L(\gamma)$, where L stands for the constant curve $t \mapsto L$.*

Proof Without loss of generality, we may assume that γ has only simple transversal crossings with $\hat{\mathscr{L}}(L)$. Then, the partition of the domain of γ in (7.8.9) can be chosen in such a way that $L \cap \gamma(t_j) = \{0\}$ for all j and every segment contains at most one crossing. By Lemma 7.8.5/1, segments without crossings do not contribute to $[\gamma : L]$. If the j-th segment contains a crossing, we write the corresponding summand in (7.8.9) in the form

$$\frac{1}{2}\Big(s\big(\gamma(t_{j-1}), L, L^j\big) - s\big(\gamma(t_j), L, L^j\big)\Big) = -\frac{1}{2}\Big(s\big(L, \gamma(t_{j-1}), L^j\big) - s\big(L, \gamma(t_j), L^j\big)\Big)$$

and apply Formula (7.8.6). To determine the quadratic form q, we choose an orthonormal basis $\{e_i\}$ in L and an appropriate chart φ_K, with $K = \{m\}$. Denote

$$A(t) := \varphi_K\big(\gamma(t)\big) \equiv \big(\varphi \circ a_K^{-1}\big)\big(\gamma(t)\big),$$

cf. (7.6.16). Hence, according to Remark 7.6.7/2, $\gamma(t)$ is spanned by the vectors $a_K(f_i + \sum_{j=1}^n A_{ij} e_j)$, $i = 1, \ldots, n$, where $f_i = J e_i$. A straightforward calculation (Exercise 7.8.6) yields that q is given by

$$q(v) = \sum_{i,j \neq m} A_{ij} v^i v^j - A_{mm} v^m v^m, \tag{7.8.10}$$

where $v = \sum_{i=1}^n v^i a_K(f_i + \sum_{j=1}^n A_{ij} e_j) \in \gamma(t)$. Thus, if the crossing is positive, the signature of $A(t)$ jumps by $+2$. Therefore, in this case, the j-th summand in

(7.8.9) is

$$\frac{1}{2}\big(s\big(\gamma(t_{j-1}), L, L^j\big) - s\big(\gamma(t_j), L, L^j\big)\big) = 1.$$

This yields the assertion. □

Remark 7.8.9

1. The Hörmander index for a 4-tuple $(L_1, L_2; M_1, M_2)$ of Lagrangian subspaces with the property that each L_i is transversal to each M_i is defined by

$$\sigma(L_1, L_2; M_1, M_2) := \frac{1}{2}\big(s(L_1, L_2, M_2) - s(L_1, L_2, M_1)\big),$$

cf. [141]. By means of this index, the sum in (7.8.9) can be written in the form

$$\begin{aligned}[L_1 : L_2] :=& \frac{1}{2}s\big(L_1(a), L_2(a), L^1\big) - \frac{1}{2}s\big(L_1(b), L_2(b), L^r\big) \\ &+ \sum_{j=1}^{r-1} \sigma\big(L_1(t_j), L_2(t_j); L^j, L^{j+1}\big).\end{aligned}$$

The Hörmander index has the following properties:

$$\sigma(L_1, L_2; M_1, M_2) = -\sigma(L_2, L_1; M_1, M_2) = -\sigma(M_1, M_2; L_1, L_2), \quad (7.8.11)$$

$$\sigma(L_1, L_2; M_1, M_2) + \sigma(L_1, L_2; M_2, M_3) = \sigma(L_1, L_2; M_1, M_3). \quad (7.8.12)$$

2. There is a number of generalizations of the concept of Maslov index. In particular, the Arnold-Leray-Maslov index is of importance. This index is defined on the universal covering manifold[11] of the Lagrange-Graßmann manifold, see [72]. There also exists a generalization to the case of arbitrary n-tuples, see [289].

Exercises

7.8.1 Determine the stabilizers and the orbits of the action of the symplectic group $\mathrm{Sp}(V, \omega)$ on $\mathscr{L}(V, \omega) \times \mathscr{L}(V, \omega)$.

7.8.2 Prove Formulae (7.8.2), (7.8.6) and (7.8.7).

7.8.3 Let $(\mathbb{R}^2, -J_1)$ be the canonical symplectic vector space in two dimensions and let $\{e, f\}$ be the canonical basis. Calculate $[L_1 : L_2]$ for $L_1(t) = \mathbb{R}f$ and $L_2(t) = \mathbb{R}(f + te)$ for all $a \le t \le b$.

7.8.4 Calculate the sum $[L_1 : L_2] + [L_2 : L_3] + [L_3 : L_1]$ for arbitrary curves $L_i : [a, b] \to \mathscr{L}(V, \omega)$.

7.8.5 Prove the following. If $L_1 = (L_1 \cap L_2) + (L_1 \cap L_3)$, then $s(L_1, L_2, L_3) = 0$.

7.8.6 Prove Formula (7.8.10).

[11] This manifold is also called the universal Maslov bundle, see [72].

Chapter 8
Symplectic Geometry

Symplectic geometry plays a tremendous role both in pure mathematics and in physics. As we will see, it provides the natural mathematical language for the study of Hamiltonian systems. In this chapter, we present the basic notions of symplectic geometry. The starting point will be the Theorem of Darboux, which states that, locally, all symplectic structures of a given dimension are equivalent to the standard symplectic vector space structure on $\mathbb{R}^{2n}$ defined in the previous chapter. Thus, in sharp contrast to the situation in Riemannian geometry, there are no local invariants. Symplectic manifolds of the same dimension can at most differ globally.[1] According to the above equivalence, many structures of local symplectic geometry have their origin in symplectic algebra.

The second elementary, but very important observation is that on symplectic manifolds the symplectic form provides a duality between smooth functions and certain vector fields, called Hamiltonian vector fields. As an immediate consequence of this duality, we obtain the notion of Poisson structure. Given the great importance of Poisson structures both in mathematics and in physics, we go beyond the symplectic case and give a brief introduction to general Poisson manifolds.

There are two classes of symplectic manifolds which are especially important in physical applications: cotangent bundles and orbits of the coadjoint representation of a Lie group. We will see that the cotangent bundle serves as a mathematical model of phase space and coadjoint orbits are relevant in the study of systems with symmetries. Both classes of examples are discussed in detail.

Moreover, in this chapter we discuss elementary properties of coisotropic submanifolds, present a number of natural generalizations of the Darboux Theorem and give an introduction to general symplectic reduction. The important special case of symplectic reduction for systems with symmetries will be dealt with in Chap. 10. The more advanced theory of Lagrangian submanifolds (including topological as-

[1] There is a huge field of research called symplectic topology, which deals with the study of global invariants of symplectic manifolds. For a nice intuitive introduction to this field we refer the reader to an article of Arnold [22]. There is a number of detailed expositions of this subject, see e.g. [206].

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*,
Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7_8, © Springer Science+Business Media Dordrecht 2013

pects and the study of singularities) is contained in Chap. 12, where it finds one of its natural applications.

Finally, the last section of this chapter is devoted to an elementary introduction to Morse theory. As we will see, the basic notions of this theory can be naturally formulated in the language of symplectic geometry. Moreover, concepts of Morse theory are of special importance in the study of symplectic manifolds and Hamiltonian systems. This is related to the above-mentioned duality between Hamiltonian vector fields and functions. This duality yields a relation between the critical points of vector fields and the critical points of functions. Thus, qualitative dynamics of a Hamiltonian vector field can be investigated using methods of Morse theory.

8.1 Basic Notions. The Darboux Theorem

Let M be a manifold. The notions of linear symplectic algebra have the following counterparts on the level of M. Every differential 2-form $\omega \in \Omega^2(M)$ induces a vertical vector bundle morphism $\omega^\flat : \mathrm{T}M \to \mathrm{T}^*M$ by

$$\left\langle \omega^\flat(X), Y \right\rangle := \omega_m(X, Y), \quad X, Y \in \mathrm{T}_m M. \tag{8.1.1}$$

The kernel and the rank of ω are defined to be the kernel and the rank of $\omega^\flat$. The 2-form ω is called non-degenerate if $\omega^\flat$ is an isomorphism. This is equivalent to the requirement that $\omega_m \in \bigwedge^2(\mathrm{T}^*M)$ be non-degenerate for all $m \in M$. If $\omega^\flat$ is an isomorphism, the inverse mapping is denoted by $\omega^\sharp$. We often write $X^\flat \equiv \omega^\flat(X)$ and $\alpha^\sharp \equiv \omega^\sharp(\alpha)$.

Definition 8.1.1 (Symplectic manifold and symplectic mapping)

1. A symplectic manifold is a pair (M, ω) consisting of a manifold M and a closed non-degenerate 2-form ω, called the symplectic form or the symplectic structure.
2. A smooth mapping $\Phi : M \to N$ of symplectic manifolds (M, ω) and (N, ρ) is called symplectic, or canonical, if $\Phi^*\rho = \omega$. If Φ is in addition a (local) diffeomorphism, it is called a (local) symplectomorphism.

If (M, ω) is a symplectic manifold, then $(\mathrm{T}_m M, \omega_m)$ is a symplectic vector space for all $m \in M$. Therefore, a symplectic manifold must have even dimension. Moreover, the tangent bundle $\mathrm{T}M$ is a symplectic vector bundle.[2] A smooth mapping $\Phi : M \to N$ is symplectic iff the linear mapping $\Phi'_m : \mathrm{T}_m M \to \mathrm{T}_{\Phi(m)} N$ is symplectic for all $m \in M$. This implies, in particular, that every symplectic mapping is an immersion, see Proposition 7.4.1/1. As in the case of symplectic mappings of symplectic vector spaces, the inverse of a symplectomorphism is symplectic. Hence, if one takes the symplectic mappings as the morphisms of symplectic manifolds,

[2] A symplectic vector bundle is a real vector bundle E endowed with a section ω in $\bigwedge^2 E^*$ which is fibrewise symplectic.

the symplectomorphisms are the corresponding isomorphisms. Finally, we note that Definition 8.1.1 extends in an obvious way to manifolds with boundary, thus yielding the notions of symplectic manifold with boundary and symplectomorphism of symplectic manifolds with boundary.

Example 8.1.2

1. Every symplectic vector space (V, ω) is a symplectic manifold, where ω is viewed as a differential 2-form by means of the identification of the tangent spaces of V with V itself. This applies in particular to the canonical symplectic vector space $(\mathbb{R}^{2n}, \omega_0)$ discussed in Example 7.1.4.
2. Let $M = \mathbb{C}^n$, viewed as a (real) manifold with the coordinates x_i given by $z_k = x_k + i x_{n+k}$, $k = 1, \dots, n$. The form
$$\omega = \frac{i}{2} \sum_k \mathrm{d}z_k \wedge \mathrm{d}\bar{z}_k$$
is symplectic. The coordinates x_i define a symplectomorphism from $(\mathbb{C}^n, \omega)$ onto $(\mathbb{R}^{2n}, \omega_0)$.
3. Every volume form on a 2-dimensional manifold is symplectic.
4. The cotangent bundle T^*Q of a manifold Q carries a natural symplectic structure, see Sect. 8.3. This fact is one of the cornerstones of the theory of Hamiltonian systems. In this context, Q plays the role of the configuration space and T^*Q is the phase space of the system.
5. Every orbit of the coadjoint representation of a Lie group carries a natural symplectic structure, see Sect. 8.4. This fact is of special importance for the theory of Hamiltonian systems with symmetries, discussed in Chap. 10.
6. The product of two symplectic manifolds (M, ω_M) and (N, ω_N) is symplectic, where the symplectic form may be taken as the sum or the difference of the pullbacks of the symplectic forms of the factors under the natural projections,
$$\omega^{\pm}_{M\times N} = \mathrm{pr}_M^*\, \omega_M \pm \mathrm{pr}_N^*\, \omega_N. \tag{8.1.2}$$

Let (M, ω) be a symplectic manifold of dimension $2n$. By Proposition 7.1.7, the $2n$-form $\omega^n = \omega \wedge \cdots \wedge \omega$ is a volume form. In particular, symplectic manifolds are orientable. The form
$$\Omega_\omega = \frac{1}{n!}(-1)^{\frac{n(n-1)}{2}}\,\omega^n \tag{8.1.3}$$
is called the canonical volume form, or the Liouville form, of (M, ω).

Proposition 8.1.3 *Every symplectic mapping between symplectic manifolds of the same dimension is a volume preserving local diffeomorphism.*

Proof Obviously, a symplectic mapping Φ preserves the canonical volume. Moreover, as already observed, Proposition 7.4.1 implies that Φ' is fibrewise injective.

For dimensional reasons, it is fibrewise bijective then and the Inverse Mapping Theorem implies that Φ is a local diffeomorphism. □

Remark 8.1.4

1. By Proposition 8.1.3, the total volume of a symplectic manifold is a global symplectic invariant.
2. Since $\mathrm{d}\omega = 0$, in a neighbourhood U of every point there exists a local 1-form β such that $\omega_{\upharpoonright U} = \mathrm{d}\beta$. Such a form is called a local symplectic potential. A symplectic manifold is said to be exact if there exists a global symplectic potential.

   Geometrically, the requirement $\mathrm{d}\omega = 0$ means that the symplectic surface area of any 3-dimensional ball is equal to zero. This yields an obstruction for (even-dimensional) manifolds to admit a symplectic structure. In more detail, let us consider the case of an orientable compact manifold. Being closed, ω defines an element $[\omega] \in H^2(M, \mathbb{R})$, which must be nonzero, because its n-th power is the class of a volume form on an orientable compact manifold and hence nonzero. Thus, orientable compact manifolds fulfilling $H^2(M, \mathbb{R}) = 0$ do not admit a symplectic structure. This applies, in particular, to any sphere S^{2n} with $n > 1$. The study of the existence and uniqueness problem of symplectic structures for compact manifolds is a field of active research in symplectic topology, see Chaps. 7 and 13 of [206]. In this context, a whole tool kit of operations for constructing symplectic manifolds has been developed. On the other hand, for noncompact manifolds (without boundary), Gromov has shown that a symplectic form exists iff the manifold admits a non-degenerate 2-form [206, Thm. 7.34]. Concerning the uniqueness aspect, it is worthwile to note that even on $\mathbb{R}^{2n}$ there exist exotic symplectic structures, not symplectomorphic to the canonical one.
3. The symplectomorphisms of (M, ω) form a group with respect to composition, denoted by $\mathrm{Symp}(M, \omega)$. This group is at the heart of symplectic topology, see [206], Chap. 10. We will briefly discuss it in Sect. 8.8.

The following theorem is fundamental for the theory of symplectic manifolds. It tells us that, locally, all symplectic manifolds of the same dimension are isomorphic. This in accordance with the corresponding statement in symplectic algebra, see Corollary 7.1.5.

Theorem 8.1.5 (Darboux) *Let (M, ω) be a symplectic manifold of dimension $2n$. For every $m \in M$, there exists a symplectomorphism from an open neighbourhood of m in M onto an open subset of the canonical symplectic vector space* $(\mathbb{R}^{2n}, \omega_0)$.

Proof The proof is based on a deformation method developed by Moser [218]. Since the statement is local, without loss of generality we may assume that $M = \mathbb{R}^{2n}$ and $m = 0$. According to Corollary 7.1.5, there exists a basis in $\mathbb{R}^{2n}$ such that $\omega(0)$ coincides with the canonical symplectic form ω_0. For $t \in [0, 1]$, let

$$\omega_t := \omega_0 + t \cdot \tilde{\omega},$$

where $\tilde{\omega} := \omega - \omega_0$. Then, $\mathrm{d}\omega_t = 0$ and $\omega_t(0) = \omega(0) = \omega_0$ is non-degenerate for all $t \in [0, 1]$. Since the function $[0, 1] \times \mathbb{R}^{2n} \ni (t, \mathbf{x}) \to \det \omega_t(\mathbf{x}) \in \mathbb{R}$ is continuous, ω_t remains non-degenerate on an open neighbourhood W of $[0, 1] \times \{0\}$ in $\mathbb{R} \times \mathbb{R}^{2n}$. Since the interval $[0, 1]$ is compact and $\mathbb{R}^{2n}$ is locally compact, the Tube Lemma of elementary topology yields an open ball K about the origin of $\mathbb{R}^{2n}$ and an open interval I containing $[0, 1]$ such that $I \times K \subset W$. On the other hand, since $\mathrm{d}\tilde{\omega} = 0$, the Poincaré Lemma yields a 1-form α on K such that $\tilde{\omega} = \mathrm{d}\alpha$. By adding an exact 1-form if necessary, we can achieve that $\alpha(0) = 0$. Since ω_t is non-degenerate on K for all $t \in I$, we can define a time-dependent vector field X on K by

$$X_t \lrcorner \omega_t = -\alpha, \quad t \in I.$$

Let Φ be the flow of X and let $\mathscr{D} \subset I \times I \times K$ be the domain[3] of Φ. Since $\alpha(0) = 0$, we have $X_t(0) = 0$ for all $t \in I$, that is, 0 is a fixed point of Φ. Hence, $[0, 1] \times [0, 1] \times \{0\} \subset \mathscr{D}$ and by applying the Tube Lemma once again, we find an open neighbourhood $U \subset K$ of 0 such that $[0, 1] \times [0, 1] \times U \subset \mathscr{D}$. Writing $\Phi_{t,0} \equiv \Phi_t$ and using (4.1.28), as well as (4.1.30) with d replaced by Φ_t^*, we compute

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\Phi_t^* \omega_t\big) = \Phi_t^*\left(\mathscr{L}_{X_t}\omega_t + \frac{\mathrm{d}}{\mathrm{d}t}\omega_t\right) = \Phi_t^*\big(\mathrm{d}(X_t \lrcorner \omega_t) + \tilde{\omega}\big) = \Phi_t^*(-\mathrm{d}\alpha + \tilde{\omega}) = 0.$$

Taken pointwise on U, this implies that $\Phi_1^*\omega = \Phi_1^*\omega_1 = \Phi_0^*\omega_0 = \omega_0$. Thus, $\Phi = \Phi_1^{-1}$ is the desired local symplectomorphism. □

Remark 8.1.6

1. By the Darboux Theorem and by Corollary 7.1.5, for every $m \in M$, there exists a local chart (U, κ) at m such that

$$\omega_{\restriction U} = \sum_{i=1}^{n} \mathrm{d}\kappa^i \wedge \mathrm{d}\kappa^{n+i}.$$

Moreover, by Proposition 7.2.9, the complementary Lagrangian subspaces $\mathbb{R}^n \times \{0\}$ and $\{0\} \times \mathbb{R}^n$ of $\mathbb{R}^{2n}$ define a symplectomorphism between the canonical symplectic vector space $(\mathbb{R}^{2n}, \omega_0)$ and the vector space $\mathbb{R}^n \oplus (\mathbb{R}^n)^*$ endowed with the natural symplectic form (7.1.5). By composing κ with this symplectomorphism, we obtain a chart taking values in $\mathbb{R}^n \oplus (\mathbb{R}^n)^*$. Motivated by applications in physics, we denote the coordinates corresponding to the factors $\mathbb{R}^n$ and $(\mathbb{R}^n)^*$ by q^i and p_i, respectively. Then,

$$\omega_{\restriction U} = \sum_{i=1}^{n} \mathrm{d}p_i \wedge \mathrm{d}q^i \equiv \mathrm{d}p_i \wedge \mathrm{d}q^i$$

[3]See Sect. 3.4 for the notation.

(if there is no danger of confusion, we will use the summation convention). The elements of the corresponding local frame in TM will be denoted by ∂_{q^i} and ∂_{p_i}. Both types of charts will be referred to as Darboux charts and the corresponding coordinates as Darboux coordinates.

2. Let (M_1, ω_1) and (M_2, ω_2) be symplectic manifolds of the same dimension and let $\Phi\colon M_2 \to M_1$ be a smooth mapping. In Darboux coordinates q^i, p_i on M_1 and $\bar{q}^i$, $\bar{p}_i$ on M_2, the condition on Φ to be symplectic reads

$$\Phi^*\big(\mathrm{d}p_i \wedge \mathrm{d}q^i\big) = \mathrm{d}\bar{p}_i \wedge \mathrm{d}\bar{q}^i. \tag{8.1.4}$$

Using the simplified notation

$$\big(q^i \circ \Phi \circ \kappa^{-1}\big) \equiv q^i, \qquad \big(p_i \circ \Phi \circ \kappa^{-1}\big) \equiv p_i,$$

where κ denotes the chart defined by $\bar{q}^i$, $\bar{p}_i$, a straightforward calculation (Exercise 8.1.1) yields that Φ is symplectic, and hence a local symplectomorphism, iff

$$\frac{\partial p_i}{\partial \bar{p}_j}\frac{\partial q^i}{\partial \bar{p}_k} - \frac{\partial p_i}{\partial \bar{p}_k}\frac{\partial q^i}{\partial \bar{p}_j} = 0 = \frac{\partial p_i}{\partial \bar{q}^j}\frac{\partial q^i}{\partial \bar{q}^k} - \frac{\partial p_i}{\partial \bar{q}^k}\frac{\partial q^i}{\partial \bar{q}^j}, \qquad \frac{\partial p_i}{\partial \bar{p}_j}\frac{\partial q^i}{\partial \bar{q}^k} - \frac{\partial p_i}{\partial \bar{q}^k}\frac{\partial q^i}{\partial \bar{p}_j} = \delta^j_k.$$

3. Let q^i and p_i be Darboux coordinates on U and let $\tilde{p}_i$ be functions on U such that q^i and $\tilde{p}_i$ are Darboux coordinates, too. Consider the functions $f_i := \tilde{p}_i - p_i$ on U. Taking q^i and p_i as coordinates on U, we find

$$0 = \mathrm{d}f_i \wedge \mathrm{d}q^i = \frac{\partial f_i}{\partial q^j}\mathrm{d}q^j \wedge \mathrm{d}q^i + \frac{\partial f_i}{\partial p_j}\mathrm{d}p_j \wedge \mathrm{d}q^i.$$

Hence, $\frac{\partial f_i}{\partial p_j} = 0$ and $\frac{\partial f_i}{\partial q^j} - \frac{\partial f_j}{\partial q^i} = 0$ for all i, j. The general solution of this system of equations is given by $f_i = \alpha_i \circ q$ with smooth functions α_i on $q(U) \subset \mathbb{R}^n$ which combine to a closed 1-form $\alpha = \alpha_i \mathrm{d}x^i$ on $q(U)$, where x^i denote the standard coordinates on $\mathbb{R}^n$.

The types of subspaces of a symplectic vector space given in Definition 7.2.2 carry over to submanifolds of a symplectic manifold:

Definition 8.1.7 Let (M, ω) be a symplectic manifold. An immersion $\varphi\colon N \to M$ is called isotropic (coisotropic, Lagrangian or symplectic) at $p \in N$ iff $\varphi'\mathrm{T}_p N$ is an isotropic (coisotropic, Lagrangian or symplectic) subspace of $(\mathrm{T}_{\varphi(p)}M, \omega_{\varphi(p)})$. If this is true for all points $p \in N$, the pair (N, φ) is called an isotropic (coisotropic, Lagrangian or symplectic) immersion.

These notions apply, in particular, to the case when (N, φ) is a submanifold of M. Since the pull-back $\varphi^*\omega$ of the symplectic form to the submanifold (N, φ) is the counterpart of the restriction ω_W to a subspace W of the symplectic vector space (V, ω), Proposition 7.2.3 yields: the submanifold (N, φ) is

(a) isotropic iff $\varphi^*\omega = 0$,
(b) Lagrange iff $\varphi^*\omega = 0$ and $2\dim N = \dim M$,
(c) symplectic iff $\varphi^*\omega$ is non-degenerate.

According to Remark 1.6.2/1, we may assume that N is given by a subset of M and that φ is the natural inclusion mapping. In this case, from (7.2.2) and (7.2.5) we obtain

$$\ker\big(\varphi^*\omega\big)_p = \mathrm{T}_p N \cap (\mathrm{T}_p N)^\omega, \tag{8.1.5}$$

$$\max(0, 2\dim N - \dim M) \leq \operatorname{rank}\big(\varphi^*\omega\big)_p \leq \dim N, \tag{8.1.6}$$

and the statements of Proposition 7.2.4 carry over pointwise. In particular, a submanifold of dimension 1 (codimension 1) is always isotropic (coisotropic) and every submanifold contained in an isotropic submanifold is isotropic. A submanifold containing a coisotropic submanifold need not be coisotropic though, because Proposition 7.2.4/4 does not apply to points outside the latter.

Finally, consider a vector subbundle $E \subset \mathrm{T}M$ over a submanifold $N \subset M$. The symplectic orthogonal of E is defined by $E^\omega := \bigcup_{p\in N} E_p^\omega$. This is a subbundle of $\mathrm{T}M$ over the submanifold N, because according to Proposition 7.2.1/1, it is mapped to the annihilator E^0 under the vertical vector bundle isomorphism $\omega^\flat : \mathrm{T}M \to \mathrm{T}^*M$. The vector subbundle E is called isotropic (coisotropic, Lagrangian or symplectic), if every fibre is isotropic (coisotropic, Lagrangian or symplectic), that is, if, respectively, $E \subset E^\omega$, $E^\omega \subset E$, $E^\omega = E$ or $E \cap E^\omega = s_0(N)$, where s_0 denotes the zero section of $\mathrm{T}M$. Thus, a submanifold $N \subset M$ is isotropic, coisotropic, Lagrangian or symplectic iff so is the vector subbundle $\mathrm{T}N \subset \mathrm{T}M$. In this case, the symplectic orthogonal $(\mathrm{T}N)^\omega$ of $\mathrm{T}N$ in $\mathrm{T}M$ is referred to as the symplectic normal bundle of the submanifold N.

Lemma 8.1.8 *Let (M, ω) be a symplectic manifold. Every Lagrangian subbundle $E \subset \mathrm{T}M$ over a submanifold $N \subset M$ admits a Lagrangian complement, that is, a Lagrangian subbundle $\tilde{E} \subset \mathrm{T}M$ over N such that $E \oplus \tilde{E} = \mathrm{T}M_{\restriction N}$.*

Proof The proof follows [305], Lecture 2. Choose an auxiliary Riemannian metric g on M. For every $m \in M$, the mapping $F_m : \mathrm{S}^2_+(\mathrm{T}^*_m M) \to \mathscr{J}(\mathrm{T}_m M, \omega_m)$ of Proposition 7.5.6 assigns to g_m an ω_m-compatible complex structure J_m on $\mathrm{T}_m M$. Since the endomorphism associated with ω_m via g_m is bijective, its absolute square with respect to g_m is strictly positive. Therefore, the square root and hence J_m depend smoothly on m. As a consequence, the J_m combine to a vertical vector bundle morphism J of $\mathrm{T}M$ satisfying $J^2 = -\mathrm{id}_{\mathrm{T}M}$.[4] It follows that $J(E)$ is a vector subbundle of $\mathrm{T}M$ over N. By Proposition 7.5.3/2, it is Lagrange and complementary to E. □

[4] A vertical vector bundle morphism of $\mathrm{T}M$ with this property is called an almost complex structure on M.

Exercises

8.1.1 Prove point 2 of Remark 8.1.6.

8.1.2 Do T^2, $S^3 \times S^1$ or $\mathbb{C}P^2$ admit a symplectic structure?

8.1.3 Show that on a symplectic manifold (M, ω) of dimension greater than 2, a function multiple $f\omega$ of ω is symplectic iff f is nonzero and locally constant.

8.2 Hamiltonian Vector Fields and Poisson Structures

Via the exterior derivative and the natural isomorphism $\omega^\sharp = (\omega^\flat)^{-1} : \mathrm{T}^*M \to \mathrm{T}M$, one can assign a vector field to every smooth function on M. This assignment is basic for the theory of Hamiltonian systems, see Chap. 9.

Definition 8.2.1 (Hamiltonian vector field) Let (M, ω) be a symplectic manifold and let $f \in C^\infty(M)$. The vector field

$$X_f := -(\mathrm{d}f)^\sharp$$

is called the Hamiltonian vector field generated by f.

The flow of X_f is referred to as the Hamiltonian flow generated by f. The definition of X_f is equivalent to

$$X_f \lrcorner \omega = -\mathrm{d}f. \tag{8.2.1}$$

In particular, the 1-form $(X_f)^\flat$ is exact.

Remark 8.2.2 Let us write down X_f in local Darboux coordinates q^i and p_i. For that purpose, we decompose $X_f = A_i \partial_{p_i} + B^i \partial_{q^i}$ with respect to the corresponding local frame $(\partial_{q^i}, \partial_{p_i})$ and determine the coefficients A_i, B^i from (8.2.1), which in coordinates reads

$$\big(A_i \partial_{p_i} + B^i \partial_{q^i}\big) \lrcorner \big(\mathrm{d}p_j \wedge \mathrm{d}q^j\big) = A_i \mathrm{d}q^i - B^i \mathrm{d}p_i = -\partial_{q^i} f \mathrm{d}q^i - \partial_{p_i} f \mathrm{d}p_i .$$

Thus,

$$X_f = (\partial_{p_i} f)\partial_{q^i} - (\partial_{q^i} f)\partial_{p_i} . \tag{8.2.2}$$

Besides the Hamiltonian vector fields, the infinitesimal symmetries of the symplectic structure are important:

Definition 8.2.3 (Symplectic vector field) A vector field X on (M, ω) is called symplectic if ω is invariant under the flow of X.

Proposition 8.2.4 *Let (M, ω) be a symplectic manifold and let X be a vector field on M. The following statements are equivalent*:

1. *X is symplectic.*
2. $\mathscr{L}_X \omega = 0$.
3. *$X^\flat$ is closed.*
4. *X is locally Hamiltonian, that is, for every $m \in M$ there exists a function f on a neighbourhood U of m such that $X_{\restriction U} = X_f$.*

Proof The equivalence of points 2 and 3 follows from

$$\mathscr{L}_X \omega = X \lrcorner \mathrm{d}\omega + \mathrm{d}(X \lrcorner \omega) = \mathrm{d}X^\flat .$$

To see that point 2 implies point 1, we apply (4.1.28) to obtain

$$\frac{\mathrm{d}}{\mathrm{d}s}\big(\Phi_s^* \omega\big)_m = \big(\Phi_s^*(\mathscr{L}_X \omega)\big)_m$$

for all s between 0 and t and all m in the domain of Φ_t. Due to $\Phi_0^* = \mathrm{id}$, integration from 0 to t yields the assertion. The remaining statements are obvious.

□

Remark 8.2.5

1. According to (8.2.1) and Proposition 8.2.4/3, every Hamiltonian vector field is symplectic.
2. Definition 8.2.3 carries over to time-dependent vector fields by requiring that $\Phi_{t_1,t_2}^* \omega = \omega$ for all pairs (t_1, t_2). Then, Proposition 8.2.4 holds with X replaced by X_t.

We denote the set of Hamiltonian vector fields on (M, ω) by $\mathfrak{X}_{\mathrm{H}}(M, \omega)$ and the set of symplectic vector fields by $\mathfrak{X}_{\mathrm{LH}}(M, \omega)$.[5]

Proposition 8.2.6 *Let (M, ω) be a symplectic manifold.*

1. *$\mathfrak{X}_{\mathrm{LH}}(M, \omega)$ is a Lie subalgebra of $\mathfrak{X}(M)$.*
2. *For $Y, Z \in \mathfrak{X}_{\mathrm{LH}}(M, \omega)$, we have*

$$[Y, Z] = X_{\omega(Y,Z)} . \tag{8.2.3}$$

In particular, $\mathfrak{X}_{\mathrm{H}}(M, \omega)$ is an ideal in $\mathfrak{X}_{\mathrm{LH}}(M, \omega)$.

Proof 1. According to Proposition 3.3.3, we have $\mathscr{L}_{[X,Y]}\omega = [\mathscr{L}_X, \mathscr{L}_Y]\omega$.

2. Using $\mathscr{L}_Y \omega = 0$ and $\mathrm{d}(Z \lrcorner \omega) = 0$, we find

$$[Y, Z] \lrcorner \omega = \mathscr{L}_Y (Z \lrcorner \omega) - Z \lrcorner \mathscr{L}_Y \omega = Y \lrcorner \mathrm{d}(Z \lrcorner \omega) + \mathrm{d}\big(Y \lrcorner (Z \lrcorner \omega)\big) = -\mathrm{d}\big(\omega(Y, Z)\big).$$

□

[5] With the index LH standing for locally Hamiltonian.

Remark 8.2.7 Viewing $\omega^\flat$ as a mapping from $\mathfrak{X}(M)$ to $\Omega^1(M)$, we obtain vector space isomorphisms

$$\mathfrak{X}_{\mathrm{H}}(M,\omega) \cong B^1(M), \qquad \mathfrak{X}_{\mathrm{LH}}(M,\omega) \cong Z^1(M)$$

and hence

$$H^1(M) \cong \mathfrak{X}_{\mathrm{LH}}(M,\omega)/\mathfrak{X}_{\mathrm{H}}(M,\omega).$$

Thus, $\mathfrak{X}_{\mathrm{H}}(M,\omega) = \mathfrak{X}_{\mathrm{LH}}(M,\omega)$ iff the first de Rham cohomology of M is trivial.

Example 8.2.8 Consider the torus T^2 with the angle coordinates x and y and symplectic form $\omega = \mathrm{d}x \wedge \mathrm{d}y$. Let

$$X = a\partial_x + b\partial_y$$

be a vector field with constant coefficients $a, b \neq 0$. Since $X^\flat = a\mathrm{d}y - b\mathrm{d}x$, we have $\mathrm{d}X^\flat = 0$ and, therefore, $X \in \mathfrak{X}_{\mathrm{LH}}(\mathrm{T}^2, \omega)$. However, as a nowhere vanishing vector field on a compact manifold, X cannot be Hamiltonian, because every smooth function on M has a maximum and a minimum and hence critical points.

Symplectic mappings can be characterized in terms of Hamiltonian vector fields:

Proposition 8.2.9 *A smooth mapping $\Phi : M \to N$ of symplectic manifolds (M,ω) and (N,ρ) is symplectic iff it is an immersion and the Hamiltonian vector fields $X_{\Phi^* f}$ on M and X_f on N are Φ-related for all $f \in C^\infty(N)$. In particular, a diffeomorphism $\Phi : M \to N$ is a symplectomorphism iff*

$$\Phi_* X_h = X_{h \circ \Phi^{-1}} \tag{8.2.4}$$

for all $h \in C^\infty(M)$.

Proof Since every symplectic mapping is an immersion, we have to show that for an immersion Φ, $\Phi^*\rho = \omega$ is equivalent to the condition that

$$\Phi' \circ X_{\Phi^* f} = X_f \circ \Phi$$

for all $f \in C^\infty(N)$. By taking the interior product of this equation with ρ and evaluating both sides on $\Phi'(Y)$, where $Y \in \mathrm{T}_m M$, we find

$$\big(\Phi^*\rho\big)_m (X_{\Phi^* f}, Y) = (X_f \lrcorner \rho)_{\Phi(m)} \big(\Phi'(Y)\big).$$

Hence, we have to show that $\Phi^*\rho = \omega$ is equivalent to

$$X_{\Phi^* f} \lrcorner \big(\Phi^*\rho\big) = -\mathrm{d}\big(\Phi^* f\big) \tag{8.2.5}$$

for all $f \in C^\infty(N)$. If $\Phi^*\rho = \omega$, then (8.2.5) holds by definition of $X_{\Phi^* f}$. Conversely, if (8.2.5) holds, then $X_{\Phi^* f} \lrcorner (\Phi^*\rho) = X_{\Phi^* f} \lrcorner \omega$ for all $f \in C^\infty(N)$. Since

Φ is an immersion, the mapping $(\Phi'_m)^{\mathrm{T}} : \mathrm{T}^*_{\Phi(m)}N \to \mathrm{T}^*_m M$ is surjective for every $m \in M$, hence every covector at m is of the form $\mathrm{d}(\Phi^* f)$ for some $f \in C^\infty(N)$. Accordingly, every tangent vector at m is of the form $X_{\Phi^* f}$ for some $f \in C^\infty(N)$. Hence, (8.2.5) implies $\Phi^*\rho = \omega$. Finally, if Φ is a diffeomorphism, by setting $h = f \circ \Phi^{-1}$ in the defining equation of Φ-relation we obtain (8.2.4). □

Since for a symplectic vector field X with flow Φ, the mappings

$$(\Phi_t)'_m : \mathrm{T}_m M \to \mathrm{T}_{\Phi_t(m)} M$$

are linear symplectomorphisms, the Symplectic Eigenvalue Theorem 7.4.3 yields

Proposition 8.2.10 *Let (M, ω) be a symplectic manifold and let $X \in \mathfrak{X}_{\mathrm{LH}}(M)$.*

1. *If m is an equilibrium of X and if μ is a characteristic exponent of m with multiplicity k, then $\overline{\mu}$ and $-\mu$ are also characteristic exponents of m with multiplicity k. The multiplicity of the characteristic exponent 0 is even.*
2. *If γ is a periodic integral curve of X and if λ is a characteristic multiplier of γ with multiplicity k, then $\overline{\lambda}$ and λ^{-1} are also characteristic multipliers of γ with multiplicity k. The multiplicity of the characteristic multiplier 1 is odd and nonzero.*

Proof 1. The characteristic exponents of m are given by the eigenvalues of the Hessian endomorphism $\mathrm{Hess}_m(X) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 (\Phi_t)'_m$. Since $\mathrm{Hess}_m(X)$ is an element of the symplectic Lie algebra $\mathfrak{sp}(\mathrm{T}_m M, \omega_m)$, the assertion follows from point 2 of the Symplectic Eigenvalue Theorem 7.4.3 and from Corollary 7.4.4/2.

2. Let T be the period of γ and let $m \in \gamma$. By Remark 3.6.11/1, the characteristic multipliers of γ are given by the eigenvalues of $(\Phi_T)'_m$, with the eigenvalue 1 corresponding to the eigenspace $\mathrm{T}_m\gamma$ omitted. Since $(\Phi_T)'_m$ is an element of the symplectic group $\mathrm{Sp}(\mathrm{T}_m M, \omega_m)$, the assertion follows from point 1 of the Symplectic Eigenvalue Theorem 7.4.3 and from Corollary 7.4.4/1. □

Remark 8.2.11

1. The characteristic multipliers of γ, with the multiplicity of the characteristic multiplier 1 reduced by 1, are called the Floquet multipliers of γ. It will be shown in Sect. 9.5 that the Floquet multipliers coincide with the eigenvalues of the tangent mapping of an isoenergetic Poincaré mapping, see Remark 9.5.2/2.
2. A subset of the set of characteristic exponents of an equilibrium m (of the set of Floquet multipliers of a periodic integral curve γ), counted with multiplicities, is called a basis set for m (for γ) if it contains one member of each of the pairs $(\mu, \overline{\mu})$ (the pairs $(\lambda, \overline{\lambda})$). By construction, a basis set has n elements ($n-1$ elements), where $2n = \dim M$.

Next, let us study the surjective mapping

$$C^\infty(M) \ni f \mapsto X_f \in \mathfrak{X}_{\mathrm{H}}(M, \omega).$$

Since $\mathfrak{X}_{\mathrm{H}}(M, \omega)$ carries a natural Lie algebra structure, it is reasonable to ask whether there is a Lie algebra structure on $C^\infty(M)$ which makes this mapping into a homomorphism. Such a structure exists, indeed:

Definition 8.2.12 (Poisson bracket) Let (M, ω) be a symplectic manifold and let $f, h \in C^\infty(M)$. The function

$$\{f, h\} := \omega(X_f, X_h)$$

is called the Poisson bracket of f and h.

Using $\omega(X_f, X_g) = -X_f \lrcorner (X_g \lrcorner \omega) = X_f \lrcorner (\mathrm{d}g) = X_f(g)$, the Poisson bracket can be rewritten as

$$\{f, g\} = X_f(g) = -X_g(f). \tag{8.2.6}$$

Proposition 8.2.13 *Let (M, ω) be a symplectic manifold.*

1. *The Poisson bracket defines on $C^\infty(M)$ the structure of a Lie algebra. With respect to this structure, the mapping $f \mapsto X_f$ is a Lie algebra homomorphism.*
2. *The Poisson bracket satisfies the Leibniz rule, that is, for all $f, g, h \in C^\infty(M)$,*

$$\{f, gh\} = \{f, g\}h + \{f, h\}g.$$

Proof 1. Obviously, the Poisson bracket is bilinear and antisymmetric and the assignment $f \mapsto X_f$ is linear. Thus, it is enough to show that the Poisson bracket fulfils the Jacobi identity and that $[X_f, X_g] = X_{\{f,g\}}$ holds for all $f, g \in C^\infty(M)$. The latter follows from (8.2.3). Using this and (8.2.6), we verify the Jacobi identity:

$$\begin{aligned}&\big\{f, \{g, h\}\big\} + \big\{g, \{h, f\}\big\} + \big\{h, \{f, g\}\big\}\\ &\quad = X_f\big(X_g(h)\big) - X_g\big(X_f(h)\big) - X_{\{f,g\}}(h) = 0.\end{aligned}$$

2. This follows from the fact that vector fields are derivations of $C^\infty(M)$. □

Remark 8.2.14

1. Due to Proposition 3.2.15 and the fact that the mapping $f \mapsto X_f$ is a Lie algebra homomorphism, two functions on M Poisson-commute iff the flows of their Hamiltonian vector fields commute.
2. Let us write down the Poisson bracket in a Darboux chart κ with coordinates q^i and p_i. Using (8.2.6) and (8.2.2), we find

$$\{f, g\} = X_f(g) = (\partial_{p_i} f)(\partial_{q^i} g) - (\partial_{q^i} f)(\partial_{p_i} g).$$

Using the simplified notation $f \equiv f \circ \kappa^{-1}$, we arrive at the standard formula

$$\{f, g\} = \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i}. \tag{8.2.7}$$

In particular, the Poisson brackets of the coordinate functions are

$$\{q^i, q^j\} = 0, \qquad \{p_i, p_j\} = 0, \qquad \{p_i, q^j\} = \delta_i{}^j.$$

Symplectic mappings can be characterized in terms of the Poisson bracket as well:

Proposition 8.2.15 *A smooth mapping $\Phi : M \to N$ of symplectic manifolds (M, ω) and (N, ρ) is symplectic iff it is an immersion and for all $f, g \in C^\infty(N)$ one has*

$$\{\Phi^* f, \Phi^* g\} = \Phi^*\{f, g\}. \tag{8.2.8}$$

Proof Due to

$$\{\Phi^* f, \Phi^* g\}(m) = \big(\Phi'(X_{\Phi^* f})_m\big)(g), \qquad \big(\Phi^*\{f, g\}\big)(m) = (X_f)_{\Phi(m)}(g),$$

the condition that $\{\Phi^* f, \Phi^* g\} = \Phi^*\{f, g\}$ for all $f, g \in C^\infty(N)$ is equivalent to the condition that the vector fields $X_{\Phi^* f}$ and X_f are Φ-related for all $f \in C^\infty(N)$. Hence, the assertion follows from Proposition 8.2.9. □

In the remaining part of this section we digress from symplectic manifolds to give a brief introduction to the more general notion of Poisson structure. For an exhaustive treatment of Poisson manifolds, we refer to the books of Vaisman [295] and Waldmann [301], where the reader can find a lot of further references. Pioneering work on this subject goes back to Sophus Lie [182–184].

Definition 8.2.16 (Poisson manifold)

1. A Poisson structure on a manifold M is a Lie algebra structure $\{,\}$ on $C^\infty(M)$ fulfilling the Leibniz rule $\{f, gh\} = g\{f, h\} + \{f, g\}h$ for all $f, g, h \in C^\infty(M)$. The pair $(M, \{,\})$ is called a Poisson manifold.
2. A smooth mapping $\Phi\colon M \to N$ of Poisson manifolds is called Poisson, or a Poisson morphism, if $\Phi^*\{f, g\} = \{\Phi^* f, \Phi^* g\}$ for all $f, g \in C^\infty(N)$.

According to Propositions 8.2.13 and 8.2.15, every symplectic manifold is Poisson and a smooth mapping of symplectic manifolds is symplectic iff it is Poisson and an immersion. The Leibniz rule implies that for every $f \in C^\infty(M)$, the mapping $\{f, \cdot\}\colon C^\infty(M) \to C^\infty(M)$ is a derivation and hence defines a vector field X_f, called the Hamiltonian vector field of f. Thus,

$$X_f(g) = \{f, g\} = -X_g(f) \tag{8.2.9}$$

for all $f, g \in C^\infty(M)$. The Jacobi identity implies

$$[X_f, X_g] = X_{\{f,g\}}, \tag{8.2.10}$$

hence the mapping $f \mapsto X_f$ is a Lie algebra homomorphism. Moreover, from

$$\{f, g\}(m) = \langle (X_f)_m, (\mathrm{d}g)_m \rangle = -\langle (X_g)_m, (\mathrm{d}f)_m \rangle$$

we read off that the value of $\{f, g\}$ at m depends on $(\mathrm{d}f)_m$ and $(\mathrm{d}g)_m$ only and that this dependence is linear. Since every element of $\mathrm{T}^*_m M$ can be written as $\mathrm{d}f(m)$ for some $f \in C^\infty(M)$, it follows that there exists a unique bivector field $\Pi \in \Gamma(\bigwedge^2 TM)$ such that

$$\{f, g\} = \Pi(\mathrm{d}f, \mathrm{d}g) \tag{8.2.11}$$

for all $f, g \in C^\infty(M)$. Π is called the Poisson tensor of $(M, \{,\})$. The Jacobi identity implies

$$\mathscr{L}_{X_f} \Pi = 0 \tag{8.2.12}$$

for all $f \in C^\infty(M)$. The Poisson tensor Π defines a vertical vector bundle morphism

$$\Pi^\sharp : \mathrm{T}^*M \to \mathrm{T}M, \qquad \langle \beta, \Pi^\sharp(\alpha) \rangle := \Pi(\alpha, \beta), \tag{8.2.13}$$

fulfilling

$$\Pi^\sharp \circ \mathrm{d}f = X_f \tag{8.2.14}$$

for all $f \in C^\infty(M)$. The rank of $(M, \{,\})$ at m is defined to be the rank of $\Pi^\sharp_m$. Since Π is antisymmetric, this is an even number. For a symplectic manifold, $\Pi^\sharp = \omega^\flat$ and Π is the image of ω under the isomorphism $\omega^\sharp \wedge \omega^\sharp : \Omega^2(M) \to \Gamma(\bigwedge^2 \mathrm{T}M)$.

Poisson morphisms can be characterized in terms of the Poisson tensor as follows. From (8.2.11) we read off that a diffeomorphism $\Phi : M \to M$ is Poisson iff $\Phi_* \Pi = \Pi$. More generally, for a smooth mapping $\Phi : M \to N$ of Poisson manifolds one finds that it is Poisson iff the corresponding Poisson tensors are Φ-related[6] (Exercise 8.2.1).

Remark 8.2.17

1. In a local chart (U, κ) on M we have

$$\Pi = \frac{1}{2} \Pi^{ij} \partial_i \wedge \partial_j, \qquad \{f, g\} = \Pi^{ij} \partial_i f \partial_j g. \tag{8.2.15}$$

   In particular,

$$\Pi^{ij} = \{\kappa^i, \kappa^j\}, \qquad \{f, g\} = \{\kappa^i, \kappa^j\} \partial_i f \partial_j g.$$

[6] The notion of Φ-relation extends in an obvious way from vector fields to multivector fields, cf. Definition 2.3.6/1.

2. By Formula (8.2.11), every bivector field Π on M defines a bilinear antisymmetric mapping $\{,\}\colon C^\infty(M)\times C^\infty(M)\to C^\infty(M)$ which fulfils the Leibniz rule. It satisfies the Jacobi identity, and hence yields a Poisson structure, iff

$$\Pi\big(\mathrm{d}f,\mathrm{d}\big(\Pi(\mathrm{d}g,\mathrm{d}h)\big)\big)+\Pi\big(\mathrm{d}g,\mathrm{d}\big(\Pi(\mathrm{d}h,\mathrm{d}f)\big)\big)+\Pi\big(\mathrm{d}h,\mathrm{d}\big(\Pi(\mathrm{d}f,\mathrm{d}g)\big)\big)=0 \tag{8.2.16}$$

for all $f,g,h\in C^\infty(M)$. In local coordinates, this condition reads

$$\Pi^{in}\partial_n\Pi^{jk}+\Pi^{jn}\partial_n\Pi^{ki}+\Pi^{kn}\partial_n\Pi^{ij}=0$$

(Exercise 8.2.2). Thus, a Poisson manifold may as well be viewed as a pair (M,Π), where Π is a bivector field fulfilling (8.2.16).

Example 8.2.18

1. Every manifold M can be equipped with the trivial Poisson structure $\{f,g\}=0$ for all $f,g\in C^\infty(M)$. In this case, $\Pi=0$.
2. Let V be a real vector space. Every element $\Pi\in\bigwedge^2(V)$ defines a constant bivector field on V. This bivector field satisfies (8.2.16) and thus defines a Poisson structure (Exercise 8.2.3).
3. Consider the dual space V^* of a real vector space V. Recall that under the natural identification of the tangent spaces of V^* with V^* itself, every bivector field Π on V^* corresponds to a smooth mapping $V^*\to\bigwedge^2 V^*$. Π is called linear if this mapping is linear. In this case, for every $v,w\in V$, the function $\xi\mapsto\Pi_\xi(v,w)$ is linear and hence corresponds to an element of V. By assigning this element to the pair (v,w) we obtain an antisymmetric bilinear mapping $[,]\colon V\times V\to V$. Thus, $[v,w]$ is defined by

$$\big\langle\xi,[v,w]\big\rangle=\Pi_\xi(v,w) \tag{8.2.17}$$

for all $\xi\in V^*$. Conversely, if an antisymmetric bilinear mapping $[,]\colon V\times V\to V$ is given, (8.2.17) defines a linear bivector field Π on V^*. Let us analyze the condition (8.2.16) on Π in terms of $[,]$. First, we observe that (8.2.16) holds iff it holds for all linear functions f,g,h on V^*. For linear functions, we have $\mathrm{d}f=f$, and the linearity of Π implies $\mathrm{d}(\Pi(\mathrm{d}f,\mathrm{d}g))=\Pi(f,g)$, where f and g are viewed as elements of V. It follows that Π satisfies (8.2.16) iff $[,]$ satisfies the Jacobi identity. In this case, the corresponding Poisson bracket of $f,g\in C^\infty(V^*)$ is given by

$$\{f,g\}(\mu)=\big\langle\mu,\big[\mathrm{d}f(\mu),\mathrm{d}g(\mu)\big]\big\rangle, \tag{8.2.18}$$

where $\mathrm{d}f(\mu)$ and $\mathrm{d}g(\mu)$ are interpreted as elements of $(V^*)^*=V$.

Thus, the Poisson structures on V^* whose Poisson tensor is linear are in bijective correspondence with Lie algebra structures on V. In particular, (8.2.17) defines a natural Poisson structure on the dual space $\mathfrak{g}^*$ of a Lie algebra $\mathfrak{g}$. This structure is usually referred to as the Lie-Poisson structure on $\mathfrak{g}^*$. In the coordinates ξ_i defined by a basis in $\mathfrak{g}^*$, the Poisson tensor and the Poisson bracket of this structure are given by

$$\Pi(\xi) = \frac{1}{2}\xi_k c_{ij}^k \frac{\partial}{\partial \xi_i} \wedge \frac{\partial}{\partial \xi_j}, \qquad \{f,g\}(\xi) = \xi_k c_{ij}^k \frac{\partial f}{\partial \xi_i}(\xi)\frac{\partial g}{\partial \xi_j}(\xi), \tag{8.2.19}$$

where c_{ij}^k are the structure constants of $\mathfrak{g}$ with respect to the dual basis in $\mathfrak{g}$.

4. We know that every volume form ω on a two-dimensional manifold M is symplectic. Let Π_ω be the corresponding Poisson tensor. Then, every bivector field is of the form $f\Pi_\omega$ for some $f \in C^\infty(M)$ and $f\Pi_\omega$ defines a Poisson structure (Exercise 8.2.4). In contrast, $f\omega$ is symplectic only if f is nowhere vanishing.

Definition 8.2.19 (Poisson vector field) A vector field X on a Poisson manifold is called Poisson iff the Poisson structure is invariant under the flow of X.

Equivalently, X is Poisson iff $\mathscr{L}_X \Pi = 0$. This follows by the same argument as the corresponding assertion about symplectic vector fields in Proposition 8.2.4. As a consequence, (8.2.12) implies that every Hamiltonian vector field is Poisson. In the special case of a symplectic manifold, a vector field is Poisson iff it is symplectic. In contrast to Proposition 8.2.4/4, however, on a Poisson manifold a Poisson vector field need not be locally Hamiltonian. For example, for the trivial Poisson structure, every vector field is Poisson, but only the zero vector field is Hamiltonian.

We conclude this section by showing that every Poisson manifold is foliated by symplectic manifolds. For that purpose, consider the characteristic distribution

$$D^\Pi := \Pi^\sharp\big(\mathrm{T}^*M\big) \subset \mathrm{T}M$$

of $(M, \{,\})$. By (8.2.14), this distribution is spanned by the Hamiltonian vector fields. For every $m \in M$, the Poisson tensor Π induces an antisymmetric 2-form ω_m on D_m^Π by

$$\omega_m(X,Y) := \Pi_m(\alpha, \beta), \tag{8.2.20}$$

where $\alpha, \beta \in \mathrm{T}^*M$ such that $X = \Pi^\sharp(\alpha)$ and $Y = \Pi^\sharp(\beta)$. The reader easily convinces himself that ω is well-defined and non-degenerate (Exercise 8.2.5). Thus, at each point m we obtain a symplectic vector space (D_m^Π, ω_m).

Theorem 8.2.20 (Symplectic Foliation Theorem) *The characteristic distribution of a Poisson manifold (M, Π) is integrable. On every integral manifold N, there exists a unique symplectic form ω_N such that the natural inclusion mapping is Poisson. This form is given by*

$$(\omega_N)_m(X_f, X_g) = \{f,g\}(m), \quad m \in N, \ f,g \in C^\infty(M). \tag{8.2.21}$$

The resulting foliation by maximal integral manifolds is called the symplectic foliation of $(M, \{,\})$ and the maximal integral manifolds are called the symplectic leaves of $(M, \{,\})$. By construction, the Hamiltonian vector fields are tangent to the symplectic leaves.

Proof According to (8.2.10), the set of Hamiltonian vector fields is involutive. Since every Hamiltonian vector field X_f is Poisson, its flow preserves Π. It follows that the rank of Π and hence the dimension of D^{Π} is constant along the integral curves of X_f. Thus, Theorem 3.5.10 implies integrability. Let N be an integral manifold. For every $m \in N$, the form $(\omega_N)_m$ coincides with the form defined in (8.2.20). Hence, it is well-defined and non-degenerate. It remains to prove that ω is closed. Since X_f is tangent to N, it induces a vector field $\tilde{X}_f$ on N. Using Proposition 4.1.6, together with Proposition 3.1.5/1 and Corollary 3.1.6, and the Jacobi identity for $\{,\}$, one finds $\mathrm{d}\omega_N(\tilde{X}_f, \tilde{X}_g, \tilde{X}_h) = 0$ for all $f, g, h \in C^\infty(M)$ (Exercise 8.2.6). This yields the assertion. □

Example 8.2.21

1. For the trivial Poisson structure $\{f, g\} = 0$ on M, the symplectic foliation is given by the points of M.
2. In Sect. 8.4 it will be shown that the symplectic leaves N of the Poisson structure on the dual space $\mathfrak{g}^*$ of a Lie algebra $\mathfrak{g}$, discussed in Example 8.2.18/3, coincide with the connected components of the coadjoint orbits and that the symplectic form ω_N is given by the Kirillov form.

Theorem 8.2.20 can be used to prove the following generalization of the Darboux Theorem to Poisson manifolds.

Theorem 8.2.22 (Splitting Theorem) *Let (M, Π) be a Poisson manifold of dimension k, let $m \in M$ and assume that Π has rank $2r$ at m. Then, there exists a local chart (U, κ) at m with coordinates $q^1, \ldots, q^r, p_1, \ldots, p_r, z^1, \ldots, z^{k-2r}$ such that*

$$\Pi_{\restriction U} = \partial_{q^i} \wedge \partial_{p_i} + \frac{1}{2}\pi^{ij}(z)\partial_{z^i} \wedge \partial_{z^j}, \qquad \pi^{ij}(0) = 0. \tag{8.2.22}$$

In these coordinates, the symplectic leaf through m is described by the equations $z^1 = z^2 = \cdots = z^{k-2r} = 0$ and q^i, p_i provide Darboux coordinates on this leaf.

Proof See [181], Theorem III.11.5, or [310]. □

Remark 8.2.23

1. The Splitting Theorem contains the following local decomposition property: let (M, Π) be a Poisson manifold of dimension k and rank $2r$ at the point m. Then, there exists a neighbourhood U of m which can be identified with the Cartesian product of a symplectic manifold V of dimension $2r$ and a Poisson manifold W of dimension $k - 2r$ whose rank vanishes at m. It was shown by Weinstein [310] that up to an isomorphism the Poisson structure on W does not depend on the chosen local coordinates.
2. A smooth function f on M is called a Casimir function for $\{,\}$ if $X_f = 0$. The set of Casimir functions is called the Poisson centre of $(M, \{,\})$. Indeed, this is the centre of the Lie algebra $(C^\infty(M), \{,\})$. One can show that in the case of

a symplectic manifold, every Casimir function is constant (Exercise 8.2.7) and that, in the general case, the restriction of a Casimir function to a symplectic leaf is constant.

Exercises

8.2.1 Show that a smooth mapping $\Phi : M \to N$ is Poisson iff the Poisson tensors of M and N are Φ-related.
8.2.2 Prove the statements of Remark 8.2.17/2.
8.2.3 Prove that every constant bivector field on a real vector space defines a Poisson structure.
8.2.4 Prove the statement of Example 8.2.18/4.
8.2.5 Show that the 2-form defined by (8.2.20) is well-defined and non-degenerate.
8.2.6 Complete the proof of Theorem 8.2.20 by showing that the 2-form ω_N is closed.
8.2.7 Prove that on a symplectic manifold, every Casimir function is constant.

8.3 The Cotangent Bundle

The cotangent bundle T^*Q of a manifold Q provides the basic model of a symplectic manifold. It has the following special properties.

(a) It carries a canonical symplectic potential.
(b) Every diffeomorphism of Q lifts to a symplectomorphism of T^*Q.
(c) It contains a variety of Lagrangian submanifolds related to the bundle structure.
(d) Since its fibres carry a natural affine structure, every 1-form on Q induces a vertical vector field on T^*Q.

In what follows, let $\pi : \mathrm{T}^*Q \to Q$ be the canonical projection. As already mentioned, in Hamiltonian mechanics, T^*Q is the phase space of a system with configuration space Q. Consequently, for the local description, we use the standard notation taken from physics: the coordinates of a local chart (U, κ) on Q are denoted by q^i. For the local frame $\{\partial_i^\kappa\}$ in $\mathrm{T}Q$ we write $\{\partial_{q^i}\}$ and for the local coframe $\{\mathrm{d}\kappa^i\}$ in T^*Q we write $\{\mathrm{d}q^i\}$. The chart κ induces a local chart κ_{T^*} on T^*U by

$$\kappa_{\mathrm{T}^*}(\xi) = \big(q^1\big(\pi(\xi)\big), \dots, q^n\big(\pi(\xi)\big), p_1(\xi), \dots, p_n(\xi)\big),$$

where the coordinate functions p_i are defined as the components of $\xi \in \mathrm{T}^*U$ with respect to the coframe $\{\mathrm{d}q^i\}$:

$$\xi = p_i(\xi)\mathrm{d}q^i.$$

Points of $\kappa_{\mathrm{T}^*}(\mathrm{T}^*U)$ will be written as pairs $(\mathbf{q}, \mathbf{p})$.

First, we show that T^*Q carries a natural exact symplectic structure.

Definition 8.3.1 (Canonical 1-form) The canonical 1-form θ, or Liouville form, on T^*Q is defined by

$$\langle \theta_\xi, X\rangle := \big\langle \xi, \pi'_\xi(X)\big\rangle,$$

where $\xi \in T^*Q$ and $X \in T_\xi(T^*Q)$.

We show that the 2-form $\omega := d\theta$ is a symplectic form on T^*Q. Obviously, $d\omega = 0$. To show that ω is non-degenerate, we calculate θ and ω in bundle coordinates q^i and p_i. The ansatz $\theta = \alpha_i dq^i + \beta^i dp_i$ and $X = A^i \partial_{q^i} + B_i \partial_{p_i}$ yields the defining equation

$$\alpha_i A^i + \beta^i B_i = p_i A^i$$

for all $A^i, B_i \in \mathbb{R}$, which is solved by $\alpha_i = p_i$ and $\beta^i = 0$. Thus, we obtain

$$\theta = p_i dq^i \quad \text{and} \quad \omega = dp_i \wedge dq^i. \tag{8.3.1}$$

In particular, ω is non-degenerate, indeed.

Definition 8.3.2 The form $\omega = d\theta$ is called the canonical symplectic form of T^*Q.

According to (8.3.1), for every chart on Q, the induced chart on T^*Q is a Darboux chart for the canonical symplectic structure.

Remark 8.3.3

1. The canonical 1-form θ is the unique 1-form on T^*Q with the property that

$$\alpha^*\theta = \alpha \tag{8.3.2}$$

   for all 1-forms α on Q (Exercise 8.3.1). As a consequence,

$$\alpha^*\omega = d\alpha. \tag{8.3.3}$$

2. Since the covectors dq^i form a basis in $T^*_x Q$ and the vectors ∂_{q^i} form a basis in $T_x Q$, the assignment $\partial_{q^i} \mapsto \partial_{q^i}$, $\partial_{p_i} \mapsto dq^i$ defines a vector space isomorphism $T_\xi(T^*Q) \to T_x Q \oplus T^*_x Q$. Since $\omega^\flat$ maps ∂_{q^i} to $-dp_i$, this isomorphism is symplectic with respect to the canonical symplectic structure (7.1.5) on $T_x Q \oplus T^*_x Q$.
3. The isomorphism $\omega^\sharp$ assigns to θ a vector field $\theta^\sharp$ on T^*Q, called the canonical vertical vector field or Liouville vector field. By definition,

$$\theta^\sharp \lrcorner \omega = \theta. \tag{8.3.4}$$

   In bundle coordinates, $\theta^\sharp$ reads $\theta^\sharp = p_i \partial_{p_i}$, which shows that it is vertical, indeed. Thus, $\theta^\sharp \lrcorner \theta = 0$, so that (4.1.24) implies

$$\mathscr{L}_{\theta^\sharp}\theta = \theta, \qquad \mathscr{L}_{\theta^\sharp}\omega = \omega. \tag{8.3.5}$$

Example 8.3.4 The special case of the cotangent bundle T^*G of a Lie group G is of particular interest. According to Proposition 5.1.6, both $\mathrm{T}G$ and T^*G are trivial. We use the (inverse) trivialization

$$\chi_{\mathrm{L}} \colon G \times \mathfrak{g}^* \to \mathrm{T}^*G, \qquad \chi_{\mathrm{L}}(a, \mu) := \mu \circ (\mathrm{L}_{a^{-1}})'_a \tag{8.3.6}$$

and write tangent vectors of $G \times \mathfrak{g}^*$ at (a, μ) in the form $(\mathrm{L}'_a A, \rho)$ with $A \in \mathfrak{g}$ and $\rho \in \mathfrak{g}^*$. Then,

$$\big(\chi_{\mathrm{L}}^*\theta\big)_{(a,\mu)}\big(L'_a A, \rho\big) = \theta_{\chi_{\mathrm{L}}(a,\mu)}\big(\chi'_{\mathrm{L}}\big(L'_a A, \rho\big)\big) = \big(\mu \circ (\mathrm{L}_{a^{-1}})'_a\big)\big(\pi' \circ \chi'_{\mathrm{L}}\big(L'_a A, \rho\big)\big).$$

Since $\pi \circ \chi_{\mathrm{L}}$ is the projection to the first component, we obtain

$$\big(\chi_{\mathrm{L}}^*\theta\big)_{(a,\mu)}\big(L'_a A, \rho\big) = \langle \mu, A \rangle. \tag{8.3.7}$$

Then, using Proposition 4.1.6 with left-invariant vector fields A, B on G and constant vector fields ρ, σ on $\mathfrak{g}^*$, we calculate (Exercise 8.3.2)

$$\omega_{(a,\mu)}\big(\big(L'_a A, \rho\big), \big(L'_a B, \sigma\big)\big) = \langle \rho, B \rangle - \langle \sigma, A \rangle - \big\langle \mu, [A, B] \big\rangle. \tag{8.3.8}$$

Next, we show that the canonical 1-form is invariant under the lift of diffeomorphisms from Q to T^*Q. This yields a special class of canonical transformations, called point transformations. Recall from Sect. 2.4 that every diffeomorphism $\varphi : Q \to Q$ induces a vector bundle isomorphism $\varphi'^{\mathrm{T}} : \mathrm{T}^*Q \to \mathrm{T}^*Q$ projecting to φ^{-1}. Let Φ be the inverse of this isomorphism. Φ is determined by the conditions $\pi \circ \Phi = \varphi \circ \pi$ and

$$\big\langle \Phi(\xi), \varphi' X \big\rangle = \langle \xi, X \rangle \tag{8.3.9}$$

for all $\xi \in \mathrm{T}^*_x Q$, $X \in \mathrm{T}_x Q$ and $x \in Q$. On the level of 1-forms α on Q, (8.3.9) can be rewritten as

$$\Phi \circ \big(\varphi^*\alpha\big) = \alpha \circ \varphi. \tag{8.3.10}$$

Definition 8.3.5 (Point transformation) The vector bundle isomorphism $\Phi = (\varphi'^{\mathrm{T}})^{-1}$ is called the point transformation induced by φ.

It turns out that point transformations can be characterized by the property that they leave the canonical 1-form θ invariant:

Proposition 8.3.6 *A fibre-preserving diffeomorphism $\Phi \colon \mathrm{T}^*Q \to \mathrm{T}^*Q$ is a point transformation iff $\Phi^*\theta = \theta$. In particular, every point transformation is a symplectomorphism.*

Proof Since Φ preserves the fibres, there is a unique diffeomorphism $\varphi : Q \to Q$ such that $\pi \circ \Phi = \varphi \circ \pi$. For $\xi \in \mathrm{T}^*Q$ and $X \in \mathrm{T}_\xi(\mathrm{T}^*Q)$ we compute

$$\big(\Phi^*\theta\big)_\xi(X) = \big\langle \Phi(\xi), \big(\pi' \circ \Phi'\big)(X) \big\rangle$$

$$= \langle \Phi(\xi), (\varphi' \circ \pi')(X) \rangle$$
$$= \langle \varphi'^{\mathrm{T}}(\Phi(\xi)), \pi'(X) \rangle.$$

Since $\theta_\xi(X) = \langle \xi, \pi'(X) \rangle$ and since π' is surjective, this implies that $\Phi^*\theta = \theta$ iff $\varphi'^{\mathrm{T}} \circ \Phi = \mathrm{id}_{\mathrm{T}^*Q}$, that is, iff Φ is the point transformation induced by φ. □

Remark 8.3.7 Let φ be a diffeomorphism of a manifold Q and let Φ be the induced point transformation. We compute the local representative of Φ with respect to local charts (U, κ) and (V, ρ) on Q with induced bundle coordinates q^i, p_i and $\bar{q}^i$, $\bar{p}_i$, respectively. The diffeomorphism φ is locally given by the functions

$$\bar{q}^i \circ \varphi \circ \kappa^{-1}.$$

Moreover, for $\xi \in \mathrm{T}^*Q$, the natural fibre coordinates $p_i(\xi)$ and $\bar{p}_i(\Phi(\xi))$ are determined by the equations $\xi = p_i(\xi)\mathrm{d}q^i$ and $\Phi(\xi) = \bar{p}_i(\Phi(\xi))\mathrm{d}\bar{q}^i$, respectively. Using (8.3.9), we obtain

$$\Phi(\xi) = p_i(\xi) \frac{\partial(q^i \circ \varphi^{-1} \circ \rho^{-1})}{\partial \bar{q}^j} \big((\rho \circ \varphi)(\pi(\xi))\big) \mathrm{d}\bar{q}^j.$$

Using the simplified notation

$$\bar{q}^i \circ \varphi \circ \kappa^{-1} \equiv \bar{q}^i, \qquad q^i \circ \varphi^{-1} \circ \rho^{-1} \equiv q^i,$$

we obtain the following local formula for Φ:

$$(\mathbf{q}, \mathbf{p}) \mapsto \left(\bar{q}(\mathbf{q}), p_j \frac{\partial q^j}{\partial \bar{\mathbf{q}}} \big(\bar{q}(\mathbf{q})\big) \right).$$

This formula can be interpreted both actively (a point transformation induced by a diffeomorphism) and passively (the change of bundle coordinates induced by a change of coordinates on Q).

Now, we turn to the discussion of Lagrangian submanifolds of cotangent bundles.

Example 8.3.8 The cotangent bundle T^*Q contains the following classes of Lagrangian submanifolds:

1. The fibres of T^*Q: let $x \in Q$ and let $i : \mathrm{T}^*_x Q \to \mathrm{T}^*Q$ be the natural inclusion mapping. Since $\mathrm{T}^*_x Q$ has half the dimension of T^*Q, it is enough to show isotropy. For $\xi \in \mathrm{T}^*_x Q$ and $X \in \mathrm{T}_\xi(\mathrm{T}^*_x Q)$, we find

$$\big(i^*\theta\big)_\xi(X) = \theta_{i(\xi)}\big(i'X\big) = \langle i(\xi), \pi' \circ i'(X) \rangle = 0.$$

   Thus, $i^*\theta = 0$ and hence $i^*\omega = 0$. In bundle coordinates the proof is even simpler: since q^i is constant along $\mathrm{T}^*_x Q$, we obtain $i^*(\mathrm{d}p_i \wedge \mathrm{d}q^i) = 0$.

2. The image of a closed 1-form: let $\alpha \in \Omega^1(Q)$. Since $\alpha(Q)$ has half the dimension of T^*Q, formula (8.3.3) implies that $\alpha(Q)$ is Lagrange iff $\mathrm{d}\alpha = 0$. In particular, the zero section and the image of an exact 1-form $\alpha = \mathrm{d}S$ is Lagrange. In the latter case, S is called a generating function for the Lagrangian submanifold $\alpha(Q)$. In bundle coordinates q^i and p_i, we have $\mathrm{d}S = \frac{\partial S}{\partial q^i}\mathrm{d}q^i$. Hence, in coordinates, $\alpha(Q)$ consists of the points $(\mathbf{q}, \mathbf{p} = \frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}))$.
3. The conormal bundle of a submanifold: let N be a submanifold of Q and let $(\mathrm{T}N)^0 \subset \mathrm{T}^*Q$ be the conormal bundle, cf. Remark 2.7.18/1. Obviously, $(\mathrm{T}N)^0$ has half the dimension of T^*Q. Moreover, the canonical 1-form θ vanishes on $(\mathrm{T}N)^0$, because π' maps vectors tangent to $(\mathrm{T}N)^0$ to vectors tangent to N. Thus, $(\mathrm{T}N)^0$ is Lagrange. Following Tulczyjew, we call it the canonical lift of N to T^*Q and denote it by $\widehat{N}$.
4. The canonical lift of the pair (N, S), where $N \subset Q$ is a submanifold and $S : N \to \mathbb{R}$ is a smooth function: define

$$\widehat{(N, S)} := \left\{\xi \in \left(\mathrm{T}^*Q\right)_{\restriction N} : \langle \xi, X\rangle = \langle \mathrm{d}S, X\rangle \text{ for all } X \in \mathrm{T}_{\pi(\xi)}N\right\}.$$

One can check that $j^*\theta = \pi_N^*\mathrm{d}S$, with j being the natural inclusion mapping of $\widehat{(N, S)}$ and $\pi_N : \widehat{(N, S)} \to N$ denoting the restriction of the canonical projection (Exercise 8.3.3). This implies that $\widehat{(N, S)}$ is Lagrange. Note that an obvious generalization is obtained by replacing $\mathrm{d}S$ by an arbitrary closed 1-form. Also note that the special cases $N = \{x\}$, $N = Q$ and $S = \text{const}$ exhaust the Lagrangian submanifolds of points 1–3.

From point 1 of Example 8.3.8 we get the following

Proposition 8.3.9 *The restriction* $(\mathrm{T}^*Q)_{\restriction N}$ *of* T^*Q *to a submanifold* N *of* Q *is a coisotropic submanifold of* T^*Q.

Proof For every $\xi \in (\mathrm{T}^*Q)_{\restriction N}$, the tangent space $\mathrm{T}_\xi((\mathrm{T}^*Q)_{\restriction N})$ contains the subspace $\mathrm{T}_\xi(\mathrm{T}^*_{\pi(\xi)}Q)$, which is Lagrange by Example 8.3.8/1. Hence, the assertion follows from Proposition 7.2.4/4. □

The following Proposition shows that point 2 of Example 8.3.8 (locally) exhausts the set of Lagrangian submanifolds in T^*Q which are transversal to the fibres.

Proposition 8.3.10 *Let* $L \subset \mathrm{T}^*Q$ *be a Lagrangian submanifold which is transversal to the fibres and which intersects every fibre at most once. Then,* $U := \pi(L)$ *is open in* Q *and there exists a closed* 1*-form* α *on* U *such that* $L = \alpha(U)$. *If* L *is in addition contractible, there exists a smooth function* $S : U \to \mathbb{R}$ *such that* $L = \mathrm{d}S(U)$.

In accordance with the terminology of Example 8.3.8/2, S is called a generating function for L. For Lagrangian submanifolds which are not transversal one has to use the more general concept of generating Morse family, see Sect. 12.4.

Proof By transversality, the restriction $\pi_{\restriction L}: L \to M$ is an immersion. Since L intersects every fibre at most once, $\pi_{\restriction L}$ is injective. Hence, U is a submanifold of Q. Since it has the same dimension as Q, it is open and $\pi_{\restriction L}$ is a diffeomorphism onto U. Thus, $\alpha := i_L \circ (\pi_{\restriction L})^{-1}$, with i_L denoting the natural inclusion mapping of L, is a local 1-form over U satisfying $L = \alpha(U)$. Since L is Lagrange, by Example 8.3.8/2, α is closed. If L is contractible, so is U. Thus, in this case, there exists a function $S: U \to \mathbb{R}$ such that $\alpha = \mathrm{d}S$ and hence $L = \mathrm{d}S(U)$. □

Finally, let us discuss the consequences of the fact that the fibres of T^*Q carry a natural affine structure. For given $x \in Q$, under the natural identification of the tangent spaces of the fibre $\mathrm{T}^*_x Q$ with $\mathrm{T}^*_x Q$ itself, every $\eta \in \mathrm{T}^*_x Q$ defines a constant vector field on $\mathrm{T}^*_x Q$, denoted by $\hat{\eta}$. Accordingly, every $\alpha \in \Omega^1(Q)$ defines a vertical vector field $\hat{\alpha}$ on T^*Q by

$$\hat{\alpha}_{\restriction \mathrm{T}^*_x Q} = \widehat{\alpha_x}. \tag{8.3.11}$$

This vector field is complete and its flow is given by

$$\Phi^\alpha : \mathbb{R} \times \mathrm{T}^*Q \to \mathrm{T}^*Q, \qquad \Phi^\alpha(t, \xi) := \xi + t\alpha\big(\pi(\xi)\big).$$

The flow induces a vertical affine transformation of T^*Q (fibre translation) by

$$\Phi_\alpha := \Phi^\alpha_1 : \mathrm{T}^*Q \to \mathrm{T}^*Q, \qquad \Phi_\alpha(\xi) = \xi + \alpha\big(\pi(\xi)\big).$$

Since fibrewise addition is commutative, for arbitrary 1-forms α and β one has

$$[\hat{\alpha}, \hat{\beta}] = 0, \qquad \Phi_\alpha \circ \Phi_\beta = \Phi_\beta \circ \Phi_\alpha = \Phi_{\beta+\alpha}.$$

One says that T^*Q acts on itself fibrewise transitively.

Proposition 8.3.11 *For every $\alpha \in \Omega^1(M)$,*

$$\Phi_\alpha^* \theta = \theta + \pi^*\alpha, \qquad \mathscr{L}_{\hat{\alpha}}\theta = \pi^*\alpha, \qquad \hat{\alpha} \lrcorner \omega = \pi^*\alpha. \tag{8.3.12}$$

Proof We have

$$\big(\Phi_\alpha^*\theta\big)_\xi(X) = \theta_{\Phi_\alpha(\xi)}\big(\Phi_\alpha' X\big) = \big\langle \xi + \alpha\big(\pi(\xi)\big), \pi' X\big\rangle = \big(\theta + \pi^*\alpha\big)_\xi(X)$$

and hence

$$\mathscr{L}_{\hat{\alpha}}\theta = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction t=0} \Phi_{t\alpha}^*\theta = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction t=0} \big(\theta + \pi^*(t\alpha)\big) = \pi^*\alpha.$$

Finally, since $\hat{\alpha}$ is vertical, $\pi^*\alpha = \mathscr{L}_{\hat{\alpha}}\theta = \hat{\alpha} \lrcorner \mathrm{d}\theta + \mathrm{d}(\hat{\alpha} \lrcorner \theta) = \hat{\alpha} \lrcorner \omega$. □

Now, let α be closed. Then, the first equation in (8.3.12) implies that Φ_α is a symplectomorphism. Moreover, by Example 8.3.8/2, the image of α is a Lagrangian submanifold of T^*Q. Along this submanifold, the tangent spaces of T^*Q decompose as

$$\mathrm{T}_{\alpha(x)}\big(\mathrm{T}^*Q\big) = \mathrm{T}_{\alpha(x)}\big(\mathrm{T}^*_x Q\big) \oplus \alpha'(\mathrm{T}_x Q) \tag{8.3.13}$$

into complementary Lagrangian subspaces. In applications, the special case of the zero 1-form is important. Its image is the zero section in T^*Q, that is, in this case (8.3.13) is a decomposition of the tangent bundle into complementary Lagrangian subspaces along the zero section. The decomposition (8.3.13) induces an identification of the canonical symplectic structure on T^*Q with the canonical symplectic structure on $\mathrm{T}_x Q \oplus \mathrm{T}^*_x Q$, given by (7.1.5).

Proposition 8.3.12 *For every $x \in Q$, the mapping*

$$\mathrm{T}_x Q \oplus \mathrm{T}^*_x Q \to \mathrm{T}_{\alpha(x)}\big(\mathrm{T}^*Q\big), \qquad (X, \eta) \mapsto \alpha'(X) + \hat{\eta}_{\alpha(x)}, \tag{8.3.14}$$

is a symplectomorphism.

Proof Clearly, the mapping (8.3.14) is a vector space isomorphism. Thus, it is enough to show that it is symplectic: by the last relation in (8.3.12), for $\sigma \in \Omega^1(Q)$ we find

$$\omega_{\alpha(x)}\big(\alpha'(X), \hat{\sigma}_{\alpha(x)}\big) = -(\hat{\sigma} \lrcorner \omega)_{\alpha(x)}\big(\alpha'(X)\big) = -\big(\pi^*\sigma\big)_{\alpha(x)}\big(\alpha'(X)\big) = -\sigma_x(X).$$

Thus, since the subspaces in the decomposition (8.3.13) are Lagrangian, we obtain

$$\omega\big(\alpha'(X) + \hat{\tau}_{\alpha(x)}, \alpha'(Y) + \hat{\sigma}_{\alpha(x)}\big) = \tau_x(Y) - \sigma_x(X) = \omega_{W\oplus W^*}\big((X, \tau_x), (Y, \sigma_x)\big),$$

with $W = \mathrm{T}_x Q$. □

Remark 8.3.13 To see how Proposition 8.3.12 is related to Proposition 7.2.9, we identify $\mathrm{T}_{\alpha(x)}(\mathrm{T}^*_x Q)$ with $\mathrm{T}^*_x Q$ and set $W = \mathrm{T}^*_x Q$ and $W' = \alpha'(\mathrm{T}_x Q)$. Then, the isomorphism $\chi : W' \to W^*$ defined by (7.2.7) is given by

$$\chi\big(\alpha'(X)\big) = -X.$$

This can be read off immediately from the above proof, using $\hat{\sigma}_{\alpha(x)} = \widehat{\sigma_x}$. Note that the roles of W and W^* have been interchanged here.

Exercises

8.3.1 Prove Remark 8.3.3/1.
8.3.2 Verify Formulae (8.3.7) and (8.3.8).
8.3.3 In Example 8.3.8/4, verify the formula $j^*\theta = \pi_N^* \mathrm{d}S$.

8.4 Coadjoint Orbits

Let G be a Lie group, let $\mathfrak{g}$ be its Lie algebra and let Ad^* denote the coadjoint representation of G on the dual space $\mathfrak{g}^*$. By a coadjoint orbit of G one means an orbit of this representation, viewed as an action of G on $\mathfrak{g}^*$. According to Corollary 6.2.9, coadjoint orbits are initial submanifolds of $\mathfrak{g}^*$. If G is compact, Corollary 6.3.5 implies that they are also closed and embedded.

In this section, we will show that every coadjoint orbit $\mathscr{O}$ of G carries a natural symplectic structure and that this structure makes it into a symplectic leaf of the natural Lie-Poisson structure of $\mathfrak{g}^*$. Since G acts transitively on $\mathscr{O}$, it is enough to define the desired symplectic form on the Killing vector fields of Ad^*. According to (6.2.3), the value at $\mu \in \mathscr{O}$ of the Killing vector field generated by $A \in \mathfrak{g}$ is given by

$$A_*(\mu) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \mathrm{Ad}^*\big(\exp(tA)\big)\mu = \mathrm{ad}^*(A)\mu. \tag{8.4.1}$$

We read off that $A_*(\mu) = 0$ iff $\langle \mu, [A, B]\rangle = 0$ for all $B \in \mathfrak{g}$. Thus, the following 2-forms on $\mathscr{O}$ are well-defined:

$$\omega^\pm_\mu(A_*, B_*) := \pm\big\langle \mu, [A, B]\big\rangle, \quad A, B \in \mathfrak{g},\ \mu \in \mathscr{O}. \tag{8.4.2}$$

These forms will be called the positive and the negative Kirillov form, respectively.

Theorem 8.4.1 (Kirillov) *The 2-forms $\omega^\pm$ are symplectic and G-invariant.*

This means that $(\mathscr{O}, \omega^\pm, \mathrm{Ad}^*)$ are symplectic G-manifolds, cf. Remark 6.1.3 and Definition 8.6.2.

Proof We give the proof for $\omega = \omega^+$. First, we prove G-invariance. Using Proposition 6.2.2/1, for $\mu \in \mathscr{O}$ and $A, B \in \mathfrak{g}$ we compute

$$\begin{aligned}
\big(\big(\mathrm{Ad}^*(a)\big)^*\omega\big)_\mu(A_*, B_*) &= \omega_{\mathrm{Ad}^*(a)\mu}\big(\big(\mathrm{Ad}^*(a)\big)_* A_*, \big(\mathrm{Ad}^*(a)\big)_* B_*\big)\\
&= \omega_{\mathrm{Ad}^*(a)\mu}\big(\big(\mathrm{Ad}(a)A\big)_*, \big(\mathrm{Ad}(a)B\big)_*\big)\\
&= \big\langle \mathrm{Ad}^*(a)\mu, \big[\mathrm{Ad}(a)A, \mathrm{Ad}(a)B\big]\big\rangle\\
&= \big\langle \mu, [A, B]\big\rangle\\
&= \omega_\mu(A_*, B_*).
\end{aligned}$$

Next, we show that ω is closed: According to Proposition 4.1.6,

$$\begin{aligned}
\mathrm{d}\omega(A_*, B_*, C_*) =& A_*\big(\omega(B_*, C_*)\big) - B_*\big(\omega(A_*, C_*)\big) + C_*\big(\omega(A_*, B_*)\big)\\
&- \omega\big([A_*, B_*], C_*\big) + \omega\big([A_*, C_*], B_*\big) - \omega\big([B_*, C_*], A_*\big).
\end{aligned}$$

By G-invariance, $\mathscr{L}_{A_*}\omega = \mathscr{L}_{B_*}\omega = \mathscr{L}_{C_*}\omega = 0$. Hence, Propositions 3.3.2 and 3.3.3/3 yield

$$A_*\big(\omega(B_*, C_*)\big) = \mathscr{L}_{A_*}\big(\omega(B_*, C_*)\big) = \omega\big([A_*, B_*], C_*\big) + \omega\big(B_*, [A_*, C_*]\big),$$

and analogous formulae for $B_*(\omega(A_*, C_*))$ and $C_*(\omega(A_*, B_*))$. Using this, Proposition 6.2.2/2 and the Jacobi identity, we obtain

$$\begin{aligned}(\mathrm{d}\omega)_\mu(A_*, B_*, C_*) &= \omega_\mu\big(B_*, [A_*, C_*]\big) - \omega_\mu\big(A_*, [B_*, C_*]\big) + \omega_\mu\big(C_*, [B_*, A_*]\big)\\ &= \big\langle\mu, \big[B, [C, A]\big]\big\rangle + \big\langle\mu, \big[A, [B, C]\big]\big\rangle + \big\langle\mu, \big[C, [A, B]\big]\big\rangle\\ &= 0.\end{aligned}$$

Thus, ω is closed, indeed. It remains to show that it is non-degenerate. Let $\mu \in \mathscr{O}$ and $A \in \mathfrak{g}$ such that $\omega_\mu(A_*, B_*) = 0$ for all $B \in \mathfrak{g}$. Since

$$\omega_\mu(A_*, B_*) = \big\langle\mu, [A, B]\big\rangle = \big\langle\mathrm{ad}^*(A)\mu, B\big\rangle,$$

this implies $\mathrm{ad}^*(A)\mu = 0$ and hence, by (8.4.1), $A_*(\mu) = 0$. □

Now, recall from Example 8.2.18/3 that $\mathfrak{g}^*$ carries a natural linear Poisson structure, namely the Lie-Poisson structure, whose Poisson tensor and Poisson bracket are given by

$$\Pi_\mu(A, B) = \big\langle\mu, [A, B]\big\rangle, \qquad \{f, g\}(\mu) = \big\langle\mu, \big[\mathrm{d}f(\mu), \mathrm{d}g(\mu)\big]\big\rangle, \tag{8.4.3}$$

respectively. Here, via the identifications $\mathrm{T}^*_\mu\mathfrak{g}^* \cong \mathfrak{g}^{**} \cong \mathfrak{g}$, the differential $\mathrm{d}f(\mu)$ is viewed as an element of $\mathfrak{g}$. We will show that the symplectic leaves of this Poisson structure coincide with the connected components of the coadjoint orbits of G and that the symplectic structure induced on a leaf coincides with the Kirillov symplectic structure (8.4.2). For that purpose, for every $A \in \mathfrak{g}$ we define a function $f_A \in C^\infty(\mathfrak{g}^*)$ by

$$f_A(\mu) := -\langle\mu, A\rangle.$$

Lemma 8.4.2 *Let* $A \in \mathfrak{g}$.

1. *The differential of* f_A *is given by* $\mathrm{d}f_A = -A$.
2. *The Hamiltonian vector field generated by* f_A *is given by* $X_{f_A} = A_*$.

Proof To determine $\mathrm{d}f_A(\mu)$ for $\mu \in \mathfrak{g}^*$, we choose a tangent vector $\sigma \in \mathrm{T}_\mu\mathfrak{g}^* \cong \mathfrak{g}^*$, represent it by the curve $t \mapsto \mu + t\sigma$ and compute

$$\big\langle\sigma, \mathrm{d}f_A(\mu)\big\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\Big\restriction_0 f_A(\mu + t\sigma) = -\frac{\mathrm{d}}{\mathrm{d}t}\Big\restriction_0 \langle\mu + t\sigma, A\rangle = -\langle\sigma, A\rangle.$$

This yields point 1. Using this, for $h \in C^\infty(\mathfrak{g}^*)$ we obtain

$$(X_{f_A})_\mu(h) = \{f_A, h\}(\mu) = \big\langle\mu, \big[\mathrm{d}f_A(\mu), \mathrm{d}h(\mu)\big]\big\rangle = -\big\langle\mu, \big[A, \mathrm{d}h(\mu)\big]\big\rangle.$$

By (8.4.1), the right hand side equals $(A_*)_\mu(h)$. This yields point 2. □

Proposition 8.4.3 *The symplectic leaves of the Poisson structure* (8.4.3) *coincide with the connected components of the coadjoint orbits and the symplectic form induced on each leaf coincides with the positive Kirillov form* (8.4.2).

Proof Since the value of the Hamiltonian vector field X_f at $\mu \in \mathfrak{g}^*$ depends on $\mathrm{d}f(\mu)$ only, the characteristic distribution of the Poisson structure Π is spanned by the Hamiltonian vector fields X_{f_A}, $A \in \mathfrak{g}$. Thus, by Lemma 8.4.2, this distribution coincides with the distribution spanned by the Killing vector fields of Ad^*, so that Theorems 6.2.8 and 8.2.20 yield that the symplectic leaves coincide with the connected components of the coadjoint orbits. To prove the second statement, let $\mathscr{O}$ be a connected component of a coadjoint orbit and let $\omega_{\mathscr{O}}$ denote the symplectic form induced on $\mathscr{O}$ by the Poisson structure, cf. (8.2.20). Using this, Lemma 8.4.2/2 and (8.4.3), for $\mu \in \mathscr{O}$ and $A, B \in \mathfrak{g}$, we obtain

$$\omega_{\mathscr{O}}(A_*, B_*)(\mu) = \omega_{\mathscr{O}}(X_{f_A}, X_{f_B})(\mu) = \{f_A, f_B\}(\mu) = \big\langle \mu, \big[\mathrm{d}f_A(\mu), \mathrm{d}f_B(\mu)\big]\big\rangle.$$

By Lemma 8.4.2/1, the right hand side equals $\langle \mu, [A, B]\rangle = \omega^+_\mu(A_*, B_*)$. □

Remark 8.4.4 Recall from Remark 5.4.11/2 that in the special case when G is semisimple, the Killing form k defines an equivariant isomorphism $F : \mathfrak{g} \to \mathfrak{g}^*$ by

$$\big\langle F(A), B\big\rangle = \mathsf{k}(A, B), \qquad A, B \in \mathfrak{g}. \tag{8.4.4}$$

By equivariance, F maps adjoint orbits $\tilde{\mathscr{O}}$ onto coadjoint orbits $\mathscr{O}$. Since $\tilde{\mathscr{O}}$ and $\mathscr{O}$ are initial submanifolds, the restriction $F : \tilde{\mathscr{O}} \to \mathscr{O}$, denoted by the same letter, is a diffeomorphism. Then,

$$\tilde{\omega}^{\pm} := F^*\omega^{\pm}$$

are Ad-invariant symplectic forms on $\tilde{\mathscr{O}}$. To compute them, let $A, B, C \in \mathfrak{g}$ and let $\tilde{B}_*$, $\tilde{C}_*$ denote the Killing vector fields generated by B and C under the adjoint representation. By (6.2.2), we have $F_*\tilde{B}_* = B_*$. Using this, we find

$$\tilde{\omega}^{\pm}_A(\tilde{B}_*, \tilde{C}_*) = \big(F^*\omega^{\pm}\big)_A(\tilde{B}_*, \tilde{C}_*) = \omega^{\pm}_{F(A)}(B_*, C_*) = \pm\big\langle F(A), [B, C]\big\rangle$$

and hence

$$\tilde{\omega}^{\pm}_A(\tilde{B}_*, \tilde{C}_*) = \pm\mathsf{k}\big(A, [B, C]\big). \tag{8.4.5}$$

If G is in addition linear and simple, then

$$\tilde{\omega}^{\pm}_A(\tilde{B}_*, \tilde{C}_*) = \pm c\,\mathrm{tr}\big(A[B, C]\big), \tag{8.4.6}$$

where the factor c is given in Example 5.4.12.

We conclude this section with a bunch of examples. In all of them, G is semisimple, so that it is enough to consider the orbits of the adjoint action.

Example 8.4.5

1. $G = \mathrm{SU}(2)$: Since every element of $\mathfrak{su}(2)$ can be diagonalized by means of an SU(2)-matrix, every orbit of the adjoint action of SU(2) contains a diagonal element. Since $\mathrm{tr}(A) = 0$, the eigenvalues are of the form $\mathrm{i}\lambda$ and $-\mathrm{i}\lambda$ with $\lambda \in \mathbb{R}$. Since the set of eigenvalues is invariant under Ad, the orbits are labelled by $\lambda \geq 0$ and the orbit corresponding to λ is given by

$$\big\{a \,\mathrm{diag}(\mathrm{i}\lambda, -\mathrm{i}\lambda)a^{\dagger} : a \in \mathrm{SU}(2)\big\}.$$

According to the Orbit Theorem 6.2.8, it is diffeomorphic to the homogeneous space of the right cosets in SU(2) of the stabilizer of $\mathrm{diag}(\mathrm{i}\lambda, -\mathrm{i}\lambda)$ under Ad. There are two distinct types of orbits. If $\lambda > 0$, the stabilizer consists of the diagonal matrices $a = \mathrm{diag}(\alpha, \overline{\alpha})$ with $\alpha \in \mathrm{U}(1)$. Hence, in this case, the orbit is diffeomorphic to the homogeneous space $\mathrm{SU}(2)/\mathrm{U}(1) \equiv \mathrm{U}(2)/(\mathrm{U}(1) \times \mathrm{U}(1))$, which according to Example 5.7.6 is diffeomorphic to $\mathbb{C}\mathrm{P}^1 = \mathrm{S}^2$. In case $\lambda = 0$, the stabilizer is SU(2) and the orbit consists of the origin alone.

That the nonzero adjoint orbits are 2-spheres can be seen alternatively by the following more explicit argument. Recall from Example 5.4.7 that the basis $\{\mathrm{I}^{\mathbb{C}}_1, \mathrm{I}^{\mathbb{C}}_2, \mathrm{I}^{\mathbb{C}}_3\}$ in $\mathfrak{su}(2)$, given in Example 5.2.8, defines a vector space isomorphism onto $\mathbb{R}^3$ which is equivariant with respect to the representation of SU(2) on $\mathbb{R}^3$ induced by the covering homomorphism $\mathrm{SU}(2) \to \mathrm{SO}(3)$ of Example 5.1.11. Thus, this isomorphism maps the nonzero adjoint orbits to the spheres S^2_r of radius $r > 0$ in $\mathbb{R}^3$, indeed. We calculate the symplectic form for the nonzero adjoint orbits: since $[\mathrm{I}^{\mathbb{C}}_i, \mathrm{I}^{\mathbb{C}}_j] = \varepsilon_{ij}{}^k \mathrm{I}^{\mathbb{C}}_k$, for $A, B, C \in \mathfrak{su}(2)$, Formula (8.4.5) yields

$$\tilde{\omega}^{\pm}_A(\tilde{B}_*, \tilde{C}_*) = \pm A^i B^j C^n \mathsf{k}\big(\mathrm{I}^{\mathbb{C}}_i, \big[\mathrm{I}^{\mathbb{C}}_j, \mathrm{I}^{\mathbb{C}}_n\big]\big) = \pm A^i B^j C^n \varepsilon_{jn}{}^l \mathsf{k}\big(\mathrm{I}^{\mathbb{C}}_i, \mathrm{I}^{\mathbb{C}}_l\big)$$

and with $\mathsf{k}(\mathrm{I}^{\mathbb{C}}_i, \mathrm{I}^{\mathbb{C}}_l) = 4\,\mathrm{tr}(\mathrm{I}^{\mathbb{C}}_i \mathrm{I}^{\mathbb{C}}_l) = -2\delta_{il}$ we obtain

$$\tilde{\omega}^{\pm}_A(\tilde{B}_*, \tilde{C}_*) = \mp 2 A^i B^j C^k \varepsilon_{ijk}.$$

Since the tangent vector $\mathbf{y}$ of S^2_r at $\mathbf{x} \in \mathrm{S}^2_r$ is the value at $\mathbf{x}$ of the Killing vector field generated by the Lie algebra element $\frac{\mathbf{x} \times \mathbf{y}}{r^2}$, we read off that via the isomorphism $\mathfrak{so}(3) \cong \mathbb{R}^3$, the Kirillov form $\tilde{\omega}^+$ gets identified with the scaled natural volume (or area) form on S^2_r:

$$\omega_{\mathbf{x}}(\mathbf{y}, \mathbf{z}) = -2\frac{\mathbf{x} \cdot (\mathbf{y} \times \mathbf{z})}{r^2}, \quad \mathbf{x} \in \mathrm{S}^2_r, \ \mathbf{y}, \mathbf{z} \perp \mathbf{x}. \tag{8.4.7}$$

2. $G = \mathrm{SU}(3)$: SU(3) is dealt with analogously to SU(2). The eigenvalues are $\mathrm{i}\lambda_1, \mathrm{i}\lambda_2, \mathrm{i}\lambda_3$ with $\lambda_k \in \mathbb{R}$ and $\lambda_1 + \lambda_2 + \lambda_3 = 0$. The orbits are given by

$$\big\{a \,\mathrm{diag}(\mathrm{i}\lambda_1, \mathrm{i}\lambda_2, \mathrm{i}\lambda_3)a^{\dagger} : a \in \mathrm{SU}(3)\big\}.$$

There are three types of orbits. If the λ_k are pairwise distinct, the stabilizer consists of the diagonal matrices $a = \mathrm{diag}(\alpha, \beta, \overline{\alpha\beta})$ with $\alpha, \beta \in \mathrm{U}(1)$. Hence, the orbit is diffeomorphic to the homogeneous space

$$\mathrm{SU}(3)/\big(\mathrm{U}(1)\times\mathrm{U}(1)\big)\equiv\mathrm{U}(3)/\big(\mathrm{U}(1)\times\mathrm{U}(1)\times\mathrm{U}(1)\big).$$

This is a flag manifold of dimension 6, cf. Example 5.7.7. If $\lambda_1=\lambda_2\neq\lambda_3$, the stabilizer consists of matrices of the form

$$\left[\begin{array}{c|c} b & \begin{matrix}0\\0\end{matrix} \\ \hline 0\,0 & \overline{\det b}\end{array}\right],\qquad b\in\mathrm{U}(2),$$

and the orbit is thus diffeomorphic to the Graßmann manifold

$$\mathrm{SU}(3)/\mathrm{U}(2)\equiv\mathrm{U}(3)/\big(\mathrm{U}(1)\times\mathrm{U}(2)\big),$$

which, in turn, is diffeomorphic to the complex projective space $\mathbb{C}\mathrm{P}^2$, cf. Example 5.7.6. Finally, if the eigenvalues λ_k are all equal, they must vanish. Then, the stabilizer is SU(3) and the orbit consists of just the origin. The computation of the Kirillov symplectic form is left to the reader (Exercise 8.4.1).

3. $G=\mathrm{SO}(3)$: As for SU(2), according to Example 5.4.7, the basis $\{\mathrm{I}_1^{\mathbb{R}},\mathrm{I}_2^{\mathbb{R}},\mathrm{I}_3^{\mathbb{R}}\}$ in $\mathfrak{so}(3)$ given there defines an isomorphism between the adjoint representation and the identical representation of SO(3). This isomorphism identifies the nonzero adjoint orbits of SO(3) with the spheres S^2_r of radius $r>0$ in $\mathbb{R}^3$. A calculation similar to that of point 1 shows that the Kirillov form $\tilde{\omega}$ on an adjoint orbit coincides via this isomorphism with the 2-form (8.4.7).
4. $G=\mathrm{SO}(4)$: This Lie group is semisimple but not simple. Recall from Example 5.4.7 that the Lie algebra isomorphism $\mathrm{d}\phi:\mathfrak{su}(2)\oplus\mathfrak{su}(2)\to\mathfrak{so}(4)$ induced by the covering homomorphism $\phi:\mathrm{SU}(2)\times\mathrm{SU}(2)\to\mathrm{SO}(4)$ of Example 5.1.11 is an isomorphism of representations of $\mathrm{SU}(2)\times\mathrm{SU}(2)$, where the representation on $\mathfrak{so}(4)$ is induced via ϕ by the adjoint representation of SO(4). Thus, $\mathrm{d}\phi$ identifies the adjoint orbits of SO(4) with those of $\mathrm{SU}(2)\times\mathrm{SU}(2)$. According to point 1, the following types of orbits occur:

$$\mathscr{O}_{(A,B)}\cong\mathrm{S}^2\times\mathrm{S}^2,\qquad\mathscr{O}_{(A,0)}\cong\mathscr{O}_{(0,A)}\cong\mathrm{S}^2,\qquad\mathscr{O}_{(0,0)}=\big\{(0,0)\big\}\tag{8.4.8}$$

with $A,B\in\mathfrak{su}(2)$, $A\neq0$, $B\neq0$. As explained there, under the Lie algebra isomorphism $\mathfrak{su}(2)\cong\mathbb{R}^3$ of Example 5.2.8, orbits of the second type are identified with spheres of fixed radius in $\mathbb{R}^3$ with the symplectic form given by (8.4.7) and orbits of the first type are identified with products thereof.

Exercises

8.4.1 Determine the Kirillov symplectic form for the coadjoint orbits of SU(3), cf. Example 8.4.5/2.

8.4.2 The Euclidean group E(3) in three dimensions is defined as the semidirect product of SO(3) with the group of translations, that is, $\mathrm{E}(3):=\mathrm{SO}(3)\ltimes\mathbb{R}^3$ with multiplication

$$(a_1,\mathbf{b}_1)(a_2,\mathbf{b}_2):=(a_1a_2,\mathbf{b}_1+a_1\mathbf{b}_2).$$

Determine the Lie algebra of E(3) and show that under the isomorphism $\mathfrak{so}(3) \cong \mathbb{R}^3$ of Example 5.2.8, the coadjoint orbits are given by

$$\mathscr{O}_{(c,d)} = \left\{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 : \mathbf{y}^2 = c^2, \mathbf{x} \cdot \mathbf{y} = cd\right\}, \quad c \geq 0,\ d \in \mathbb{R}.$$

8.4.3 For the Lie group G of real upper triangular $(n \times n)$-matrices with unit determinant, prove the following.
(a) The Lie algebra $\mathfrak{g}$ of G consists of the real upper triangular matrices with trace 0.
(b) By means of the Ad-invariant scalar product $\langle A, B\rangle = \mathrm{tr}(AB)$, $\mathfrak{g}^*$ can be identified with the space of real lower triangular matrices with trace 0.
(c) Under this identification, the coadjoint action takes the form

$$\mathrm{Ad}^*(g)A = \left(gAg^{-1}\right)_-,$$

with the minus sign meaning that all entries above the main diagonal are set to zero. Find the explicit matrix representation for elements of the orbit through

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}.$$

8.5 Coisotropic Submanifolds and Contact Structures

Coisotropic submanifolds play an important role in symplectic reduction, in the theory of integrable systems and especially in Hamilton-Jacobi theory. The results to be derived in this section apply in particular to Lagrangian submanifolds.

Let (M, ω) be a symplectic manifold and let (N, φ) be a coisotropic submanifold of M. For simplicity, we assume that N is given by a subset of M and that φ is the natural inclusion mapping, cf. Remark 1.6.2/1. By definition of coisotropy, we have $(\mathrm{T}N)^\omega \subset \mathrm{T}N$. This implies that $(\mathrm{T}_pN)^\omega$ coincides with the characteristic subspace $F_p^{\omega_N}$ of the 2-form $\omega_N = \varphi^*\omega$, see Definition 4.2.18: indeed, since $\mathrm{d}\omega_N = 0$, Formula (8.1.5) yields

$$\ker(\omega_N)_p = \mathrm{T}_pN \cap (\mathrm{T}_pN)^\omega = (\mathrm{T}_pN)^\omega. \tag{8.5.1}$$

Moreover, by point 2 of Proposition 7.2.1, the dimension of $(\mathrm{T}_pN)^\omega$ is equal to the codimension of N and hence independent of p. This shows that the characteristic distribution D^{ω_N} of ω_N is regular and that $(\mathrm{T}N)^\omega$ coincides with this distribution.

Definition 8.5.1 The distribution $D^{\omega_N} = (\mathrm{T}N)^{\omega}$ is called the characteristic distribution of N. Sections of D^{ω_N} are called characteristic vector fields.

Note that D^{ω_N} is isotropic as a vector subbundle of $\mathrm{T}M$ and that $(D^{\omega_N})^{\omega} = \mathrm{T}N$. Moreover, Proposition 4.2.20 yields

Corollary 8.5.2 *The characteristic distribution D^{ω_N} of a coisotropic submanifold is integrable.*

Integral submanifolds of D^{ω_N} are called characteristics of N.

Proposition 8.5.3 *Let N be an initial coisotropic submanifold of the symplectic manifold (M, ω) and let L be a Lagrangian submanifold of M contained in N. Then, L is a union of characteristics of N.*[7]

Proof By Proposition 1.6.14, L is a submanifold of N. Hence, $\mathrm{T}L \subset (\mathrm{T}N)_{\restriction L}$ and thus

$$\left(D^{\omega_N}\right)_{\restriction L} = \left((\mathrm{T}N)_{\restriction L}\right)^{\omega} \subset (\mathrm{T}L)^{\omega} = \mathrm{T}L.$$

It follows that, through every point of L, there exists a characteristic of N which is contained in L. □

Lemma 8.5.4 *Let (M, ω) be a symplectic manifold and let $N \subset M$ be a coisotropic submanifold. A Hamiltonian vector field X_f on M restricts to*

1. *a vector field on N iff f is constant on the characteristics of N,*
2. *a characteristic vector field on N iff f is locally constant on N.*

Proof For all $m \in N$ and $Y \in \mathrm{T}_m N$, we have

$$\omega_m\left((X_f)_m, Y\right) = -Y(f).$$

First, this implies that $(X_f)_{\restriction N}$ takes values in $\mathrm{T}N = (D^{\omega_N})^{\omega}$ iff $Y(f) = 0$ for all $Y \in D^{\omega_N}$, that is, iff f is constant along the characteristics of N. Second, this implies that $(X_f)_{\restriction N}$ takes values in $D^{\omega_N} = (\mathrm{T}N)^{\omega}$ iff $Y(f) = 0$ for all $Y \in \mathrm{T}N$, that is, iff f is locally constant on N. □

We recall that in Sect. 2.7 we had characterized the tangent bundle of an embedded submanifold $N \subset M$ and its annihilator by means of the ideal $C^{\infty}_N(M)$ of smooth functions on M, vanishing on N:

$$\mathrm{T}_p N = \left\{X \in \mathrm{T}_p M \colon X(f) = 0 \text{ for all } f \in C^{\infty}_N(M)\right\}, \tag{8.5.2}$$

[7] Note that the statement of the proposition is not about the maximal integral manifolds of D^{ω_N}.

$$(\mathrm{T}_p N)^0 = \{\alpha \in \mathrm{T}^*_p M : \alpha = \mathrm{d}f(p) \text{ for some } f \in C^\infty_N(M)\}. \tag{8.5.3}$$

According to Proposition 7.2.1/1, the isomorphism $\omega^\flat : \mathrm{T}M \to \mathrm{T}^*M$ yields an identification of the symplectic normal bundle $(\mathrm{T}N)^\omega \subset \mathrm{T}M_{\restriction N}$ with the annihilator $(\mathrm{T}N)^0$ of $\mathrm{T}N$ in $\mathrm{T}^*M_{\restriction N}$. Since this isomorphism identifies X_f with $-\mathrm{d}f$, we get the following characterization of the normal bundle:

$$(\mathrm{T}_p N)^\omega = \{X \in \mathrm{T}_p M : X = X_f(p) \text{ for some } f \in C^\infty_N(M)\}. \tag{8.5.4}$$

These observations yield the following criteria for coisotropy [311]:

Proposition 8.5.5 *For an embedded submanifold N of a symplectic manifold (M, ω), the following statements are equivalent*:

1. *N is coisotropic.*
2. *For every $f \in C^\infty_N(M)$, the Hamiltonian vector field X_f is tangent to N.*
3. *$C^\infty_N(M)$ is a Poisson subalgebra of $C^\infty(M)$.*

Proof The submanifold N is coisotropic iff $(\mathrm{T}N)^\omega \subset \mathrm{T}N$, that is, according to (8.5.4), iff X_f is tangent to N for all $f \in C^\infty_N(M)$. Thus, points 1 and 2 are equivalent. By (8.5.2), point 2 is equivalent to the requirement that $X_f(g)_{\restriction N} = \{f, g\}_{\restriction N} = 0$ for all $f, g \in C^\infty_N(M)$. This is equivalent to point 3. □

Now, let $F : M \to \mathbb{R}^r$ be a smooth mapping for which 0 is a regular value and let $N = F^{-1}(0)$. In this context, one has the following criterion for coisotropy. Let F_a denote the components of F and recall from the Level Set Theorem that for every $m \in N$, one has

$$\mathrm{T}_m N = \ker F'_m. \tag{8.5.5}$$

Proposition 8.5.6 *$N = F^{-1}(0)$ is coisotropic iff the functions F_a are in involution on N, that is,*

$$\{F_a, F_b\}_{\restriction N} = 0.$$

Proof By Proposition 8.5.5/3, if N is coisotropic, then $\{F_a, F_b\}_{\restriction N} = 0$. Conversely, if $\{F_a, F_b\}_{\restriction N} = 0$, then $\langle \mathrm{d}F_b, X_{F_a}\rangle(m) = 0$ for all $m \in N$. In view of (8.5.5), this yields that the vector fields X_{F_a} are tangent to N. On the other hand, (8.5.5) implies that the differentials $\mathrm{d}F_a$ span the annihilator $(\mathrm{T}N)^0$ and, consequently, that the Hamiltonian vector fields X_{F_a} span $(\mathrm{T}N)^\omega$. Thus, $(\mathrm{T}N)^\omega \subset \mathrm{T}N$. □

Remark 8.5.7

1. If we assume that $F : M \to \mathbb{R}^r$ is a submersion and that $\{F_a, F_b\} = 0$ holds on the whole of M, we obtain a foliation of M by coisotropic submanifolds (level sets) of codimension r. The special case $r = \frac{1}{2}\dim M$ yields a foliation by Lagrangian submanifolds. This is the situation of an integrable system. Chap. 11 is devoted to the study of such systems.

2. It is easy to show that all coisotropic submanifolds can be described locally as level sets in the sense of Proposition 8.5.6. More precisely, let $N \subset M$ be a coisotropic submanifold. Then, for every point $p \in N$, there exists a neighbourhood V of p in N, a neighbourhood U of p in M and a submersion $F : U \to \mathbb{R}^r$ such that $V = F^{-1}(0)$ and $\{F_a, F_b\}_{\restriction V} = 0$ (Exercise 8.5.1).

Example 8.5.8 Let $N \subset M$ be an embedded submanifold of M.

1. The restriction $\mathrm{T}^*M_{\restriction N}$ of T^*M to N is coisotropic. This has been shown before (Proposition 8.3.9).
2. Let $\pi : Q \to N$ be a surjective submersion. Then, $\mathrm{V}Q := \ker \pi'$ is a vertical subbundle of $\mathrm{T}Q$, cf. Example 2.7.7. Its annihilator $\mathrm{V}^0 Q$ is the union of the conormal bundles of the fibres of π. According to Example 8.3.8/3, these conormal bundles are Lagrangian submanifolds. Thus, Proposition 7.2.4/4 implies that $\mathrm{V}^0 Q$ is coisotropic.

By Proposition 7.2.4/2, a submanifold of codimension 1 of a symplectic manifold (M, ω) is always coisotropic. Such submanifolds play a prominent role in the theory of Hamiltonian systems and in Hamilton-Jacobi theory, see Chaps. 9 and 12. Often they carry the additional structure of a so-called contact manifold. To explain this, let us start by considering the following special case. Assume that (M, ω) is a $2n$-dimensional exact symplectic manifold, that is, there exists a global potential 1-form β such that $\omega = \mathrm{d}\beta$. Then, there exists a unique vector field Z on M such that

$$Z \lrcorner \omega = \beta. \tag{8.5.6}$$

Remark 8.5.9 The vector field Z defined by (8.5.6) fulfils

$$\mathscr{L}_Z \omega = \omega. \tag{8.5.7}$$

A vector field with this property is said to be of Liouville type, cf. Remark 8.3.3/3 for the cotangent bundle case.

Now, let (P, ι) be a hypersurface[8] in M which is transversal to Z. Then, the 1-form

$$\alpha := \iota^* \beta$$

on P has the property that $\alpha \wedge (\mathrm{d}\alpha)^n$ is a volume form on P, because

$$\beta \wedge (\mathrm{d}\beta)^{n-1} = (Z \lrcorner \omega) \wedge \omega^{n-1} = \frac{1}{n} Z \lrcorner \big(\omega^n\big) \tag{8.5.8}$$

and because P is transversal to Z. It follows that α is nonzero on the characteristic distribution $D^{\omega_P} = (\mathrm{T}P)^\omega = \ker \mathrm{d}\alpha$ and that $\mathrm{d}\alpha$ is non-degenerate on the hyperplane distribution[9] $\ker \alpha$. In view of the integrability criterion of Proposition 4.7.6,

[8] That is, an embedded submanifold of codimension 1.

[9] A hyperplane distribution on P is a regular distribution $E \subset \mathrm{T}P$ of codimension one.

the latter means that the hyperplane distribution $\ker\alpha$ is maximally non-integrable. Indeed, for any $X, Y \in \Gamma(\ker\alpha)$, we have

$$d\alpha(X, Y) = -\alpha\big([X, Y]\big).$$

One says that α is a contact form on P and calls the pair (P, α) an exact, or strict, contact structure.

Given the great importance contact geometry has acquired in recent years, especially in connection with symplectic topology, we will discuss the basics here. There is a huge literature where one can find detailed presentations, e.g. [24], [102], [103], [181] and [206], see also [57].

The abstract notion of contact manifold is defined without taking recourse to a symplectic manifold. Thus, let P be a manifold and let E be a hyperplane distribution on P. Motivated by the above special case, we take the property of E to be maximally non-integrable as a defining condition and we formulate this condition in terms of locally defining 1-forms. A 1-form α on $U \subset P$ is said to be locally defining for E if

$$\ker\alpha = E_{\restriction U}.$$

It is said to be globally defining for E if in addition $U = P$. Locally defining 1-forms exist in a neighbourhood of every point of P: consider the quotient vector bundle $\mathscr{L} := \mathrm{T}P/E$, called the characteristic line bundle of E. The natural projection $\mathrm{pr} : \mathrm{T}P \to \mathscr{L}$ is a vertical vector bundle morphism. Let $\mathrm{pr}^{\mathrm{T}} : \mathscr{L}^* \to \mathrm{T}^*P$ be the dual morphism. Every local section $\tilde{\alpha}$ of $\mathscr{L}^*$ over $U \subset P$ defines a local 1-form α on U by $\alpha = \mathrm{pr}^{\mathrm{T}} \circ \tilde{\alpha}$. By construction, $\alpha_{\restriction E} = 0$. Hence, α is locally defining for E iff $\tilde{\alpha}$ is nowhere vanishing.

Definition 8.5.10 (Contact manifold) Let P be a manifold.

1. A contact form on P is a 1-form α for which $(d\alpha)_{\restriction \ker\alpha}$ is non-degenerate.
2. A contact structure on P is a hyperplane distribution E with the property that any locally defining 1-form is a local contact form on P. The pair (P, E) is called a contact manifold.
3. (P, E) is said to be exact or strict if there exists a globally defining contact form.
4. A smooth mapping $\Phi : P_1 \to P_2$ of contact manifolds (P_1, E_1) and (P_2, E_2) is called a contact mapping if $\Phi'(E_1) = E_2$. If Φ is in addition a diffeomorphism, it is called a contactomorphism.

Remark 8.5.11

1. Let α be a contact form on P. Since it is a 1-form, the subspace $\ker\alpha_p \subset \mathrm{T}_pP$ has codimension 1 if $\alpha_p \neq 0$ and codimension 0 otherwise. Since $d\alpha$ is non-degenerate on $\ker\alpha_p$, Proposition 7.1.2 implies that $\ker\alpha_p$ has even dimension. Hence, the codimension must be the same for all $p \in P$. It follows that
   (a) α is nowhere vanishing,
   (b) $\ker\alpha$ is a hyperplane distribution on P and hence an exact contact structure,

(c) P has odd dimension.
Moreover, non-degeneracy implies $\ker(\mathrm{d}\alpha)_p \cap \ker\alpha_p = \{0\}$. Since a 2-form on an odd-dimensional manifold must have nontrivial kernel, it follows that $\ker(\mathrm{d}\alpha)$ is a vector bundle complement of $\ker\alpha$ in $\mathrm{T}P$. Thus, α defines a splitting

$$\mathrm{T}P = \ker\alpha \oplus \ker\mathrm{d}\alpha. \tag{8.5.9}$$

2. Let (P, E) be a contact structure. Then, point 1 yields that E has even dimension and P has odd dimension. Moreover, every locally defining contact form for E on $U \subset P$ defines a splitting of $\mathrm{T}P$ over U given by (8.5.9). This implies that over U, the characteristic line bundle $\mathscr{L}$ of E can be identified with $\ker(\mathrm{d}\alpha)$.
3. Let E be a hyperplane distribution on P. If α is a locally defining 1-form for E over U, then so is $f\alpha$ for every nowhere vanishing function $f \in C^\infty(U)$. Conversely, for any two locally defining 1-forms α_1, α_2 over U, there exists $f \in C^\infty(U)$, necessarily nowhere vanishing, such that $\alpha_2 = f\alpha_1$. It follows that E is a contact structure iff for every $p \in P$ there exists a locally defining contact form at p. Thus, contact structures on P correspond bijectively to equivalence classes of local contact forms on P under the equivalence relation $\alpha_1 \sim \alpha_2$ iff $\alpha_2 = f\alpha_1$ for some smooth function f on the common domain of α_1 and α_2.
4. By Remark 2.7.11/3, every hyperplane distribution E admits a vector bundle complement in $\mathrm{T}P$ and by identifying the characteristic line bundle $\mathscr{L}$ with such a complement one obtains a splitting

$$\mathrm{T}P \cong E \oplus \mathscr{L}. \tag{8.5.10}$$

For example, if we choose an auxiliary Riemannian metric g on P, we can identify $\mathscr{L}$ with the orthogonal complement $E^\perp$ of E in $\mathrm{T}P$ and thus realize the splitting (8.5.10) in the form $\mathrm{T}P = E \oplus E^\perp$. This way, every nowhere vanishing local section $\tilde{\alpha}$ of $\mathscr{L}^*$ corresponds via g to a nowhere vanishing local section Y of $E^\perp$ and the 1-form $\alpha = \mathrm{pr}^{\mathrm{T}} \circ \tilde{\alpha}$ is given in terms of Y by $\alpha = \mathrm{g}(Y, \cdot)$.

Proposition 8.5.12 *A* 1*-form* α *on* P *is a contact form iff* $\alpha \wedge (\mathrm{d}\alpha)^n$ *is a volume form on* P.

Proof First, assume that α is a contact form. Since $(\mathrm{d}\alpha)_{\restriction\ker\alpha}$ is non-degenerate, Proposition 7.1.7 yields that $(\mathrm{d}\alpha)^n_{\restriction\ker\alpha_p} \neq 0$ for all $p \in P$. The decomposition (8.5.9) implies $\alpha_{\restriction\ker(\mathrm{d}\alpha)_p} \neq 0$ and hence $\alpha_p \wedge (\mathrm{d}\alpha)^n_p \neq 0$ for all $p \in P$. Thus, $\alpha \wedge (\mathrm{d}\alpha)^n$ is a volume form. Conversely, assume that $\alpha \wedge (\mathrm{d}\alpha)^n$ is a volume form and let $p \in P$. By Proposition 2.7.5, we find a basis $\{e_1, \ldots, e_{2n}, f\}$ in T_pP such that $\ker\alpha_p$ is spanned by $e_1, \ldots, e_{2n}$. Then,

$$\alpha_p \wedge (\mathrm{d}\alpha)^n_p(e_1, \ldots, e_{2n}, f) = \alpha_p(f)(\mathrm{d}\alpha)^n_p(e_1, \ldots, e_{2n}) \neq 0$$

and hence $(\mathrm{d}\alpha)^n_p(e_1, \ldots, e_{2n}) \neq 0$. Thus, $(\mathrm{d}\alpha)_{\restriction\ker\alpha}$ is non-degenerate. □

Remark 8.5.13 In analogy to the special case discussed in the beginning, Proposition 8.5.12 states that a hyperplane distribution is a contact structure iff it is maximally non-integrable.

Proposition 8.5.12 implies

Corollary 8.5.14 *Let (P, E) be an exact contact manifold with globally defining contact form α. Then, $\alpha \wedge (\mathrm{d}\alpha)^n$ is a volume form on P. In particular, every exact contact manifold is orientable.*

More generally, one has the following result on the orientability of contact manifolds.

Proposition 8.5.15 *Let (P, E) be a $(2n+1)$-dimensional contact manifold.*

1. *If n is odd, P is orientable.*
2. *If n is even, P is orientable iff (P, E) is exact.*

Proof See Exercise 8.5.2 or [102]. □

Now, let us derive a criterion for exactness. Recall from Remark 2.7.11/4 that a hyperplane distribution E is called coorientable if the characteristic line bundle $\mathscr{L}$ is orientable. For dimensional reasons this means that it admits a global nowhere vanishing section and hence that it is trivial. Since every such section induces a globally defining 1-form for E, we have

Proposition 8.5.16 *A contact manifold (P, E) is exact iff E is coorientable, that is, iff the characteristic line bundle $\mathscr{L} = \mathrm{T}P/E$ is trivial.*

Remark 8.5.17 (Reeb vector field) Let (P, E) be an exact contact manifold and let α be a defining 1-form. Then, there exists a unique vector field R_α on P such that

$$R_\alpha \lrcorner \alpha = 1, \qquad R_\alpha \lrcorner \mathrm{d}\alpha = 0. \tag{8.5.11}$$

It is called the Reeb vector field associated with α. According to (8.5.11), it spans $\ker \mathrm{d}\alpha$ and is transversal to E. By (4.1.28) and (4.1.24), the flow Φ of R_α satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t} \Phi_t^* \alpha = \Phi_t^* \mathscr{L}_{R_\alpha} \alpha = \Phi_t^* \big(R_\alpha \lrcorner \mathrm{d}\alpha + \mathrm{d}(R_\alpha \lrcorner \alpha) \big) = 0,$$

that is, it leaves α and hence E invariant. Thus, R_α is an example of a contact vector field (a vector field whose flow preserves the contact structure E).

Example 8.5.18

1. Let $P = \mathbb{R}^{2n+1}$ with the standard coordinates $x^1, \dots, x^n, y^1, \dots, y^n$ and z. Then, the 1-form
$$\alpha := \mathrm{d}z + \sum_{i=1}^{n} x^i \mathrm{d}y^i \tag{8.5.12}$$
is a contact form, called the standard contact form on $\mathbb{R}^{2n+1}$ (Exercise 8.5.3).
2. Consider $\mathbb{R}^{2n+2}$ with the standard coordinates $x^1, \dots, x^{n+1}$ and $y^1, \dots, y^{n+1}$ and define
$$\alpha := \frac{1}{2} \sum_{i=1}^{n+1} \bigl(x^i \mathrm{d}y^i - y^i \mathrm{d}x^i\bigr). \tag{8.5.13}$$
Let $P = \mathrm{S}^{2n+1}$ and let $\iota : \mathrm{S}^{2n+1} \to \mathbb{R}^{2n+2}$ be the natural inclusion mapping. Then, $\iota^*\alpha$ is a contact form, called the standard contact form on S^{2n+1} (Exercise 8.5.4).
3. Let Q be a manifold of dimension n. The projective cotangent bundle of Q is defined by
$$\mathrm{P}^* Q := \bigl(\mathrm{T}^* Q \setminus s_0\bigr)/\sim, \tag{8.5.14}$$
where s_0 denotes the zero section in $\mathrm{T}^* Q$ and $\xi \sim \xi'$ iff $\xi = a\xi'$ for some $a \in \mathbb{R}$. This is a locally trivial fibre bundle over Q with typical fibre $\mathbb{R}\mathrm{P}^{n-1}$ and projection π induced from that of $\mathrm{T}^* Q$. For $[\xi] \in \mathrm{P}^* Q$, we take the mapping
$$\bigl(\pi'_{[\xi]}\bigr)^{\mathrm{T}} : \mathrm{T}^*_{\pi([\xi])} Q \to \mathrm{T}^*_{[\xi]}\bigl(\mathrm{P}^* Q\bigr)$$
and define
$$E_{[\xi]} := \ker\bigl\{\bigl(\pi'_{[\xi]}\bigr)^{\mathrm{T}}(\xi)\bigr\}, \tag{8.5.15}$$
with ξ being an arbitrary representative of $[\xi] \in \mathrm{P}^* Q$. This is a hyperplane distribution on $\mathrm{P}^* Q$. We leave it to the reader to prove that the canonical 1-form θ on $\mathrm{T}^* Q$ descends to a globally defining contact form for E (Exercise 8.5.5). Thus, $(\mathrm{P}^* Q, E)$ is an exact contact manifold.
4. Similarly, the cotangent sphere bundle of Q is defined by
$$\mathrm{S}^* Q := \bigl(\mathrm{T}^* Q \setminus s_0\bigr)/\sim, \tag{8.5.16}$$
where $\xi \sim \xi'$ iff $\xi = a\xi'$ for some positive real number a. This is a locally trivial fibre bundle over Q with typical fibre S^{n-1} and projection induced from that of $\mathrm{T}^* Q$. As in the previous example, Formula (8.5.15), with π interpreted as the projection of $\mathrm{S}^* Q$, defines an exact contact structure on $\mathrm{S}^* Q$ with contact form induced from the canonical 1-form θ on $\mathrm{T}^* Q$.
5. The following example plays an important role in applications. Given a $2n$-dimensional exact symplectic manifold (M, ω) with potential 1-form β, define

$P = M \times \mathbb{R}$ and[10]

$$\alpha := \mathrm{d}t - \beta, \tag{8.5.17}$$

where t denotes the standard coordinate on $\mathbb{R}$. Since

$$\alpha \wedge (\mathrm{d}\alpha)^n = (-1)^n \mathrm{d}t \wedge (\mathrm{d}\beta)^n$$

and since the right hand side is a volume form on P, α is a contact form on P.

6. As a consequence of point 2 of Proposition 8.5.15, one can use manifolds which are not orientable to construct contact manifolds which are not exact. For example, let $P = \mathbb{R}^{n+1} \times \mathbb{R}P^n$ with the standard coordinates $x^1, \ldots, x^{n+1}$ on $\mathbb{R}^{n+1}$ and the homogeneous coordinates $[y^1 : \ldots : y^{n+1}]$ on $\mathbb{R}P^n$ and define

$$E := \ker\left\{\sum_{i=1}^{n+1} y^i \mathrm{d}x^i\right\}.$$

One can check that E is a contact structure on P. From Example 4.2.5/3 we know that $\mathbb{R}P^n$ is not orientable for even n. Hence, in this case, E is not exact. For an exhaustive discussion of this contact structure, we refer the reader to [102], Example 2.14 and Proposition 2.15. Let us add that in the case where n is odd, P is orientable but E is not coorientable and, therefore, (P, E) is not exact, too.

Now, let us return to the discussion of hypersurfaces of symplectic manifolds. To begin with, we discuss the procedure of symplectization for an exact contact manifold (P, E). Choose a globally defining 1-form α, and endow $P \times \mathbb{R}$ with the exact 2-form

$$\omega_\alpha := \mathrm{d}\big(\mathrm{e}^t \alpha\big) = \mathrm{e}^t (\mathrm{d}t \wedge \alpha + \mathrm{d}\alpha), \tag{8.5.18}$$

where t denotes the standard coordinate on $\mathbb{R}$. By assumption, we have the decomposition (8.5.9). Since $\mathrm{d}\alpha$ is non-degenerate on $\ker \alpha = E$ and α is non-degenerate on $\ker(\mathrm{d}\alpha)$, and since $\mathrm{d}t$ is non-degenerate on $\mathbb{R}$, ω_α is non-degenerate and hence symplectic. The symplectic manifold $(P \times \mathbb{R}, \omega_\alpha)$ is referred to as a symplectization of (P, E). Since P embeds into $P \times \mathbb{R}$ as the hypersurface $P \times \{0\}$, we obtain

Proposition 8.5.19 *Every exact contact manifold can be embedded as a hypersurface in an exact symplectic manifold.*

Moreover, since every contact manifold is locally exact, by applying the procedure of symplectization locally, we obtain the following contact counterpart of the Darboux Theorem.

[10] We omit the natural projections to the factors of the direct product.

Proposition 8.5.20 *Let (P, E) be a $(2n+1)$-dimensional contact manifold. For every $p \in P$, there exist local coordinates $x^1, \ldots, x^n$, $y^1, \ldots, y^n$ and z at p such that*

$$\alpha = \sum_{i=1}^{n} x^i \mathrm{d}y^i + \mathrm{d}z$$

is a locally defining 1-form for E.

Example 8.5.21 For $P = \mathrm{S}^{2n-1}$ with the contact structure of Example 8.5.18/2, the manifold $\mathrm{S}^{2n-1} \times \mathbb{R}$, endowed with the symplectic form (8.5.18), is symplectomorphic to $\mathbb{R}^{2n} \setminus \{0\}$ with the standard symplectic structure (Exercise 8.5.6).

Remark 8.5.22

1. Let (P, E) be an exact contact manifold and let $\iota : P \to P \times \mathbb{R}$ be the natural embedding of P into its symplectization, given by $\iota(p) = (p, 0)$. Since $\iota^*\omega = \mathrm{d}\alpha$, the splitting (8.5.9) is a decomposition of $\mathrm{T}P$ into the non-integrable contact structure $E = \ker \alpha$ and the integrable characteristic distribution $D^{\omega_P} = \ker \mathrm{d}\alpha$ of the coisotropic submanifold P. Moreover, there exists a natural transversal Liouville vector field, namely $Z = \frac{\partial}{\partial t}$. Indeed, $Z \lrcorner \omega = \mathrm{e}^t \alpha$.
2. In the case where (P, E) is not necessarily exact, the symplectization is constructed as follows. Let s_0 denote the zero section of T^*P, let E^0 denote the annihilator of E in T^*P and let θ be the canonical 1-form on T^*P. Define

$$M := E^0 \setminus s_0, \qquad \beta := \iota^*\theta,$$

where $\iota : M \to \mathrm{T}^*P$ denotes the natural inclusion mapping. The action of the multiplicative group $\mathbb{R}_* := \mathbb{R} \setminus \{0\}$ on M by scalar multiplication turns M into a principal $\mathbb{R}_*$-bundle over P with projection π induced from the canonical projection of T^*P by restriction. Using that every locally defining 1-form α for E over U induces a local trivialization $P \times \mathbb{R}_* \to \pi^{-1}(U)$ by $(m, t) \mapsto t\alpha_m$, as well as Proposition 8.5.12, one can show that $(\mathrm{d}\beta)^n$ is a volume form. Hence, $\mathrm{d}\beta$ is a symplectic form on M (Exercise 8.5.7). Thus, in the general situation, symplectization yields a principal bundle over P whose bundle manifold M is symplectic, and P is only locally embedded into M as a hypersurface. Since locally defining 1-forms for E define local trivializations of M, in the case where (P, E) is exact with global contact form α, M is trivial and hence decomposes into the two connected components $M_\pm$ containing the images of $\pm\alpha$. Both $M_\pm$ contain P as a submanifold, embedded via $\pm\alpha$. Moreover, the mappings $P \times \mathbb{R} \to M_\pm$, defined by $(m, t) \mapsto \pm \mathrm{e}^{t/2}\alpha_m$, are symplectomorphisms making the diagram

$$\begin{array}{ccc} P \times \mathbb{R} & \xrightarrow{\hspace{3em}} & M_\pm \\ & {\scriptstyle i}\nwarrow \quad \nearrow{\scriptstyle \pm\alpha} & \\ & P & \end{array}$$

commutative.

Finally, we consider a generalization of the class of examples we started with.

Definition 8.5.23 (Hypersurface of contact type) Let (M, ω) be a symplectic manifold. A compact hypersurface (P, ι) of M is said to be of contact type iff P admits a contact form α such that $\mathrm{d}\alpha = \iota^*\omega$.

Thus, every hypersurface of contact type is an exact contact manifold.

Lemma 8.5.24 *Let (M, ω) be a symplectic manifold and let (P, ι) be a hypersurface of contact type with contact form α. There exists a 1-form λ on a neighbourhood U of P in M such that*

$$\mathrm{d}\lambda = \omega_{\restriction U}, \qquad \iota^*\lambda = \alpha.$$

Proof Since P is compact, it is embedded. Since it admits a global contact form, it is orientable. Thus, P admits a transversal vector field, constructed for example by means of the following data:

(a) a covering by oriented charts (U, κ) mapping $P \cap U$ to an open subset of $\mathbb{R}^{2n-1} \times \{0\} \subset \mathbb{R}^{2n}$,
(b) a subordinate partition of unity,
(c) the unit vector field on $\mathbb{R}^{2n}$ given by the standard basis element $\mathbf{e}_{2n}$.

Since P is compact, the flow of a transversal vector field defines a diffeomorphism $\phi : P \times (-\varepsilon, \varepsilon) \to U$ with $\phi(m, 0) = m$ for all $m \in P$. Let $\mathrm{pr}_1 : P \times (-\varepsilon, \varepsilon) \to P$ denote the natural projection and define a 1-form on U by

$$\tau := \left(\mathrm{pr}_1 \circ \phi^{-1}\right)^* \alpha.$$

Consider the 2-form $\omega_{\restriction U} - \mathrm{d}\tau$. Since $\mathrm{pr}_1 \circ \phi^{-1} \circ \iota = \mathrm{id}_P$, we have $\iota^*(\omega_{\restriction U} - \mathrm{d}\tau) = 0$. Hence, the generalized Poincaré Lemma 4.3.14 implies the existence of a 1-form β on U such that $\omega_{\restriction U} - \mathrm{d}\tau = \mathrm{d}\beta$ and $\iota^*\beta = 0$. Then, $\lambda := \tau + \beta$ has the desired properties. □

Proposition 8.5.25 *A compact hypersurface (P, ι) of a symplectic manifold (M, ω) is of contact type iff on some neighbourhood of P in M there exists a Liouville vector field which is transversal to P.*

Proof First, assume that there exists a transversal Liouville vector field Z on a neighbourhood U of P in M. Define $\alpha := Z \lrcorner \omega$. Using the Liouville property (8.5.7), we find

$$\mathrm{d}\alpha = \mathrm{d}(Z \lrcorner \omega) = \mathscr{L}_Z \omega = \omega.$$

Now, (8.5.8) and transversality imply that $\iota^*(\alpha \wedge (\mathrm{d}\alpha)^n)$ is a volume form on P. Hence, $\iota^*\alpha$ is a contact form on P. Conversely, assume that $\alpha \in \Omega^1(P)$ is a contact form with the property $\mathrm{d}\alpha = \iota^*\omega$. We define a vector field Z on a neighbourhood U

of P by $Z \lrcorner \omega = \lambda$, where λ is the 1-form provided by Lemma 8.5.24. This vector field has the desired properties: due to

$$\mathscr{L}_Z \omega = \mathrm{d}(Z \lrcorner \omega) = \mathrm{d}\lambda = \omega,$$

it is Liouville and due to $\alpha = \iota^* \lambda$, it is transversal. □

Corollary 8.5.26 *Let (M, ω) be a symplectic manifold and let $P \subset M$ be a hypersurface of contact type with contact form α. Let $U \subset M$ be a neighbourhood of P on which there exists a Liouville vector field Z transversal to P.*

1. *U contains a neighbourhood V of P in M which is foliated by hypersurfaces of contact type modelled on P.*
2. *The flow of Z yields isomorphisms of the characteristic line bundles.*

Proof Let $\iota : P \to M$ denote the natural inclusion mapping.

1. Since P is compact and since Z is transversal to P, by restriction, the flow Φ of Z induces a diffeomorphism from $(-\varepsilon, \varepsilon) \times P$ onto some open neighbourhood V of P in U. For $t \in (-\varepsilon, \varepsilon)$, we define

$$P_t := \Phi_t(P), \qquad \alpha_t := \Phi_{-t}^* \alpha.$$

Since $\mathrm{d}\alpha_t = (\iota \circ \Phi_{-t})^* \omega$ and since $\iota \circ \Phi_{-t} : P_t \to M$ is the natural inclusion mapping, P_t is a hypersurface of contact type with globally defining 1-form α_t.

2. Let $t \in (-\varepsilon, \varepsilon)$. According to Remark 8.5.11/2, the characteristic line bundle of P_t can be identified with $\ker \mathrm{d}\alpha_t$. From $\mathscr{L}_Z \omega = \omega$ we obtain $\Phi_t^* \omega = e^t \omega$. For $p \in P$, $X \in \ker(\mathrm{d}\alpha)_p$ and $Y \in \mathrm{T}_p P$, we find $\mathrm{d}\alpha(X, Y) = \iota^* \omega(X, Y) = 0$ and hence

$$0 = \big(e^t \iota^* \omega\big)(X, Y) = \big((\Phi_t \circ \iota)^* \omega\big)(X, Y) = \omega\big((\Phi_t \circ \iota)' X, (\Phi_t \circ \iota)' Y\big).$$

It follows that $(\Phi_t \circ \iota)' X \in \ker \mathrm{d}\alpha_t$. Thus, by restriction, $(\Phi_t \circ \iota)'$ induces an isomorphism from $\ker \mathrm{d}\alpha$ onto $\ker \mathrm{d}\alpha_t$. □

For a deeper discussion of hypersurfaces of contact type we refer to [139]. We will meet them again in Sects. 9.3 and 9.4.

Exercises

8.5.1 Prove Remark 8.5.7/2 by induction on the number of functions in involution.

8.5.2 Use Proposition 8.5.12 and a partition of unity to prove Proposition 8.5.15. *Hint*. For local contact forms α and functions f, $(f\alpha) \wedge (\mathrm{d}(f\alpha))^n = f^{n+1} \alpha \wedge (\mathrm{d}\alpha)^n$.

8.5.3 Prove that (8.5.12) is a contact form on $\mathbb{R}^{2n+1}$. Calculate it in the polar coordinates r_i, φ_i on the planes (x_i, y_i).

8.5.4 Prove that $\iota^* \alpha$, with α defined by (8.5.13), is a contact form on S^{2n+1}.

8.5.5 Prove that (8.5.15) defines a contact structure on the projective cotangent bundle $\mathrm{P}^* Q$.

8.5.6 Prove that the symplectization of S^{2n+1} yields $\mathbb{R}^{2n} \setminus \{0\}$, cf. Example 8.5.21. *Hint*. Use the logarithm of the radius function.

8.5.7 Provide proofs for the statements made in Remark 8.5.22/2.

8.6 Generalizations of the Darboux Theorem

In this section we discuss natural generalizations of the classical Darboux Theorem, which are important both in the theory of integrable systems and in the theory of symmetry reduction of Hamiltonian systems. Moreover, they have proved useful in other applications, too, like e.g. in canonical realizations of Lie algebras, see [74]. Most of the results below belong to Weinstein, see [305]. As in the proof of the classical Darboux Theorem, the deformation method of Moser plays a central role.

Theorem 8.6.1 (Weinstein) *Let M be a manifold, endowed with two symplectic forms ω_0 and ω_1 and let $N \subset M$ be an embedded submanifold, on which ω_0 and ω_1 coincide, that is,*

$$\omega_0(X, Y) = \omega_1(X, Y)$$

for all $X, Y \in \mathrm{T}_m M$ and $m \in N$. Then, there exists a diffeomorphism Φ of open neighbourhoods U and $\Phi(U)$ of $N \subset M$ with the properties

$$\Phi^* \omega_1 = \omega_0, \qquad \Phi_{\restriction N} = \mathrm{id}\,.$$

Moreover, Φ can be chosen so that $\Phi'_m = \mathrm{id}_{\mathrm{T}_m M}$ for all $m \in N$.

Proof By the Tubular Neighbourhood Theorem for embedded submanifolds, cf. Remark 6.4.7, there exists an open neighbourhood V of $N \subset M$ which can be diffeomorphically identified with a neighbourhood of the zero section s_0 of a vector bundle (E, N, π). In what follows, we view V in this way. Let $\mu_s : E \to E$ be the fibrewise multiplication by $s \in \mathbb{R}$ and let V be chosen so that $\mu_s(V) \subset V$ for all $s \in [0, 1]$. Let $\Delta := \{(s, t) \in [0, 1] \times [0, 1) : t \leq s\}$. There exists an open neighbourhood $\mathscr{D}$ of $\Delta \times V$ in $\mathbb{R}^2 \times V$ such that the mapping

$$\phi : \mathscr{D} \to V, \qquad \phi(s, t, m) \equiv \phi_{s,t}(m) := \mu_{(1-s)(1-t)^{-1}}(m),$$

is well defined. We have $\phi_{0,0} = \mathrm{id}_V$ and $\phi_{1,t} = s_0 \circ \pi$, and $\phi_{s,t}$ is a diffeomorphism onto its image for all $s < 1$. Thus, for $s < 1$, ϕ defines a time-dependent flow (Exercise 8.6.1) on V. Let Y be the corresponding time-dependent vector field,

$$Y_s(m) = \frac{\mathrm{d}}{\mathrm{d}\tau}_{\restriction 0} \phi_{s+\tau,s}(m) = -\mu_{(1-s)^{-1}}(m). \tag{8.6.1}$$

Obviously, Y_s vanishes on N. Using (4.1.28) and denoting $\phi_s \equiv \phi_{s,0}$, for $m \in V$ we compute:

$$\begin{aligned}(\omega_0 - \omega_1)(m) &= \big(\phi_1^*(\omega_1 - \omega_0) - (\omega_1 - \omega_0)\big)(m) \\ &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\big(\phi_s^*(\omega_1 - \omega_0)\big)(m)\mathrm{d}s\end{aligned}$$

$$= \int_0^1 \big(\phi_s^*\big(\mathscr{L}_{Y_s}(\omega_1 - \omega_0)\big)\big)(m)\mathrm{d}s$$

$$= \int_0^1 \mathrm{d}\big(\phi_s^*\big(Y_s \lrcorner(\omega_1 - \omega_0)\big)\big)(m)\mathrm{d}s. \tag{8.6.2}$$

Thus, by (4.1.30),

$$\omega_0 - \omega_1 = \mathrm{d}\alpha, \quad \text{where } \alpha := \int_0^1 \phi_s^*\big(Y_s \lrcorner(\omega_1 - \omega_0)\big)\mathrm{d}s. \tag{8.6.3}$$

Although Y_s is not defined for $s = 1$, the integral exists, because ϕ_s' amounts to multiplication by $(1-s)$ and hence the factor $(1-s)^{-1}$ in (8.6.1) cancels. Note that α vanishes on N. Now, as in the proof of the Darboux Theorem 8.1.5, we consider the family of 2-forms

$$\omega_t := \omega_0 + t(\omega_1 - \omega_0), \quad t \in [0,1]. \tag{8.6.4}$$

By analogous arguments, one can show that ω_t is non-degenerate on an open neighbourhood W of $[0,1] \times N$ in $[0,1] \times V$. Thus, one can define a time-dependent vector field X on W by

$$X_t \lrcorner \omega_t = \alpha. \tag{8.6.5}$$

There exists an open neighbourhood U_0 of N in V such that $[0,1] \times \{0\} \times U_0$ is contained in the domain of the flow Φ of X.[11] Writing $\Phi_t \equiv \Phi_{t,0}$, one finds

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi_t^*\omega_t = \Phi_t^*\left(\mathscr{L}_{X_t}\omega_t + \frac{\mathrm{d}}{\mathrm{d}t}\omega_t\right) = \Phi_t^*(\mathrm{d}\alpha + \omega_1 - \omega_0) = 0$$

on $[0,1] \times U_0$. Since $\Phi_0 = \mathrm{id}$, this implies $\Phi_1^*\omega_1 = \omega_0$. Since X_t vanishes on N for all t, we have $\Phi_{1\restriction N} = \mathrm{id}_N$. Thus, $\Phi = \Phi_1 : U_0 \to U_1 = \Phi_1(U_0)$ is the desired diffeomorphism.

To see that $(\Phi_t)'_m = \mathrm{id}_{\mathrm{T}_m M}$ for all $m \in N$, we choose a local chart with coordinates x^i. Since $\omega_0 - \omega_1$ and Y_s vanish on N, the coefficients of the 1-form $Y_s \lrcorner(\omega_1 - \omega_0)$ are sums of products of two functions vanishing on N. Then, the first derivatives of these coefficients vanish on N, too, and this is also true for the coefficients of α. Using this, we conclude that the partial derivatives of the coefficients of X_t vanish on N as well. Differentiating the defining equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t,m) = X_t\big(\Phi(t,m)\big)$$

with respect to x^i we obtain a differential equation for the matrix of $(\Phi_t)'_m$. Since the partial derivatives of the coefficients of X_t vanish on N and since $\Phi_0 = \mathrm{id}$, the solution is $(\Phi_t)'_m = \mathrm{id}_{\mathrm{T}_m M}$ for all t, hence in particular for $t = 1$. □

[11] U_0 is the union of the subsets of V obtained by applying the Tube Lemma to every point of N.

Let us consider the special case of an orbit of a symplectic Lie group action.[12]

Definition 8.6.2 (Symplectic Lie group action) Let (M, ω) be a symplectic manifold. An action Ψ of a Lie group G on M is called symplectic, or canonical, if

$$\Psi_a^* \omega = \omega \tag{8.6.6}$$

for all $a \in G$. In this case, the tuple (M, ω, Ψ) is called a symplectic G-manifold.

In what follows, we assume that Ψ is a left action. The following corollary is a direct consequence of Theorem 8.6.1, with the submanifold N being an orbit of Ψ.

Corollary 8.6.3 (Equivariant Darboux Theorem) *Let a manifold M be endowed with two symplectic forms ω_0 and ω_1 and a proper action Ψ of a Lie group G which is symplectic with respect to both ω_0 and ω_1. If ω_0 and ω_1 coincide on an orbit O of Ψ, that is, if*

$$\omega_0(X, Y) = \omega_1(X, Y)$$

for all $m \in O$ and $X, Y \in \mathrm{T}_m M$, there exist Ψ-invariant open neighbourhoods U_0 and U_1 of O and a Ψ-equivariant diffeomorphism $\Phi : U_0 \to U_1$ with the property

$$\Phi^* \omega_1 = \omega_0, \qquad \Phi_{\restriction O} = \mathrm{id}\,.$$

The diffeomorphism Φ can be chosen so that $\Phi'_m = \mathrm{id}_{\mathrm{T}_m M}$ for all $m \in O$.

Proof The proof is completely analogous to that of Theorem 8.6.1, with the general Tubular Neighbourhood Theorem replaced by Theorem 6.4.3. Obviously, ω_t is invariant under Ψ. Since the fibrewise multiplication μ_s commutes with Ψ, the flow ϕ and thus the vector field Y are invariant, too. This implies G-invariance of α and X, and thus of the symplectomorphism Φ. □

Another special case is that of a Lagrangian submanifold L. This is important in the theory of integrable systems. By Lemma 8.1.8, the tangent bundle $\mathrm{T}L$ admits a Lagrangian complement E in $\mathrm{T}M_{\restriction L}$, so that

$$\mathrm{T}M_{\restriction L} = \mathrm{T}L \oplus E. \tag{8.6.7}$$

Pointwise application of Proposition 7.2.9 yields the vertical isomorphism of vector bundles

$$\chi : E \to \mathrm{T}^* L, \qquad \chi(Z) = \omega(Z, \cdot). \tag{8.6.8}$$

For all $m \in L$, this isomorphism induces symplectomorphisms

$$\Psi_m^L : \mathrm{T}_m M \to \mathrm{T}_m L \oplus \mathrm{T}_m^* L, \qquad \Psi_m^L(X) := X_L \oplus \chi(X_E), \tag{8.6.9}$$

[12] This notion has been already announced in Remark 6.1.3.

with X_L and X_E being the components of X with respect to the decomposition (8.6.7). The reader may compare this with Proposition 8.3.12 by putting $M = \mathrm{T}^*Q$ and taking for L the image of a closed 1-form on Q.

Theorem 8.6.4 (Weinstein) *Let L be an embedded Lagrangian submanifold of a symplectic manifold (M,ω). Then, there exists a symplectomorphism Φ of an open neighbourhood of L in M onto an open neighbourhood of the zero section s_0 of T^*L such that $\Phi_{\restriction L} = s_0$. For every Lagrangian complement E of $\mathrm{T}L$ in $\mathrm{T}M_{\restriction L}$, Φ can be chosen so that for all $m \in L$,*

$$\Phi'_m(E_m) = \mathrm{T}_{\Phi(m)}\big(\mathrm{T}^*_m L\big). \tag{8.6.10}$$

Proof Let E be a Lagrangian complement of $\mathrm{T}L$ in $\mathrm{T}M_{\restriction L}$ and let $\chi : E \to \mathrm{T}^*L$ be the vector bundle isomorphism defined by (8.6.8). The Tubular Neighbourhood Theorem for embedded submanifolds implies the existence of a diffeomorphism φ of an open neighbourhood U of $L \subset M$ onto an open neighbourhood $\varphi(U)$ of the zero section e_0 in E such that

$$\varphi_{\restriction L} = e_0, \qquad \varphi'_{\restriction E_m} = \mathrm{id}_{\mathrm{T}_{\varphi(m)}E_m} \tag{8.6.11}$$

for all $m \in L$. Thus, $\psi = \chi \circ \varphi$ is a diffeomorphism of U onto an open neighbourhood $\psi(U)$ of the zero section s_0 in T^*L with the property $\psi_{\restriction L} = s_0$. Let θ denote the canonical 1-form on T^*L. We show that $\psi^*\mathrm{d}\theta$ and ω coincide on L and apply Theorem 8.6.1. For that purpose, we show that $\psi'_m : \mathrm{T}_m M \to \mathrm{T}_{s_0(m)}(\mathrm{T}^*L)$ is symplectic for all $m \in L$. A brief computation using (8.6.11) and $\chi \circ e_0 = s_0$ shows that for $X \in \mathrm{T}_m M$,

$$\psi'_m(X) = \chi'_{e_0(m)} \circ \varphi'_m(X) = s'_0(X_L) + \widehat{\chi(X_E)}_{s_0(m)}, \tag{8.6.12}$$

where $\widehat{\chi(X_E)}$ is the constant vector field on $\mathrm{T}^*_m L$ defined by $\chi(X_E) \in \mathrm{T}^*_m L$ (Exercise 8.6.2). Hence, ψ'_m coincides with the composition

$$\mathrm{T}_m M \to \mathrm{T}_m L \oplus \mathrm{T}^*_m L \to \mathrm{T}_{s_0(m)}\big(\mathrm{T}^*L\big), \tag{8.6.13}$$

where the first mapping is the symplectomorphism (8.6.9) and the second one is the symplectomorphism provided by Proposition 8.3.12. Now, we apply Theorem 8.6.1. Composing ψ with the local symplectomorphism provided by this theorem, we obtain the desired symplectomorphism Φ. □

Finally, let us assume that the symplectic manifold (M,ω) is endowed with the additional structure of an integrable Lagrangian foliation $\mathscr{F}$, that is, a foliation by Lagrangian submanifolds. Let D be the corresponding Lagrangian distribution, given by the tangent spaces of the leaves of $\mathscr{F}$. The standard example of a Lagrangian foliation is given by the fibres of the cotangent bundle of a manifold. We say that a Lagrangian submanifold L of M and a Lagrangian foliation $\mathscr{F}$ of M

are transversal if L is transversal to every leaf of $\mathscr{F}$. In this case, for dimensional reasons, one has

$$\mathrm{T}_m M = \mathrm{T}_m L \oplus D_m$$

for every $m \in L$. In the presence of a transversal Lagrangian foliation, the symplectomorphism Φ of Theorem 8.6.4 can be chosen in such a way that it respects the fibre structure:

Corollary 8.6.5 *Let (M, ω) be a symplectic manifold, let L be an embedded Lagrangian submanifold and let $\mathscr{F}$ be a Lagrangian foliation of M transversal to L. Then, the symplectomorphism Φ of Theorem 8.6.4 can be chosen in such a way that the intersection of any leaf of $\mathscr{F}$ with the domain of Φ is mapped to a fibre of T^*L.*

Proof Let D be the distribution defined by $\mathscr{F}$. In the notation of the proof of Theorem 8.6.4, we choose $E = D_{\restriction L}$. Since χ is fibre-preserving, in a first step one must show that φ can be chosen in such a way that $\psi = \chi \circ \varphi$ maps leaves of $\mathscr{F}$ to fibres of T^*L. This follows from the general Tubular Neighbourhood Theorem, see Remark 6.4.7. In a second step, one must show that the mapping Φ of Theorem 8.6.1 maps every leaf of $\mathscr{F}$ to itself. For that purpose, it is enough to show that, in the notation of the proof of that theorem, the time-dependent vector field X defined by (8.6.5) lies in D: by construction, the vector field Y_s, defined by (8.6.1), lies in D. Since D is Lagrange with respect to both $\psi^*\mathrm{d}\theta$ and ω, the 1-form α defined by (8.6.3) vanishes on all vectors of D. Thus, the vector field X_t lies in the ω_t-orthogonal complement of D. Since D is also Lagrange with respect to ω_t, it follows that X_t lies in D. □

The fibrewise transitive action of the cotangent bundle on itself, defined in Sect. 8.3, carries over to the case under consideration. Let us recall that a foliation $\mathscr{F}$ on a manifold M is called simple if there exists a surjective submersion $\rho\colon M \to B$ onto another manifold B such that for every point $m \in M$ the leaves of $\mathscr{F}$ are the closed submanifolds $\rho^{-1}(b)$, $b \in B$, cf. Example 3.5.20/2. We may view B as the space of leaves, $B = M/\mathscr{F}$. Every $\xi \in \mathrm{T}^*_b B$, $b \in B$, defines a vector field $\hat{\xi}$ on the leaf $\rho^{-1}(b)$ by

$$\hat{\xi}_m \lrcorner \omega_m = \left(\rho'_m\right)^{\mathrm{T}}(\xi). \tag{8.6.14}$$

In what follows, we assume that the flow ϕ^ξ of $\hat{\xi}$ is complete for every $\xi \in \mathrm{T}^*B$. For example, this is the case if the leaves of $\mathscr{F}$ are compact. Then, we can define $\phi_\xi = \phi^\xi_1$, which yields a diffeomorphism of the leaf $\rho^{-1}(b)$. Thus, every fibre $\mathrm{T}^*_b B$ acts on the leaf $\rho^{-1}(b)$ as a vector group. Now, let

$$s\colon B \to s(B) = L \subset M$$

be a Lagrangian section of ρ, that is, a Lagrangian submanifold such that $\rho \circ s = \mathrm{id}_B$. Then,

$$\Phi\colon \mathrm{T}^*B \to M, \qquad \Phi(\xi) := \phi_\xi\big(s \circ \rho(\xi)\big), \tag{8.6.15}$$

defines a symplectomorphism of a neighbourhood of the zero section of T^*B onto a neighbourhood of $L = s(B)$ in M, (Exercise 8.6.3). This way we obtain an alternative proof of Theorem 8.6.4, with the Lagrangian complement E being replaced by the integrable Lagrangian distribution D.

Let us add that the assignment $\xi \mapsto \hat{\hat{\xi}}$ extends to 1-forms on B by assigning to $\alpha \in \Omega^1(B)$ the vertical vector field $\hat{\alpha}$ on M given by

$$\hat{\alpha}_m := \widehat{\alpha_{\rho(m)}}$$

for all $m \in M$, cf. (8.3.11). Then,

$$\hat{\alpha} \lrcorner \omega = \rho^* \alpha \tag{8.6.16}$$

and the flow ϕ^α of $\hat{\alpha}$ is given leafwise by $\phi_t^\alpha(m) = \phi_t^{\alpha_{\rho(m)}}(m)$. It defines a diffeomorphism of M by $\phi_\alpha := \phi_1^\alpha$. This diffeomorphism is vertical with respect to ρ and satisfies $\phi_\alpha(m) = \phi_{\alpha_{\rho(m)}}(m)$ for all $m \in M$. As in the case of the cotangent bundle, one can show that

$$\phi_\alpha^* \omega = \omega + \rho^* \mathrm{d}\alpha, \qquad \mathscr{L}_{\hat{\alpha}} \omega = \rho^* \mathrm{d}\alpha, \qquad [\hat{\alpha}, \hat{\beta}] = 0 \tag{8.6.17}$$

(Exercise 8.6.4).

Exercises

8.6.1 Complete the proof of Theorem 8.6.1 by verifying the properties of the 2-parameter family of mappings $\{\phi_{s,t}\}$ and proving formula (8.6.1).

8.6.2 Prove Formula (8.6.12).

8.6.3 Show that the mapping Φ, defined by (8.6.15), is a symplectomorphism of a neighbourhood of the zero section of T^*L onto a neighbourhood of L in M.

8.6.4 Prove the formulae in (8.6.17).

8.6.5 Let Φ be the local symplectomorphism of Corollary 8.6.5 and let $\beta := \Phi^*\theta$. Show that

$$\mathrm{d}\beta = \omega, \quad D \subset \ker \beta, \quad \beta_{\restriction L} = 0.$$

8.7 Symplectic Reduction

The general theory of symplectic reduction goes back to Benenti and Tulczyjew [37, 42]. In this section we present the main aspects of this theory without going into all details, for which we refer the reader to the above papers. An exhaustive treatment can also be found in [181]. The particularly important case of symplectic reduction for systems with symmetries will be dealt with in Chap. 10.

Definition 8.7.1 (Symplectic reduction) Let (M, ω) be a symplectic manifold, let (N, ι_N) be a submanifold and let (P, ω_P) be another symplectic manifold. A mapping $\pi : N \to P$ is called a symplectic reduction if it is a surjective submersion fulfilling

$$\pi^*\omega_P = \iota_N^*\omega. \tag{8.7.1}$$

In this case, the pair (P, ω_P) is called a reduced symplectic manifold for N. A symplectic reduction is called strict if N is coisotropic.

Comparing this definition with that of linear symplectic reduction in Sect. 7.3, we see that P is the geometric counterpart of the quotient vector space $\hat{W}$ given by (7.3.1) and ω_P corresponds to the symplectic form ω_W defined by (7.3.2). In particular, the case of strict reduction corresponds to the case where W is coisotropic, cf. (7.3.3). For what follows, we denote

$$\omega_N := \iota_N^*\omega.$$

Let F^{ω_N} be the family of characteristic subspaces and let D^{ω_N} be the characteristic distribution of ω_N, cf. Definition 4.2.18. Recall from Proposition 4.2.20 that D^{ω_N} is integrable and that the foliation consisting of the maximal integral manifolds is referred to as the characteristic foliation $\mathscr{F}^{\omega_N}$ of ω_N. Moreover, recall from Example 3.5.20/2 that if $\mathscr{F}^{\omega_N}$ is simple, the corresponding space of leaves is unique up to diffeomorphisms. We say that a surjective submersion $\pi : N \to P$ has connected fibres if $\pi^{-1}(p)$ is connected for every $p \in P$.

Proposition 8.7.2 *Let (M, ω) be a symplectic manifold and (N, ι_N) a submanifold.*

1. *If $\pi : N \to P$ is a symplectic reduction onto a symplectic manifold (P, ω_P), then*

$$\operatorname{rank}\omega_N = \dim P, \qquad D^{\omega_N} = \ker\omega_N = \ker\pi'. \tag{8.7.2}$$

   In particular, D^{ω_N} is regular. If π has connected fibres, $\mathscr{F}^{\omega_N}$ is simple.
2. *If the rank of ω_N is constant and if $\mathscr{F}^{\omega_N}$ is simple, the space of leaves P admits a unique symplectic form ω_P such that the canonical projection $\pi : N \to P$ is a symplectic reduction. If $\tilde{\pi} : N \to \tilde{P}$ is another symplectic reduction onto a symplectic manifold $(\tilde{P}, \tilde{\omega})$, then there exists a surjective local symplectomorphism $\chi : P \to \tilde{P}$ such that $\tilde{\pi} = \chi \circ \pi$. If $\tilde{\pi}$ has connected fibres, χ is a symplectomorphism.*

Proof Due to $\mathrm{d}\omega_N = 0$, we have $F^{\omega_N} = \ker\omega_N$.

1. Since ω_P is non-degenerate, (8.7.1) implies $\ker\omega_N = \ker\pi'$. First, this yields

$$\operatorname{rank}\omega_N = \dim N - \dim\ker\omega_N = \dim\ker\pi' + \dim P - \dim\ker\omega_N = \dim P.$$

Second, this implies that $\ker\omega_N = F^{\omega_N}$ has constant rank and hence, by the remarks after Definition 4.2.18, that it coincides with D^{ω_N}. In particular, D^{ω_N} is regular and coincides with $\ker\pi'$. Since π is a submersion, the maximal integral manifolds of the distribution $\ker\pi'$ are the connected components of the fibres of π. Therefore, if the fibres of π are connected, they coincide with the maximal integral manifolds of $\ker \pi' = D^{\omega_N}$ and hence with the leaves of $\mathscr{F}^{\omega_N}$. Thus, in this case, $\mathscr{F}^{\omega_N}$ is simple.

2. Using an atlas on N adapted to the foliation $\mathscr{F}^{\omega_N}$, one can check that for any two pairs of tangent vectors $X_i, Y_i \in \mathrm{T}_{m_i} N$, $i = 1, 2$, satisfying $\pi' X_1 = \pi' X_2$ and $\pi' Y_1 = \pi' Y_2$, one has

$$\omega_N(X_1, Y_1) = \omega_N(X_2, Y_2), \tag{8.7.3}$$

see Exercise 8.7.1. Since in addition π is a submersion, Eq. (8.7.1), taken pointwise, defines a unique 2-form $(\omega_P)_p$ on $\mathrm{T}_p P$ for every $p \in P$. Since π admits local sections, these 2-forms combine to a unique differential 2-form ω_P on P. Since

$$\pi^* \mathrm{d}\omega_P = \mathrm{d}\big(\pi^* \omega_P\big) = \mathrm{d}\omega_N = 0$$

and, again, since π is a submersion, ω_P is closed. Due to $\operatorname{rank} \omega_P = \operatorname{rank} \omega_N = \dim P$, it is non-degenerate and hence symplectic. Now, let $\tilde{\pi} : N \to \tilde{P}$ be another symplectic reduction. Then, by point 1,

$$\ker \pi' = \ker \omega_N = \ker \tilde{\pi}'.$$

Since P is the space of leaves of $\mathscr{F}^{\omega_N}$, the fibres of π are the maximal integral manifolds of the distribution $\ker \pi'$ and hence the connected components of the fibres of $\tilde{\pi}$. In particular, the fibres of π are contained in those of $\tilde{\pi}$. Hence, there exists a unique mapping $\chi : P \to \tilde{P}$ such that $\tilde{\pi} = \chi \circ \pi$. Since π admits local sections, this mapping is smooth. Using

$$\pi^* \omega_P = \omega_N = \tilde{\pi}^* \tilde{\omega}_P = \pi^* \big(\chi^* \tilde{\omega}_P\big)$$

and the fact that π is a submersion, we conclude that $\chi^* \tilde{\omega}_P = \omega_P$. Since P and $\tilde{P}$ have the same dimension, according to Proposition 8.1.3, χ is a local symplectomorphism. It is surjective, because $\tilde{\pi}$ is surjective. Obviously, if $\tilde{\pi}$ has connected fibres, the same argument applied backwards yields that χ is a symplectomorphism. □

The next Proposition characterizes symplectic reductions in terms of their graphs.

Proposition 8.7.3 *Let (M, ω) be a symplectic manifold, let (N, ι_N) be a submanifold and let $\pi : N \to P$ be a surjective submersion onto a symplectic manifold (P, ω_P). The mapping π is a symplectic reduction iff its graph Γ_π is an isotropic submanifold of $M \times P$ endowed with the symplectic product structure $\omega^-_{M \times P}$ defined by* (8.1.2)*. The reduction is strict iff Γ_π is Lagrange.*

Proof We calculate

$$(\iota_N \times \pi)^* \big(\mathrm{pr}_M^* \, \omega - \mathrm{pr}_P^* \, \omega_P\big) = \omega_N - \pi^* \omega_P .$$

Since $\iota_N \times \pi$ is a diffeomorphism from N onto Γ_π, (8.7.1) holds iff Γ_π is isotropic. By Proposition 7.2.4/3, N is coisotropic iff $\operatorname{rank}\omega_N = 2\dim N - \dim M$. By Formula (8.7.2), we have $\operatorname{rank}\omega_N = \dim P$. Thus, N is coisotropic iff

$$\dim N = \frac{1}{2}(\dim M + \dim P),$$

that is, iff Γ_π is Lagrange. □

Finally, in applications in physics, it often happens that there is a submanifold of the symplectic manifold (the phase space), which is invariant under the flow of the Hamiltonian vector field of a given function (a physical observable). We show that if the assumptions of symplectic reduction are fulfilled, one obtains a reduction of the flow to a lower dimensional manifold (the reduced phase space).

Proposition 8.7.4 *Let (M, ω) be a symplectic manifold, let $N \subset M$ be a submanifold and let $\pi : N \to P$ be a symplectic reduction with connected fibres onto a symplectic manifold (P, ω_P). Let $f \in C^\infty(M)$ and assume that N is invariant under the flow of the Hamiltonian vector field X_f. Then, there exists a unique function $\hat{f} \in C^\infty(P)$ such that*

$$\iota_N^* f = \pi^* \hat{f}.$$

The vector field on N induced by X_f is π-related to $X_{\hat{f}}$.

Proof Since the characteristic distribution of ω_N is given by $\ker\omega_N = \mathrm{T}N \cap (\mathrm{T}N)^\omega$ and since Proposition 7.2.1/1 implies $(\mathrm{T}N)^\flat = ((\mathrm{T}N)^\omega)^0$, we have

$$(\mathrm{T}N)^\flat \subset (\ker\omega_N)^0.$$

Thus, if $(X_f)_{\restriction N}$ takes values in $\mathrm{T}N$, then $(\mathrm{d}f)_{\restriction N}$ takes values in $(\ker\omega_N)^0$ and hence f is constant on each leaf of the characteristic foliation $\mathscr{F}^{\omega_N}$. Since π is a surjective submersion with connected fibres, there exists a unique smooth function $\hat{f} \in C^\infty(P)$ such that $\pi^*\hat{f} = \iota_N^* f \equiv f_N$. Then, $\pi^*\mathrm{d}\hat{f} = \mathrm{d}f_N$. Using $\omega_N = \pi^*\omega_P$ and the fact that the vector field $\tilde{X}_f$ induced by X_f on N is ι_N-related to X_f, we obtain

$$\tilde{X}_f \lrcorner \big(\pi^*\omega_P\big) = \tilde{X}_f \lrcorner \omega_N = -\mathrm{d}f_N = -\pi^*\mathrm{d}\hat{f} = \pi^*(X_{\hat{f}} \lrcorner \omega_P).$$

Evaluating both sides on an arbitrary tangent vector $Y \in \mathrm{T}N$ and using that ω_P is non-degenerate and that π is a submersion, we conclude

$$\pi' \circ \tilde{X}_f = X_{\hat{f}} \circ \pi,$$

that is, $\tilde{X}_f$ is π-related to $X_{\hat{f}}$. □

Remark 8.7.5

1. For $\hat{f}$ to exist, it suffices to require that f be constant on the leaves of the characteristic foliation $\mathscr{F}^{\omega_N}$. In this case, $(\mathrm{d}f)_{\restriction N}$ takes values in $(\ker \omega_N)^0$ and thus Proposition 7.2.1/5 implies that $\tilde{X}_f$ takes values in
$$(\ker \omega_N)^\omega = \big(\mathrm{T}N \cap (\mathrm{T}N)^\omega\big)^\omega = \mathrm{T}N + (\mathrm{T}N)^\omega .$$
2. By Proposition 3.2.13, the flows Φ^f of $\tilde{X}_f$ and $\Phi^{\hat{f}}$ of $X_{\hat{f}}$ fulfil
$$\Phi_t^{\hat{f}} \circ \pi = \pi \circ \Phi_t^f$$
on the domain of Φ_t^f. Thus, by symplectic reduction, the problem of finding the integral curves of the Hamiltonian vector field X_f through points of N has been reduced to the corresponding problem on a space of lower dimension. For this reason, such reductions are enormously important in physical applications, see Chap. 10.

Exercises

8.7.1 Complete the proof of Proposition 8.7.2/2 by verifying (8.7.3).
Hint. Show that in a local chart (U, κ) on N adapted to the foliation $\mathscr{F}^{\omega_N}$, with κ^i, $i \le k$ yielding coordinates on the leaves and κ^i, $i > k$ labelling the leaves, one has $\omega_{\restriction U} = \sum_{k<i<j<\dim N} \omega_{ij} \mathrm{d}\kappa^i \wedge \mathrm{d}\kappa^j$, where the functions ω_{ij} are constant on the leaves.

8.8 Symplectomorphisms and Generating Functions

In this section we will show that symplectomorphisms may be viewed as Lagrangian submanifolds and that they can be (locally) generated by functions. Using this fact, we will gain some insight into the structure of the group of symplectomorphisms.

The following proposition is a special case of Proposition 8.7.3.

Proposition 8.8.1 *Let (M_1, ω_1) and (M_2, ω_2) be symplectic manifolds of the same dimension. A diffeomorphism $\Phi : M_1 \to M_2$ is a symplectomorphism iff its graph is a Lagrangian submanifold of $M_1 \times M_2$ endowed with the product symplectic structure $\omega^-_{M_1 \times M_2}$ defined by* (8.1.2).

Proof Let $\iota_\Phi : \Gamma_\Phi \to M_1 \times M_2$ be the natural inclusion mapping and let pr_i be the canonical projections onto the factors of $M_1 \times M_2$. Note that $\Phi \circ \mathrm{pr}_1 \circ \iota_\Phi = \mathrm{pr}_2 \circ \iota_\Phi$. Then,
$$\begin{aligned}
\iota_\Phi^* \omega^-_{M_1 \times M_2} &= \iota_\Phi^* \big(\mathrm{pr}_2^* \omega_2 - \mathrm{pr}_1^* \omega_1\big) \\
&= (\mathrm{pr}_2 \circ \iota_\Phi)^* \omega_2 - (\mathrm{pr}_1 \circ \iota_\Phi)^* \omega_1 \\
&= (\mathrm{pr}_1 \circ \iota_\Phi)^* \big(\Phi^* \omega_2 - \omega_1\big).
\end{aligned}$$

Since $\mathrm{pr}_1 \circ \iota_\Phi$ is a diffeomorphism, we conclude that $\iota_\Phi^* \omega^-_{M_1 \times M_2} = 0$ iff $\omega_1 = \Phi^* \omega_2$. In this case, Γ_Φ is Lagrange, because $\dim \Gamma_\Phi = \dim M_1 = \frac{1}{2} \dim(M_1 \times M_2)$. □

Now, let $\Phi : M_1 \to M_2$ be a symplectomorphism. By the Poincaré Lemma, locally there exists a 1-form τ on $M_1 \times M_2$ such that $\omega^-_{M_1 \times M_2} = \mathrm{d}\tau$. Since $\iota_\Phi^* \omega^-_{M_1 \times M_2} = 0$, $\iota_\Phi^* \tau$ is closed. Applying once again the Poincaré Lemma, we find an open subset $U \subset \Gamma_\Phi$ and a smooth function $S : U \to \mathbb{R}$ fulfilling

$$\left(\iota_\Phi^* \tau\right)_{\restriction U} = -\mathrm{d}S. \tag{8.8.1}$$

The function S is called a generating function of the canonical transformation Φ. Note that S depends on the choice of the potential τ. Let us discuss the choices commonly used in the physics literature. For that purpose, let q^i, p_i and $\bar{q}^i$, $\bar{p}_i$ be local Darboux coordinates on M_1 and M_2, respectively. The induced coordinates on $M_1 \times M_2$ will be denoted by $q^i \equiv q^i \circ \mathrm{pr}_1$, $p_i \equiv p_i \circ \mathrm{pr}_1$, $\bar{q}^i \equiv \bar{q}^i \circ \mathrm{pr}_2$ and $\bar{p}_i \equiv \bar{p}_i \circ \mathrm{pr}_2$. In these coordinates, we have

$$\omega = \mathrm{d}\bar{p}_i \wedge \mathrm{d}\bar{q}^i - \mathrm{d}p_i \wedge \mathrm{d}q^i.$$

The following choices for τ occur frequently:

$$\bar{p}_i \mathrm{d}\bar{q}^i - p_i \mathrm{d}q^i, \qquad -\bar{q}^i \mathrm{d}\bar{p}_i - p_i \mathrm{d}q^i, \qquad \bar{p}_i \mathrm{d}\bar{q}^i + q^i \mathrm{d}p_i, \qquad -\bar{q}^i \mathrm{d}\bar{p}_i + q^i \mathrm{d}p_i.$$

The corresponding generating functions S are said to be of the i-th kind, $i = 1, \ldots, 4$. Which choice is convenient depends on Φ. If, for instance, the functions $q^i \circ \iota_\Phi$ and $\bar{q}^i \circ \iota_\Phi$ define a local chart λ on Γ_Φ, one may use $\tau = \bar{p}_i \mathrm{d}\bar{q}^i - p_i \mathrm{d}q^i$, so that the corresponding generating function S is of the first kind. It can be determined as follows. Using the simplified notation

$$p_i \equiv p_i \circ \iota_\Phi \circ \lambda^{-1}, \qquad \bar{p}_i \equiv \bar{p}_i \circ \iota_\Phi \circ \lambda^{-1}, \tag{8.8.2}$$

and writing $S \equiv S \circ \lambda^{-1}$, from (8.8.1) we read off the relations

$$p_i = \frac{\partial S}{\partial q^i}, \qquad \bar{p}_i = -\frac{\partial S}{\partial \bar{q}^i}, \tag{8.8.3}$$

taught in the standard course in classical mechanics. Since the functions (8.8.2) are determined by Φ, (8.8.3) is a system of first order partial differential equations for a function in the $2n$ variables q^i and $\bar{q}^i$. This system has a unique solution, up to an additive constant.

By analogy, if the combinations of coordinates given by $(q, \bar{p})$, $(p, \bar{q})$ or $(p, \bar{p})$ yield local charts on Γ_Φ, one may work with generating functions of the second, third or fourth kind. The corresponding systems of differential equations read, respectively,

$$\bar{q}^i = \frac{\partial S}{\partial \bar{p}_i}, \qquad p_i = \frac{\partial S}{\partial q^i}, \tag{8.8.4}$$

$$q^i = -\frac{\partial S}{\partial p_i}, \qquad \bar{p}_i = -\frac{\partial S}{\partial \bar{q}^i}, \tag{8.8.5}$$

$$q^i = -\frac{\partial S}{\partial p_i}, \qquad \bar{q}^i = \frac{\partial S}{\partial \bar{p}_i}. \tag{8.8.6}$$

An exhaustive discussion of all possible coordinate systems on Γ_Φ, and thus of all generating functions, can be found in the book of Arnold [18].

Conversely, given a smooth function S in the $2n$ variables q^i and $\bar{q}^i$ fulfilling

$$\det\left(\frac{\partial^2 S}{\partial q^i \partial \bar{q}^j}\right) \neq 0, \tag{8.8.7}$$

the relations (8.8.3) define $2n$ functions p_i and $\bar{p}_i$ in these variables. By (8.8.7), the matrix of partial derivatives $\frac{\partial p_i}{\partial \bar{q}^j}$ is invertible. Hence, the Inverse Mapping Theorem implies that the mapping $\bar{\mathbf{q}} \mapsto p(\mathbf{q}, \bar{\mathbf{q}})$ can be locally inverted for every fixed $\mathbf{q}$, thus yielding a mapping $(\mathbf{q}, \mathbf{p}) \mapsto \bar{q}(\mathbf{q}, \mathbf{p})$. By plugging this into the functions $\bar{p}_i$ we finally arrive at a mapping

$$(\mathbf{q}, \mathbf{p}) \mapsto \big(\bar{q}(\mathbf{q}, \mathbf{p}), \bar{p}(\mathbf{q}, \mathbf{p})\big)$$

which via the chosen coordinates q^i, p_i on M_1 and $\bar{q}^i$, $\bar{p}_i$ on M_2 defines a local diffeomorphism $\Phi : M_1 \to M_2$. This diffeomorphism is symplectic:

$$\begin{aligned}
&\iota_\Phi^*\big(\mathrm{d}\bar{p}_i \wedge \mathrm{d}\bar{q}^i - \mathrm{d}p_i \wedge \mathrm{d}q^i\big) \\
&\quad = \left(\frac{\partial \bar{p}_i}{\partial q^j}\mathrm{d}q^j + \frac{\partial \bar{p}_i}{\partial \bar{q}^j}\mathrm{d}\bar{q}^j\right) \wedge \mathrm{d}\bar{q}^i - \left(\frac{\partial p_i}{\partial q^j}\mathrm{d}q^j + \frac{\partial p_i}{\partial \bar{q}^j}\mathrm{d}\bar{q}^j\right) \wedge \mathrm{d}q^i \\
&\quad = \left(\frac{\partial^2 S}{\partial q^j \partial \bar{q}^i}\mathrm{d}q^j + \frac{\partial^2 S}{\partial \bar{q}^j \partial \bar{q}^i}\mathrm{d}\bar{q}^j\right) \wedge \mathrm{d}\bar{q}^i - \left(\frac{\partial^2 S}{\partial q^j \partial q^i}\mathrm{d}q^j + \frac{\partial^2 S}{\partial \bar{q}^j \partial q^i}\mathrm{d}\bar{q}^j\right) \wedge \mathrm{d}q^i \\
&\quad = 0.
\end{aligned}$$

Analogously, by interpreting S as a generating function of the second, third or fourth kind, one obtains local symplectomorphisms defined, respectively, by the relations (8.8.4), (8.8.5) or (8.8.6).

Remark 8.8.2 The definition of generating function given by (8.8.1) generalizes to arbitrary Lagrangian submanifolds (L, i) of a symplectic manifold (M, ω). Since $i^*\omega = 0$, locally there exists a potential τ of ω fulfilling $\mathrm{d}(i^*\tau) = 0$. Thus, the Poincaré Lemma yields an open subset $U \subset L$ and a function $S : U \to \mathbb{R}$ fulfilling (8.8.1). For the special case of a Lagrangian submanifold of a cotangent bundle $M = \mathrm{T}^*Q$ which is transversal to the fibres, the function S can be viewed as a function on Q. This case was dealt with in Proposition 8.3.10.

Example 8.8.3 The concept of generating functions of Lagrangian submanifolds is useful in thermodynamics: for one mol of an ideal gas, the phase space is the open subset $\mathbb{R}^4_+ \subset \mathbb{R}^4$ with global coordinates p, V, T, S and symplectic form

$$\omega = \mathrm{d}V \wedge \mathrm{d}p + \mathrm{d}T \wedge \mathrm{d}S.$$

The equations of state

$$p \cdot V = R \cdot T, \qquad p \cdot V^{\gamma} = k \cdot \exp \frac{S}{c_V}, \qquad c_V = \frac{R}{\gamma - 1}$$

define a 2-dimensional Lagrangian submanifold. All the thermodynamical potentials are generating functions of this Lagrangian submanifold, see [160].

In the remaining part of this section we present some elementary facts concerning the structure of the group of symplectomorphisms $\mathrm{Symp}(M, \omega)$ of a symplectic manifold (M, ω). An exhaustive discussion can be found in [206, Ch. 10]. First, we equip $\mathrm{Symp}(M, \omega)$ with an appropriate topology, called the C^1-topology. To define it, recall that the compact-open topology on the space $C^\infty(N, P)$ of smooth mappings from a smooth manifold N to a smooth manifold P is generated by finite intersections of subsets $V^0(K, U)$, where $K \subset N$ is compact, $U \subset P$ is open and $V^0(K, U)$ consists of the mappings $\varphi : N \to P$ satisfying $\varphi(K) \subset U$. For convenience, below we will refer to this topology as the C^0-topology. Since P, being a manifold, is metrizable, this topology coincides with the topology of uniform convergence on compact sets.[13] Now, the C^1-topology on $C^\infty(N, P)$ is defined to be the initial topology induced by the assignment

$$C^\infty(N, P) \to C^\infty(\mathrm{T}N, \mathrm{T}P), \qquad \varphi \mapsto \varphi',$$

where $C^\infty(\mathrm{T}N, \mathrm{T}P)$ is endowed with the C^0-topology. That means, it is generated by finite intersections of subsets $V^1(K, U)$, where $K \subset \mathrm{T}N$ is compact, $U \subset \mathrm{T}P$ is open and $V^1(K, U)$ is the preimage of $V^0(K, U)$ under the above mapping, that is, it consists of all mappings $\varphi : N \to P$ satisfying $\varphi'(K) \subset U$.[14] Since the projections $\mathrm{T}N \to N$ and $\mathrm{T}P \to P$ preserve compactness and openness, respectively, the C^1-topology is stronger than the C^0-topology. In particular, every C^0-open subset is also C^1-open and C^1-convergence implies C^0-convergence. We note that the composition mapping

$$C^\infty(N, P) \times C^\infty(P, Q) \to C^\infty(N, Q), \qquad (\varphi, \psi) \mapsto \psi \circ \varphi,$$

[13] A sequence of mappings $\varphi_i : N \to P$ converges to a mapping $\varphi : N \to P$ iff for every compact $K \subset N$ and every $\varepsilon > 0$ there exists n_0 such that $\sup_{m \in K} \mathrm{d}(\varphi_n(m), \varphi(m)) < \varepsilon$ for all $n > n_0$; here d is some metric on P, compatible with the topology.

[14] Equivalently, a sequence $\{\varphi_i\}$ converges to φ in the C^1-topology iff φ_i' converges to φ' in the C^0-topology on $C^\infty(\mathrm{T}N, \mathrm{T}P)$.

is continuous in the respective C^0-topologies. The same is true for the inversion mapping $\varphi \mapsto \varphi^{-1}$, defined on the subset of diffeomorphisms from N to P. Due to $(\psi \circ \varphi)' = \psi' \circ \varphi'$ and $(\varphi^{-1})' = (\varphi')^{-1}$, these mappings are also continuous in the C^1-topology. Thus, $\mathrm{Symp}(M, \omega)$ is a topological group in both the C^0- and the C^1-topologies. If M is compact, it can be equipped with the structure of a Fréchet-Lie group, but this requires the use of a more sophisticated topology, which is beyond our scope. We will comment on this below. Rather, as an application of Proposition 8.6.4 we will show that, in the C^1-topology, $\mathrm{Symp}(M, \omega)$ for compact M is locally homeomorphic to the space $Z^1(M)$ of closed 1-forms on M, endowed with the C^1-topology induced from $C^\infty(M, \mathrm{T}^*M)$. This result belongs to Weinstein [303, §6]. By continuity of the group multiplication, it suffices to construct a local homeomorphism in a neighbourhood of id_M.

Proposition 8.8.4 (Weinstein) *Let (M, ω) be a compact symplectic manifold. There exists a homeomorphism from an arcwise connected open C^1-neighbourhood of* id_M *in* $\mathrm{Symp}(M, \omega)$ *onto an arcwise connected open C^1-neighbourhood of the zero* 1-*form in the space of closed one-forms $Z^1(M)$.*

Proof Let $\Delta : M \to M \times M$, $\Delta(m) := (m, m)$, be the diagonal mapping and let $\mathrm{pr}_i : M \times M \to M$ denote the natural projections. For $\varphi \in \mathrm{Symp}(M, \omega)$, let Γ_φ denote the graph of φ. Γ_φ is the image of the graph mapping $\mathrm{gr}_\varphi := (\mathrm{id}_M \times \varphi) \circ \Delta$.

By Proposition 8.8.1, Γ_φ is an embedded Lagrangian submanifold of $M \times M$ for all $\varphi \in \mathrm{Symp}(M, \omega)$. If $\varphi = \mathrm{id}_M$, then Γ_φ coincides with the submanifold (M, Δ). Theorem 8.6.4 implies that there exists a symplectomorphism Ψ from an open neighbourhood U of $\Gamma_{\mathrm{id}_M} = \Delta(M)$ in $M \times M$ onto an open neighbourhood V of the zero section s_0 in T^*M, satisfying

$$\Psi \circ \Delta = s_0.$$

Then, $\Psi(\Gamma_\varphi)$ is a Lagrangian submanifold of T^*M for all $\varphi \in \mathrm{Symp}(M, \omega)$ such that $\Gamma_\varphi \subset U$. By Proposition 8.3.10, $\Psi(\Gamma_\varphi)$ is the image of a closed 1-form on M iff it intersects each fibre of T^*M transversally and exactly once. Thus, to construct the desired mapping from symplectomorphisms to closed 1-forms, it suffices to show that there exists an open neighbourhood $\mathscr{U}$ of id_M in $\mathrm{Symp}(M, \omega)$ in the C^1-topology such that for all $\varphi \in \mathscr{U}$, the graph Γ_φ is contained in U and the Lagrangian submanifold $\Psi(\Gamma_\varphi)$ of T^*M intersects each fibre of T^*M transversally and exactly once. For that purpose, let $\tilde{U} \subset \mathrm{T}(M \times M)$ be obtained from the preimage of U under the natural projection $\mathrm{T}(M \times M) \to M \times M$ by removing the preimage under Ψ' of the vertical distribution on T^*M. This is an open subset of $\mathrm{T}(M \times M)$. Since M is compact, $\mathrm{T}M$ contains a compact subset $\tilde{K}$ which generates $\mathrm{T}M$ under scalar multiplication by real numbers and does not intersect the zero section (e.g., the unit sphere bundle with respect to some Riemannian metric on M). Then, $V^1(\tilde{K}, \tilde{U})$ is an open neighbourhood of Δ in $C^\infty(M, M \times M)$ in the C^1-topology. Define

$$\mathscr{U} := \left\{\varphi \in \mathrm{Symp}(M, \omega) : \mathrm{gr}_\varphi \in V^1(\tilde{K}, \tilde{U})\right\}.$$

Since the assignment $\varphi \mapsto \mathrm{gr}_\varphi$ is the restriction to $\{\Delta\} \times \{\mathrm{id}_M\} \times \mathrm{Symp}(M, \omega)$ of the mapping

$$C^\infty(M, M \times M) \times C^\infty(M, M) \times C^\infty(M, M) \to C^\infty(M \times M, M \times M),$$
$$(\psi_1, \psi_2, \psi_3) \mapsto (\psi_2 \times \psi_3) \circ \psi_1,$$

it is C^1-continuous. Hence, $\mathscr{U}$ is an open neighbourhood of id_M in $\mathrm{Symp}(M, \omega)$ in the C^1-topology. By construction, for all $\varphi \in \mathscr{U}$, the graph Γ_φ is contained in U and $\Psi(\Gamma_\varphi)$ is transversal to the fibres of T^*M. Since M is compact, it is clear that by possibly shrinking $\mathscr{U}$ we may achieve that $\Psi(\Gamma_\varphi)$ intersects the fibres at most once. Then, the mapping $\pi \circ \Psi \circ \mathrm{gr}_\varphi$, with the canonical projection $\pi : \mathrm{T}^*M \to M$, is an injective immersion, and hence an embedding of M into M. Using the theorem on invariance of domain, stated in Footnote 38 on page 159, one can show that the image of this embedding must be M (Exercise 8.8.3). Therefore, $\Psi(\Gamma_\varphi)$ intersects each fibre of T^*M at least once, and hence exactly once. Thus, $\mathscr{U}$ has the desired properties. Now, for every $\varphi \in \mathscr{U}$, Proposition 8.3.10 yields a closed 1-form α on M such that $\Psi(\Gamma_\varphi) = \alpha(M)$. Explicitly, α is given by

$$\alpha = \Psi \circ \mathrm{gr}_\varphi \circ \lambda^{-1} \quad \text{with } \lambda := \pi \circ \Psi \circ \mathrm{gr}_\varphi .$$

The assignment $\varphi \mapsto \alpha$ is C^1-continuous, because it decomposes into a sequence of composition and inversion mappings. To prove that it can be made into a homeomorphism, we show that there exists an open neighbourhood $\mathscr{V}$ of the zero section in $Z^1(M)$ in the C^1-topology such that, for all $\alpha \in \mathscr{V}$, $\alpha(M) \subset V$ and $\Psi^{-1}(\alpha(M))$ intersects each of the submanifolds $\{m\} \times M$ and $M \times \{m\}$, $m \in M$, transversally and exactly once. $\mathscr{V}$ can be constructed in the same way as $\mathscr{U}$, one just replaces $\mathrm{Symp}(M, \omega) \subset C^\infty(M, M)$ by $Z^1(M) \subset C^\infty(M, \mathrm{T}^*M)$, φ by α, gr_φ by $\Psi^{-1} \circ \alpha$ and the preimage of the vertical distribution of T^*M under Ψ' by the two distributions $\ker \mathrm{pr}'_i$, $i = 1, 2$. Then, for all $\alpha \in \mathscr{V}$, the mappings $\mathrm{pr}_i \circ \Psi^{-1} \circ \alpha : M \to M$ are bijective immersions and hence diffeomorphisms. Define

$$\varphi := \mathrm{pr}_2 \circ \Psi^{-1} \circ \alpha \circ \mu^{-1} \quad \text{with } \mu := \mathrm{pr}_1 \circ \Psi^{-1} \circ \alpha.$$

Since the assignment $\alpha \mapsto \varphi$ decomposes into a sequence of composition and inversion mappings, it is C^1-continuous. A straightforward calculation shows that this assignment is inverse to the assignment $\varphi \mapsto \alpha$ constructed above. Finally, we intersect $\mathscr{U}$ with the image of $\mathscr{V}$ and $\mathscr{V}$ with the image of $\mathscr{U}$. Since $Z^1(M)$ is locally arcwise connected, the neighbourhoods so obtained can be shrunk so that they become arcwise connected. □

Proposition 8.8.4 yields

Corollary 8.8.5 *The group of symplectomorphisms of a compact symplectic manifold is locally arcwise connected.*

Remark 8.8.6 Since the group multiplication in $\mathrm{Symp}(M,\omega)$ is C^1-continuous, Proposition 8.8.4 provides an atlas modelling $\mathrm{Symp}(M,\omega)$ for compact M on the infinite-dimensional vector space $Z^1(M) \cong \mathfrak{X}_{\mathrm{LH}}(M,\omega)$ endowed with the C^1-topology.[15] However, since $\mathfrak{X}_{\mathrm{LH}}(M,\omega)$ is not complete in this topology, this does not supply a differentiable structure on $\mathrm{Symp}(M,\omega)$. To obtain completeness and thus a differentiable structure one can either enlarge the spaces by passing to forms and symplectomorphisms of the differentiability class C^1. Then, $\mathrm{Symp}(M,\omega)$ is a smooth manifold modelled on the Banach space of closed 1-forms on M of class C^1. In this setup, the multiplication mapping in $\mathrm{Symp}(M,\omega)$ turns out to be continuous but not differentiable [263, Thm. 2.1]. Or one can construct appropriate topologies, thus turning $\mathfrak{X}_{\mathrm{LH}}(M,\omega)$ into a Fréchet space[16] and $\mathrm{Symp}(M,\omega)$ into a smooth manifold modelled on $\mathfrak{X}_{\mathrm{LH}}(M,\omega)$ [262], see also [82, 213, 230]. In this structure, the multiplication and inversion mappings are smooth and hence $\mathrm{Symp}(M,\omega)$ is an infinite-dimensional Fréchet Lie group.[17] The corresponding Lie algebra can be naturally identified with $\mathfrak{X}_{\mathrm{LH}}(M,\omega)$ with the ordinary commutator of vector fields and the exponential mapping being given by

$$\exp : \mathfrak{X}_{\mathrm{LH}}(M,\omega) \to \mathrm{Symp}(M,\omega), \qquad \exp(X) := \Phi_1^X. \tag{8.8.8}$$

In this sense, independent of topologies or differentiable structures, one may speak of $\mathfrak{X}_{\mathrm{LH}}(M,\omega)$ as the Lie algebra of $\mathrm{Symp}(M,\omega)$. We note that, in contrast to the finite-dimensional case, the exponential mapping (8.8.8) is not a local diffeomorphism between neighbourhoods of the origin in $\mathfrak{X}_{\mathrm{LH}}(M,\omega)$ and the unit element of $\mathrm{Symp}(M,\omega)$, see [99] and [230].

Let us denote the arcwise connected component of $\mathrm{Symp}(M,\omega)$ containing id_M by $\mathrm{Symp}_0(M,\omega)$. This is a normal subgroup of $\mathrm{Symp}(M,\omega)$. Note that if M is compact, Corollary 8.8.5 implies that this is also the connected component of id_M, because in a locally arcwise connected space, the arcwise connected components coincide with the connected components. By definition, for every $\varphi \in \mathrm{Symp}_0(M,\omega)$ there exists a C^1-continuous curve $\Phi : [0,1] \to \mathrm{Symp}_0(M,\omega)$ such that $\Phi_0 = \mathrm{id}_M$ and $\Phi_1 = \varphi$. It is not hard to see that Φ can be chosen so that the induced homotopy $\Phi : [0,1] \times M \to M$ is smooth. Since the latter runs through the diffeomorphisms of M, it is commonly referred to as a smooth isotopy. According to Remark 3.4.5/2, Φ is the flow of the time-dependent symplectic vector field X defined by

$$X_t(m) = \frac{\mathrm{d}}{\mathrm{d}s}\restriction_t \Phi_s \circ \Phi_{-t}(m), \quad m \in M,\ t \in [0,1]. \tag{8.8.9}$$

---

[15] Since M is compact, the C^1-topology on $\mathfrak{X}_{\mathrm{LH}}(M,\omega)$ allows for a norm, defined by taking the maximum over the usual C^1-norms of the local representatives of α in some chosen finite atlas, see e.g. [233, 234].

[16] A complete metrizable locally convex vector space.

[17] A Lie group modelled on a Fréchet space.

By Proposition 8.2.4/4 and Remark 8.2.5/2, X is locally Hamiltonian for every t. This observation motivates the study of isotopies generated by time-dependent vector fields for which X is Hamiltonian for all t.

Definition 8.8.7 (Hamiltonian diffeomorphism) Let (M, ω) be a symplectic manifold.

1. A smooth isotopy $\Phi : [0, 1] \times M \to M$ in $\mathrm{Diff}(M)$ with associated time-dependent vector field X is called Hamiltonian if there exists a smooth function $H : [0, 1] \times M \to \mathbb{R}$ such that $X_t = X_{H_t}$ for all t. We say that Φ is generated by H.
2. A diffeomorphism φ of M is called Hamiltonian if there exists a Hamiltonian isotopy from id_M to φ. The set of Hamiltonian diffeomorphisms is denoted by $\mathrm{Ham}(M, \omega)$.

Obviously, every Hamiltonian diffeomorphism is a symplectomorphism. In fact, we have

Proposition 8.8.8 $\mathrm{Ham}(M, \omega)$ *is a normal subgroup of* $\mathrm{Symp}_0(M, \omega)$.

Proof Let Φ and Ψ be Hamiltonian isotopies, generated by the time-dependent Hamiltonians H and K, respectively. Using (8.2.4), we compute

$$\frac{\mathrm{d}}{\mathrm{d}t}(\Psi_t \circ \Phi_t)(m) = \big(X_{K_t} + (\Psi_t)_* X_{H_t}\big)\big(\Psi_t \circ \Phi_t(m)\big) = X_{K_t + H_t \circ \Psi_t^{-1}}\big(\Psi_t \circ \Phi_t(m)\big).$$

Hence, $(t, m) \mapsto \Psi_t \circ \Phi_t(m)$ is a Hamiltonian isotopy, generated by the Hamiltonian $(t, m) \mapsto K_t(m) + H_t \circ \Psi_t^{-1}(m)$. Similarly, one shows that $(t, m) \mapsto \Phi_t^{-1}(m)$ and $(t, m) \mapsto \varphi \circ \Phi_t \circ \varphi^{-1}(m)$, for any symplectomorphism φ, are Hamiltonian isotopies generated by the time-dependent Hamiltonians $(t, m) \mapsto -H_t \circ \Phi_t(m)$ and $(t, m) \mapsto H_t \circ \varphi^{-1}(m)$, respectively (Exercise 8.8.4). □

As a consequence of Proposition 8.8.8, one has the following sequence of subgroup inclusion mappings:

$$\mathrm{Ham}(M, \omega) \hookrightarrow \mathrm{Symp}_0(M, \omega) \hookrightarrow \mathrm{Symp}(M, \omega) \hookrightarrow \mathrm{Diff}^+(M) \hookrightarrow \mathrm{Diff}(M), \tag{8.8.10}$$

with $\mathrm{Diff}^+(M)$ denoting the group of diffeomorphisms which preserve the orientation defined by the natural volume form of ω. This sequence is studied in symplectic topology.

Next, we show that the local homeomorphism of Proposition 8.8.4 maps Hamiltonian isotopies to families of exact 1-forms. We start with a lemma on exact symplectic manifolds (which by Remark 8.1.4/2 are necessarily noncompact).

Lemma 8.8.9 *Let* $(M, \mathrm{d}\sigma)$ *be an exact symplectic manifold. An isotopy* Φ *in* $\mathrm{Diff}(M)$ *with* $\Phi_0 = \mathrm{id}_M$ *is Hamiltonian iff there exists a smooth function*

$f : [0, 1] \times M \to \mathbb{R}$ *such that for every* $t \in [0, 1]$,

$$\Phi_t^* \sigma - \sigma = \mathrm{d} f_t. \tag{8.8.11}$$

Proof Let X be the time-dependent vector field associated with Φ, cf. Remark 3.4.5/2. Using (4.1.28), we compute

$$\frac{\mathrm{d}}{\mathrm{d}t} \Phi_t^* \sigma = \Phi_t^* \mathscr{L}_{X_t} \sigma = \Phi_t^* \big(X_t \lrcorner \mathrm{d}\sigma + \mathrm{d}(X_t \lrcorner \sigma) \big). \tag{8.8.12}$$

If Φ is Hamiltonian with generating time-dependent Hamiltonian H, this implies

$$\frac{\mathrm{d}}{\mathrm{d}t} \Phi_t^* \sigma = \Phi_t^* \big(\mathrm{d}(X_{H_t} \lrcorner \sigma) - \mathrm{d} H_t \big).$$

Integration of these two families of 1-forms yields (8.8.11) with

$$f_t = \int \big((X_t \lrcorner \sigma - H_t) \circ \Phi_t \big) \mathrm{d}t.$$

Conversely, if f is given such that (8.8.11) holds, for every t we define

$$H_t := X_t \lrcorner \sigma - \big(\Phi_t^{-1} \big)^* \left(\frac{\mathrm{d}}{\mathrm{d}t} f_t \right).$$

This yields a smooth function $H : [0, 1] \times M \to M$. A straightforward calculation using (4.1.30) and (8.8.12) shows that $X_{H_t} = X_t$ for all t. Hence, Φ is Hamiltonian with generating time-dependent Hamiltonian H. □

Proposition 8.8.10 *Let* (M, ω) *be a compact symplectic manifold. The local homeomorphism of Proposition* 8.8.4 *maps Hamiltonian isotopies of* M *to smooth families of exact* 1*-forms on* M *and vice versa.*

Proof We adopt the notation of the proof of Proposition 8.8.4. According to that proof, for every smooth isotopy Φ in $\mathscr{U}$ there exists a smooth family $\{\lambda_t\}$ of diffeomorphisms of M and a smooth family $\{\alpha_t\}$ of closed 1-forms on M such that

$$\Psi \circ \mathrm{gr}_{\Phi_t} = \alpha_t \circ \lambda_t$$

for all t.[18] Denote the canonical 1-form on T^*M by θ. Since $\Psi^*\theta$ is a potential for ω on $U \subset M \times M$ and since $\mathrm{id}_M \times \Phi$ is a smooth isotopy of the exact symplectic manifold (U, ω), Lemma 8.8.9 yields that $\mathrm{id}_M \times \Phi$, and hence Φ, is Hamiltonian iff there exists a smooth function $f : [0, 1] \times M \times M \to \mathbb{R}$ such that for all t

$$(\mathrm{id}_M \times \Phi_t)^* \circ \Psi^*(\theta) - \Psi^*\theta = \mathrm{d} f_t.$$

[18] With Φ_t playing the role of φ in the proof of Proposition 8.8.4.

Applying Δ^* to this equation and using $\Delta^* \circ \Psi^*(\theta) = s_0^*\theta = 0$, we obtain

$$\mathrm{gr}_{\Phi_t}^* \circ \Psi^*(\theta) = \Delta^* \circ (\mathrm{id} \times \Phi_t)^* \circ \Psi^*(\theta) = \mathrm{d}(f_t \circ \Delta).$$

By the defining property (8.3.2) of θ, the left hand side yields

$$\mathrm{gr}_{\Phi_t}^* \circ \Psi^*(\theta) = \lambda_t^*\big(\alpha_t^*\theta\big) = \lambda_t^*\alpha_t,$$

so that $\alpha_t = \mathrm{d}(f_t \circ \Delta \circ \lambda_t^{-1})$. Thus, Φ is Hamiltonian iff α_t is exact for all t. □

As a consequence of Proposition 8.8.10, if the first de Rham cohomology group $H^1(M)$ of M is trivial, $\mathrm{Ham}(M,\omega)$ coincides with $\mathrm{Symp}_0(M,\omega)$, that is, in this case, the first inclusion mapping in the sequence (8.8.10) is surjective and, by Remark 8.8.6, the Lie algebra of $\mathrm{Ham}(M,\omega)$ may be identified with $\mathfrak{X}_{\mathrm{H}}(M,\omega)$. For an arbitrary compact manifold, the situation is more complicated. The homeomorphism of Proposition 8.8.4 need not always map Hamiltonian diffeomorphisms to exact 1-forms, no matter how C^1-close to id_M they are. Rather, one can show that a symplectomorphism φ in the domain of this homeomorphism is Hamiltonian iff the cohomology class of its image belongs to a certain countable subgroup of $H^1(M)$, called the flux group [206, Lemma 10.16]. This group is the image of the fundamental group of $\mathrm{Symp}_0(M,\omega)$ under the so-called flux homomorphism which assigns to the homotopy class of a closed curve in $\mathrm{Symp}_0(M,\omega)$, represented by a symplectic isotopy Φ, the cohomology class of the closed 1-form $\int (X_t \lrcorner \omega)\mathrm{d}t$. Here, X is the time-dependent vector field associated with Φ via (8.8.9), the integral is defined pointwise and the resulting 1-form is indeed closed due to (4.1.30). For details, like the proof that this mapping is well-defined and a group homomorphism, see [206, Chap. 10] or [242, Chap. 14]. As a consequence of the facts that the homeomorphism of Proposition 8.8.4 maps Hamiltonian diffeomorphisms to the flux group and that the latter is countable, one obtains

Proposition 8.8.11 *If M is compact, every smooth isotopy in* $\mathrm{Ham}(M,\omega)$ *is Hamiltonian.*

Proof See [206, Prop. 10.17]. If the first de Rham cohomology group of M is trivial, the assertion follows from Proposition 8.8.10 (Exercise 8.8.5). □

Remark 8.8.12 Let (M,ω) be a compact symplectic manifold.

1. On the basis of Proposition 8.8.11, one may give the following intuitive argument showing that the tangent space at id_M of $\mathrm{Ham}(M,\omega)$ is given by $\mathfrak{X}_{\mathrm{H}}(M,\omega)$ [242, §1.4]: every curve in $\mathrm{Ham}(M,\omega)$ through id_M is a Hamiltonian isotopy and hence its tangent vector at id_M is a Hamiltonian vector field. Conversely, the flow of any Hamiltonian vector field yields a Hamiltonian isotopy and thus a curve in $\mathrm{Ham}(M,\omega)$ through id_M. Thus, according to Remark 8.8.6 and by analogy with the finite-dimensional situation of Proposition 5.6.5, we conclude that

Ham(M,ω) is the Lie subgroup of Symp(M,ω) associated with the Lie subalgebra $\mathfrak{X}_{\mathrm{H}}(M,\omega)$ of $\mathfrak{X}_{\mathrm{LH}}(M,\omega)$. It was proved by Ono [231] that the flux group is discrete.[19] Thus, one may shrink the C^1-neighbourhood of id_M of Proposition 8.8.4 so that all Hamiltonian diffeomorphisms in that neighbourhood are mapped to exact forms. Hence, for compact M, Ham(M,ω) is a C^1-closed Lie subgroup of Symp$_0(M,\omega)$.

2. If M is connected, the Hamiltonian function corresponding to a Hamiltonian vector field is unique up to a constant. One way to obtain uniqueness consists in considering the space $\mathscr{A}(M)$ of smooth functions on M with zero mean[20] with respect to the canonical volume form Ω of ω, see [242]. This is a Poisson subalgebra of $C^\infty(M)$ (Exercise 8.8.6). Thus, in this case the Lie algebra of Ham(M,ω) can be identified with $\mathscr{A}(M)$. From the proof of Proposition 8.8.8 we read off that the adjoint action of Ham(M,ω) on $\mathscr{A}(M)$ is given by

$$\mathrm{Ham}(M,\omega)\times\mathscr{A}(M)\to\mathscr{A}(M),\qquad (\Phi,f)\mapsto f\circ\Phi^{-1}. \tag{8.8.13}$$

The following fundamental algebraic statements about Ham(M,ω) were proved by Banyaga [30, 31].

Proposition 8.8.13 (Banyaga)

1. *If the symplectic manifold* (M,ω) *is compact*, Ham(M,ω) *is a simple*[21] *group*.
2. *If two symplectic manifolds* (M_1,ω_1) *and* (M_2,ω_2) *have isomorphic groups of Hamiltonian symplectomorphisms*, *there exists a diffeomorphism* $\Phi: M_1\to M_2$ *and a number* $c\neq 0$ *such that* $\Phi^*\omega_2=c\omega_1$.

The second statement says that the algebraic structure of the group of Hamiltonian diffeomorphisms determines the symplectic structure up to a constant factor.

Remark 8.8.14 In case M is not compact, the above results on Symp(M,ω) and Ham(M,ω) for compact M carry over to the group Symp$^c(M,\omega)$ of compactly supported symplectomorphisms[22] and the subgroup Ham$^c(M,\omega)$ of elements which can be joined to id_M by a Hamiltonian isotopy whose generating time-dependent Hamiltonian has compact support. Here, the topology is defined to be the inductive limit, taken over the directed set of compact subsets K of M, of the topologies on the subgroups of symplectomorphisms with support in K. Correspondingly, one has to use compactly supported forms. Thus, in particular, Symp$^c(M,\omega)$ is locally homeomorphic to the space of compactly supported closed 1-forms and its Lie algebra consists of compactly supported symplectic vector fields. The Lie algebra of

---

[19] This was the affirmative answer to the so-called flux conjecture, see also [179]. For noncompact M, the flux group need not be discrete, see [205].

[20] That is, $\int_M f\,\Omega=0$.

[21] A group G which does not contain normal subgroups besides $\{\mathbb{1}\}$ and G.

[22] That is, symplectomorphisms which outside a compact set coincide with the identical mapping.

$\mathrm{Ham}^c(M, \omega)$ coincides with the space of compactly supported Hamiltonian vector fields. It can be identified with the space $\mathscr{A}^c(M)$ of compactly supported functions on M with zero mean. For details we refer to [206] and [242]. An exhaustive treatment of the case $M = \mathbb{R}^{2n}$ can be found in [139].

We conclude this section with a remark on the geometric structure of $\mathrm{Ham}(M, \omega)$. This structure will show up again in the next chapter. Assume that M is noncompact and consider the group $\mathrm{Ham}^c(M, \omega)$ of compactly supported Hamiltonian symplectomorphisms. In 1990, Hofer [134] observed that there exists a norm on the Lie algebra $\mathscr{A}^c(M)$ of $\mathrm{Ham}^c(M, \omega)$. It can be defined by

$$\|H\| := \sup_{m\in M} H(m) - \inf_{m\in M} H(m). \tag{8.8.14}$$

This norm is invariant under the adjoint action (8.8.13). The corresponding length function for a smooth Hamiltonian isotopy Φ in $\mathrm{Ham}^c(M, \omega)$ generated by the time-dependent Hamiltonian H is given by

$$l(\Phi) = \int_0^1 \|H_t\| \mathrm{d}t$$

and the distance between two Hamiltonian diffeomorphisms φ and ψ is defined by

$$\rho(\varphi, \psi) = \inf\bigl(l(\Phi)\bigr),$$

with the infimum taken over all Hamiltonian isotopies Φ fulfilling $\Phi_0 = \varphi$ and $\Phi_1 = \psi$. It is easy to show that ρ defines a bi-invariant pseudo-metric on $\mathrm{Ham}^c(M, \omega)$, but it is hard to prove that $\rho(\varphi, \psi) = 0$ implies $\varphi = \psi$. See [134] for the case $M = \mathbb{R}^{2n}$ and [178] for the general case. For a further discussion of the Hofer metric we refer to the book of Polterovich [242]. The case $M = \mathbb{R}^{2n}$ is dealt with in great detail in the book of Hofer and Zehnder, see [139]. The Hofer norm plays a fundamental role in symplectic topology and in the study of global existence questions in the theory of Hamiltonian systems. In particular, it gives rise to a certain symplectic invariant, the so-called Hofer-Zehnder capacity, which turns out to be one of the basic tools for studying the Weinstein conjecture, see Sect. 9.4.

Exercises

8.8.1 Let $(M_1, \omega_1) = (\mathbb{R} \times \mathbb{R}_+, \mathrm{d}p \wedge \mathrm{d}q)$ and $(M_2, \omega_2) = (\mathbb{R}^2 \setminus \{0\}, \mathrm{d}\bar{p} \wedge \mathrm{d}\bar{q})$ and let $\Phi\colon M_1 \to M_2$, $\Phi(q, p) = (\bar{q}, \bar{p})$ be given by

$$\bar{q} = \sqrt{\frac{p}{\pi\omega}} \sin(2\pi q), \qquad \bar{p} = \sqrt{\frac{p\omega}{\pi}} \cos(2\pi q).$$

(a) Show that the restrictions of Φ to the subsets $(x, x+1) \times \mathbb{R}_+ \subset M_1$, $x \in \mathbb{R}$, are symplectomorphisms onto their images. Determine these images.
(b) Find the generating function of the first kind for the restriction of Φ to $(0, 1) \times \mathbb{R}_+$ and the generating function of the second kind for the restriction to $(-\frac{1}{2}, \frac{1}{2}) \times \mathbb{R}_+$.

8.8.2 Prove that the equations of state of an ideal gas define a Lagrangian submanifold of the symplectic manifold postulated in Example 8.8.3.
8.8.3 Use the theorem on invariance of domain, stated in Footnote 38 on page 159, to prove that every embedding of a compact manifold into another compact manifold of the same dimension and the same number of connected components is surjective.
8.8.4 Complete the proof of Proposition 8.8.8.
8.8.5 Use Proposition 8.8.10 to prove Proposition 8.8.11 under the assumption that the first de Rham cohomology group of M is trivial.
8.8.6 Let (M,ω) be a symplectic manifold. Show that the space $\mathscr{A}(M)$ of smooth functions on M with zero mean with respect to the canonical volume form of ω is a Poisson subalgebra of $C^\infty(M)$.

8.9 Elementary Morse Theory

In this section, we will discuss some elements of Morse theory, which will be used later on. Let $f : M \to \mathbb{R}$ be a smooth function. Recall that $m \in M$ is a critical point of f iff

$$\mathrm{d}f(m) = 0.$$

Let M_f be the set of critical points of f. The image $\mathrm{d}f(M)$ of the 1-form $\mathrm{d}f$ and the image $s_0(M)$ of the zero section s_0 in T^*M are embedded submanifolds of T^*M which intersect over every critical point.

Definition 8.9.1 (Morse function) A smooth function $f\colon M \to \mathbb{R}$ is called a Morse function if $\mathrm{d}f(M)$ and $s_0(M)$ are transversal, that is, if

$$\mathrm{T}_\xi\big(\mathrm{T}^*M\big) = \mathrm{T}_\xi\big(\mathrm{d}f(M)\big) + \mathrm{T}_\xi\big(s_0(M)\big) \tag{8.9.1}$$

for all $\xi \in \mathrm{d}f(M) \cap s_0(M)$.

By Corollary 1.8.5, for a Morse function f the intersection $\mathrm{d}f(M) \cap s_0(M)$ is an embedded submanifold. Since $\mathrm{d}f(M)$ and $s_0(M)$ are Lagrange, it has dimension zero. Thus, for a Morse function f all critical points are isolated. In particular, if M is compact, the number of critical points of a Morse function is finite. Moreover, for every critical point m, the tangent spaces of $\mathrm{d}f(M)$ and $s_0(M)$ at $s_0(m)$ are complementary Lagrangian subspaces of $\mathrm{T}_{s_0(m)}(\mathrm{T}^*M)$.

Let us derive a criterion for transversality. According to Proposition 8.3.12, for every $m \in M_f$, the zero section s_0 induces a natural symplectomorphism

$$\Phi_{s_0(m)}\colon \mathrm{T}_{s_0(m)}\mathrm{T}^*M \to \mathrm{T}_mM \oplus \mathrm{T}^*_mM. \tag{8.9.2}$$

Let $\mathrm{pr}_2\colon \mathrm{T}_mM \oplus \mathrm{T}^*_mM \to \mathrm{T}^*_mM$ be the projection onto the second component of the direct sum. The tangent mapping

$$(\mathrm{d}f)'_m\colon \mathrm{T}_mM \to \mathrm{T}_{s_0(m)}\mathrm{T}^*M$$

induces a linear mapping

$$\mathrm{Hess}_m(f)\colon \mathrm{T}_m M \to \mathrm{T}_m^* M, \qquad \mathrm{Hess}_m(f) := \mathrm{pr}_2 \circ \Phi_{s_0(m)} \circ (\mathrm{d}f)'_m. \tag{8.9.3}$$

This mapping or, equivalently, the corresponding bilinear form

$$\mathrm{Hess}_m(f)\colon \mathrm{T}_m M \times \mathrm{T}_m M \to \mathbb{R}$$

is called the Hessian of f at the critical point m. Obviously, $\mathrm{d}f(M)$ and $s_0(M)$ are transversal at a critical point $m \in M_f$ iff $(\mathrm{d}f)'_m$ has maximal rank, that is, iff the Hessian of f is non-degenerate at that point. Such a critical point is called non-degenerate.

Let $X, Y \in \mathrm{T}_m M$. Using $\Phi_{s_0(m)} \circ (s_0)'_m(Y) = (Y, 0)$ and that $\Phi_{s_0(m)}$ is symplectic, we obtain

$$\mathrm{Hess}_m(f)(X, Y) = \omega_{s_0(m)}\big((\mathrm{d}f)'_m(X), (s_0)'_m(Y)\big). \tag{8.9.4}$$

Since the images of $\mathrm{d}f$ and s_0 and the fibres of T^*M are Lagrange, we have

$$\omega_{s_0(m)}\big((\mathrm{d}f)'_m(X), (s_0)'_m(Y)\big) = \omega_{s_0(m)}\big((\mathrm{d}f)'_m(Y), (s_0)'_m(X)\big),$$

that is, the Hessian is symmetric. In a local chart (U, κ) at m, we obtain

$$\mathrm{Hess}_m(f)(\partial_i, \partial_j) = (\partial_i \partial_j f)(m) \equiv \frac{\partial^2 (f \circ \kappa^{-1})}{\partial x^i \partial x^j}\big(\kappa(m)\big).$$

By the Theorem of Sylvester, for every symmetric bilinear form H on a vector space V there exists a maximal subspace of V on which H is negative definite. The dimension of this subspace is called the index of H.

Definition 8.9.2 (Morse index) Let $f\colon M \to \mathbb{R}$ be a Morse function and let $m \in M_f$. The index $\mathrm{ind}_f(m)$ of f at m is defined as the index of $\mathrm{Hess}_m(f)$.

Theorem 8.9.3 (Morse Lemma) *Let $f\colon M \to \mathbb{R}$ be a smooth function and let $m_0 \in M_f$ be a non-degenerate critical point. Then, there exists a local chart mapping m_0 to* 0 *such that the local representative of f is given by*

$$f(\mathbf{x}) = f(0) + \frac{1}{2}\mathrm{Hess}_0(f)(\mathbf{x}, \mathbf{x}). \tag{8.9.5}$$

The following proof is due to Palais [235]. We follow the presentation in [32].

Proof As in the proof of the Darboux-Theorem and its generalizations, we use the deformation method of Moser. Since the statement is local, without loss of generality, we may assume $M = \mathbb{R}^n$, $m_0 = 0$ and $f(m_0) = 0$. Denote

$$h(\mathbf{x}) = \frac{1}{2}\mathrm{Hess}_0(f)(\mathbf{x}, \mathbf{x}), \qquad f_t = tf + (1-t)h.$$

We seek a time-dependent vector field X whose flow $\Phi_t \equiv \Phi_{t,0}$ satisfies $\Phi_t^* f_t = h$. Then, Φ_1^{-1} yields the desired local chart. Differentiating this equation with respect to t, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi_t^* f_t = \Phi_t^*\big(X_t(f_t)\big) + \Phi_t^*\left(\frac{\mathrm{d}}{\mathrm{d}t} f_t\right) = \Phi_t^*\big(X_t(f_t) + f - h\big) = 0.$$

Hence, X_t must satisfy

$$X_t(f_t) = h - f. \tag{8.9.6}$$

The left hand side of this equation can be rewritten as follows: expand $X_t = X_t^i \frac{\partial}{\partial x^i}$ and integrate the equality

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(X_t^i(\mathbf{x})\frac{\partial f_t}{\partial x^i}(s\mathbf{x})\right) = X_t^i(\mathbf{x})\frac{\partial^2 f_t}{\partial x^i \partial x^j}(s\mathbf{x})x^j$$

with respect to s from 0 to 1 to obtain

$$\big(X_t(f_t)\big)(\mathbf{x}) = B^t_{ij}(\mathbf{x})X_t^i(\mathbf{x})x^j, \qquad B^t_{ij}(\mathbf{x}) = \int_0^1 \frac{\partial^2 f_t}{\partial x^i \partial x^j}(s\mathbf{x})\mathrm{d}s.$$

For the right hand side of (8.9.6), Taylor's Theorem yields

$$(h - f)(\mathbf{x}) = F_{ij}(\mathbf{x})x^i x^j$$

on some neighbourhood U of the origin, where F_{ij} are smooth functions on U. Thus, (8.9.6) holds on U if the functions X_t^i satisfy

$$B^t_{ij}(\mathbf{x})X_t^i(\mathbf{x}) = F_{ij}(\mathbf{x})x^i$$

for all j. Since

$$B^t_{ij}(0) = \frac{\partial^2 f_t}{\partial x^i \partial x^j}(0) = \mathrm{Hess}_0(f)_{ij}$$

is invertible for all t, and since $[0, 1]$ is compact, by possibly shrinking U we can achieve that $B^t_{ij}(\mathbf{x})$ can be inverted for all $\mathbf{x} \in U$ and $t \in [0, 1]$ and thus obtain the desired vector field X satisfying (8.9.6). Finally, since $X_t(0) = 0$, and again since $[0, 1]$ is compact, by possibly further shrinking U we can achieve that U is contained in the domain of Φ_t for all $t \in [0, 1]$. This completes the proof. □

By successively diagonalizing the symmetric bilinear form $\mathrm{Hess}_0(f)$, the local representative of f can be brought to the following canonical form (Exercise 8.9.1):

Corollary 8.9.4 *Let $f : M \to \mathbb{R}$ be a smooth function, let $m_0 \in M_f$ be a non-degenerate critical point and let i be the index of f at m_0. Then, there exists a local chart mapping m_0 to 0 such that the local representative of f is given by*

$$f(\mathbf{x}) = f(0) - \frac{1}{2}\big(x_1^2 + \cdots + x_i^2\big) + \frac{1}{2}\big(x_{i+1}^2 + \cdots + x_n^2\big). \tag{8.9.7}$$

Remark 8.9.5 The set of Morse functions is open and dense in the space of smooth functions [212, §6]. In particular, Morse functions exist on every manifold.

Next, we will see how the critical points of a smooth function $f\colon M \to \mathbb{R}$ are related to the topology of the subsets

$$M^a := f^{-1}(-\infty, a] = \{m \in M : f(m) \le a\} \tag{8.9.8}$$

and

$$M^{[a,b]} := f^{-1}[a, b] = \{m \in M : a \le f(m) \le b\}. \tag{8.9.9}$$

If a is a regular value of f, the Level Set Theorem yields that $f^{-1}(a)$ is an embedded submanifold of M of codimension 1. Moreover, $f^{-1}(-\infty, a)$ is an open submanifold of M. Using local charts on M mapping $f^{-1}(a)$ to the subspace $\{0\} \times \mathbb{R}^{n-1}$ and local charts on $f^{-1}(-\infty, a)$ taking values in $\mathbb{R}^n_-$, one can construct an atlas on M^a with values in $\mathbb{R}^n_-$ and thus show that M^a is a manifold with boundary, where the boundary is given by $f^{-1}(a)$. Analogously, $M^{[a,b]}$ is a manifold with boundary, provided neither a nor b are critical values of f, and the boundary is given by $f^{-1}(a) \cup f^{-1}(b)$. If M is compact, it is reasonable to imagine f as a height function: while the parameter runs through the real numbers (starting at $-\infty$), M^a grows (starting from the empty set) to the whole manifold. Using methods of Morse theory, it is possible to describe this process of growing. The simplest statement of this type is

Proposition 8.9.6 (Morse Isotopy Lemma) *Let M be a smooth manifold and let $f\colon M \to \mathbb{R}$ be a smooth function. Let $a < b$ and assume that $M^{[a,b]}$ is compact and does not contain any critical point of f. Then, M^a is diffeomorphic to M^b.*

Proof Choose a Riemannian metric g on M and consider the gradient vector field ∇f of f, defined by

$$\nabla f := \mathrm{g}^{-1} \circ (\mathrm{d}f),$$

cf. Formula (4.5.14). Then,

$$\mathrm{g}(\nabla f, X) = X(f) \tag{8.9.10}$$

for all $X \in \mathfrak{X}(M)$. Since $[a, b]$ does not contain any critical value, ∇f is nowhere vanishing on $M^{[a,b]}$. This remains so on some neighbourhood U of $M^{[a,b]}$ in M. Since $M^{[a,b]}$ is compact, U can be chosen to have compact closure. On U, we can define

$$\hat{X}_m := \frac{\nabla f(m)}{\| \nabla f(m) \|^2}, \quad m \in U. \tag{8.9.11}$$

To extend $\hat{X}$ to a vector field on M, we choose a smooth function $\chi : M \to \mathbb{R}$ satisfying $\chi_{\restriction M^{[a,b]}} = 1$ and $\chi_{\restriction (M \setminus U)} = 0$ and define a vector field Y on M by

$$Y_m := \begin{cases} \chi(m)\hat{X}_m & m \in U, \\ 0 & m \notin U. \end{cases}$$

Since U has compact closure and hence Y has compact support, the flow Φ of Y is complete. Using (8.9.10), for $m \in M^{[a,b]}$ we compute

$$\frac{\mathrm{d}(f \circ \Phi_t)(m)}{\mathrm{d}t} = Y_m(f) = \hat{X}_m(f) = 1.$$

Thus, $f(\Phi_t(m)) = f(m) + t$. We conclude that $\Phi_{(b-a)}\colon M \to M$ is a diffeomorphism mapping M^a onto M^b. □

Remark 8.9.7 By means of the flow Φ constructed in the proof of Proposition 8.9.6, one can define a mapping

$$h : [0, 1] \times M^b \to M^b, \qquad h(t, m) := \begin{cases} m & m \in M^a, \\ \Phi_{t(a-f(m))}(m) & m \in M^b \setminus M^a. \end{cases}$$

This mapping is continuous and satisfies

$$h(0, \cdot) = \mathrm{id}_{M^b}, \qquad h\big(1, M^b\big) \subset M^a, \qquad h(t, \cdot)_{\restriction M^a} = \mathrm{id}_{M^a}.$$

One says that h is a strong deformation retraction from M^b to M^a and that M^a is a strong deformation retract of M^b.

Next, we discuss the question how the topological structure of M^b differs from that of M^a if a and b are regular values and $M^{[a,b]}$ contains a single critical point. In general, this question can be answered only up to homotopy-equivalence, that is, the best one can say is how the homotopy type of M^b differs from that of M^a:

Proposition 8.9.8 *Let M be a smooth manifold and let $f\colon M \to \mathbb{R}$ be a smooth function. Let $a < b$ be regular values of f and assume that $M^{[a,b]}$ is compact and contains a single critical point m of f. Let $i := \mathrm{ind}_m(f)$. Define subsets $\tilde{M}^a$ and $\tilde{M}^b$ of M as follows: choose a local chart (U, κ) at m such that the local representative of f has the canonical form given in Corollary 8.9.4, choose $\varepsilon > 0$ such that $\kappa(U)$ contains the ball of radius 2ε about the origin and let*

$$\tilde{M}^a := M^{f(m)-\varepsilon^2}, \qquad \tilde{M}^b := \tilde{M}^a \cup \kappa^{-1}\big(\mathrm{e}^i_{2\varepsilon}\big),$$

where

$$\mathrm{e}^i_{2\varepsilon} := \big\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| < 2\varepsilon, x_{i+1} = \cdots = x_n = 0\big\},$$

see Fig. 8.1. Then, M^a is homotopy-equivalent to $\tilde{M}^a$ and M^b is homotopy-equivalent to $\tilde{M}^b$ (in the relative topology induced from M).

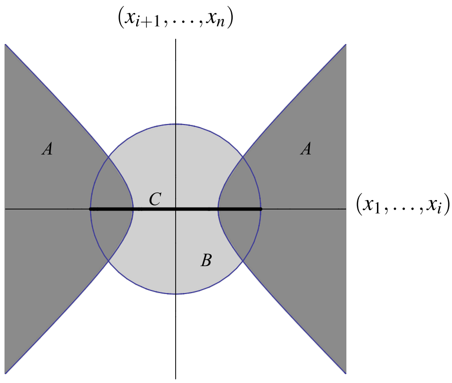


Fig. 8.1 Construction of the subset $\tilde{M}^b$ in Proposition 8.9.8 as the union of the subset $\tilde{M}^a$, labelled by A, and the i-cell $e^i_{2\varepsilon}$, labelled by C. B denotes the ball of radius 2ε about the origin

Proof See [212], Theorem 3.2. □

The set $e^i_{2\varepsilon}$ is called a cell of dimension i and the process of passing from $\tilde{M}^a$ to $\tilde{M}^b$ is commonly referred to as attaching a cell of dimension i. While $\tilde{M}^a$ is a manifold with boundary, $\tilde{M}^b$ need not be so. It is however obvious that the cell $e^i_{2\varepsilon}$ may be thickened and $\tilde{M}^b$ may be smoothened so as to render a manifold with boundary.

In case M is compact and hence every smooth function is bounded, Propositions 8.9.6 and 8.9.8 can be used to construct a topological space which is homotopy-equivalent to M as follows. Choose f so that for every singular value there is exactly one critical point. Denote the singular values by a_i with $a_1 < \cdots < a_r$. Starting at $a = -\infty$ with the empty set, successively attach an appropriate cell each time a passes a singular value. In more detail, since a_1 is a minimum, we must take a zero-dimensional cell $*$. For $a_1 \leq a < a_2$, each M^a is homotopy equivalent to this cell. When passing the singular value a_2 we must attach a cell of dimension equal to the index of the corresponding critical point by identifying the boundary of the closure of this cell with $*$. Continuing this procedure, we obtain a topological space which is called a cell complex, or more precisely a CW-complex, see [55]. In the last step, since a_r is a maximum, we have to attach a cell of the dimension of M. Using this construction, one can show that every compact manifold is homotopy-equivalent to a cell complex with one cell in dimension i for each critical point of index i of some Morse function on M. This result carries over to noncompact manifolds. For a thorough discussion, see [212], §3 and §6.

Example 8.9.9 We consider an upright 2-torus T^2 with the height function f as shown in Fig. 8.2, thereby leaving the necessary computations to the reader (Exercise 8.9.3). There are four critical values $a_1, \dots, a_4$, corresponding to four critical points $m_1, \dots, m_4$. The critical point m_1 is a minimum and hence has index 0, m_2 and m_3 are saddle points and hence have index 1 and m_4 is a maximum and hence has index 2. Thus, starting with a single point (0-cell), first we have to attach a 1-cell, then once again a 1-cell, and finally a 2-cell, see the figure. The way how each

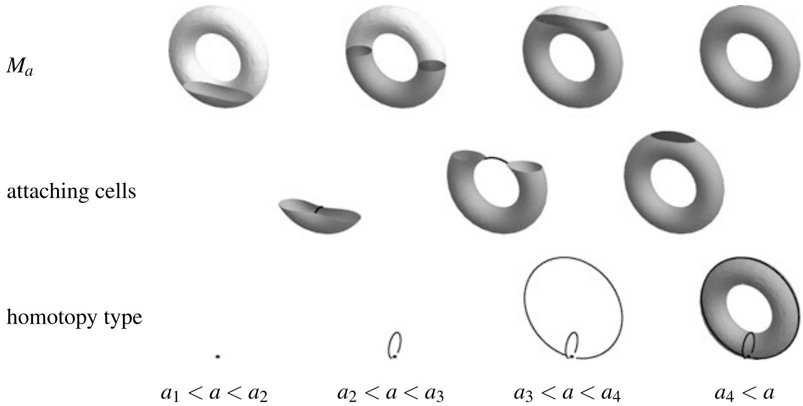


Fig. 8.2 The subsets M_a, their homotopy types in terms of cell complexes and the operations of attaching cells for T^2, cf. Example 8.9.9

of these cells is attached to the cell complex obtained before follows from the way this cell is attached to the corresponding subset M_a according to Proposition 8.9.8.

Remark 8.9.10

1. In some situations, the indices of the critical points of a Morse function contain more information than just the homotopy type of M. For example, the Reeb Theorem states that if M is a compact manifold and if it admits a Morse function with exactly two critical points, then M is homeomorphic to the sphere $\mathrm{S}^{\dim M}$, see e.g. [212], §4. It need not be diffeomorphic to $\mathrm{S}^{\dim M}$, though.
2. By means of a so-called Morse-Smale pair[23] (f, g) one can construct a homology theory for M, called Morse homology, see for example [51] or [146]. The chains of this homology theory are generated by the critical points of f, ordered by their index, and the boundary operator assigns to a critical point m_0 of index i the sum over all critical points m of index $i-1$, weighted with the signed number of flow lines between m_0 and m. It can be shown [146] that Morse homology coincides with singular homology.[24] In particular, it does not depend on the chosen Morse-Smale pair (f, g). These facts indicate that Morse theory yields a mighty tool for investigating the topology of manifolds. For an axiomatic approach to Morse homology in the sense of Eilenberg and Steenrod we refer to [268].
3. As a consequence of either the construction of a homotopy-equivalent cell complex for M or of the equivalence of Morse homology and singular homology, one obtains the Morse inequalities. Let $b_i = \dim H^i(M)$ be the Betti numbers

[23] That is, f is a Morse function and g is a Riemannian metric on M such that for every pair of critical points m_1, m_2 the stable manifold of m_1 with respect to the gradient vector field ∇f is transversal to the unstable manifold of m_2.

[24] See e.g. [55] for this notion.

of M, see Sect. 4.3, and let m_i be the number of critical points with index i. Let $n = \dim M$. Then:

$$m_i \geq b_i,$$
$$m_k - m_{k-1} + \cdots + (-1)^k m_0 \geq b_k - b_{k-1} + \cdots + (-1)^k b_0,$$
$$\sum_{i=0}^{n} (-1)^i m_i = \sum_{i=0}^{n} (-1)^i b_i.$$

4. There are interesting applications of Morse theory in the theory of Lie groups and symmetric spaces, see [97] and [212]. As an example, consider a classical Lie group $G \subset M_n(\mathbb{K})$, with $\mathbb{K} = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$. Endow $M_n(\mathbb{K})$ with the standard scalar product

$$(A, B) = \operatorname{Re}\operatorname{tr}\big(A^\dagger B\big).$$

Every $A \in M_n(\mathbb{K})$ defines a function on G by

$$f_A : G \to \mathbb{R}, \qquad f_A(a) := \operatorname{Re}\operatorname{tr}(Aa). \tag{8.9.12}$$

One can show that if A is diagonal with pairwise distinct entries, f_A is a Morse function. It is not hard to compute the critical points and their indices. It is also interesting to compute the critical set for the case $A = \mathbb{1}$, see Exercise 8.9.4.

In applications, e.g. in the theory of systems with symmetries, one often faces the situation that one has to work with a restricted class of functions which cannot be assumed to have isolated critical points. The following definition yields a reasonable generalization to this case:

Definition 8.9.11 (Morse-Bott function) Let M be a smooth manifold. A smooth function $f : M \to \mathbb{R}$ is called a Morse-Bott function if the following conditions are fulfilled:

1. The critical set M_f is a disjoint union of connected embedded submanifolds, called critical submanifolds.
2. For every critical submanifold N and every $m \in N$, the bilinear form induced by $\operatorname{Hess}_m(f)$ on the normal space $\mathrm{N}_m N = \mathrm{T}_m M / \mathrm{T}_m N$ is non-degenerate.[25]

As an example, we consider a torus lying on a plane. Its height function is obviously a Morse-Bott function with two critical submanifolds: a circle, on which the height function is maximal, and a circle in the plane, on which it is minimal.

As before, the index of a critical point $m \in N$ is defined as the index of $\operatorname{Hess}_m(f)$. Since N is connected, the rank of $\operatorname{Hess}_m(f)$ is constant along N. Hence, the index

[25] This is equivalent to the condition $\ker(\operatorname{Hess}_m(f)) = \mathrm{T}_m N$.

depends only on the submanifold N. Since single points are embedded submanifolds, every Morse function is a Morse-Bott function. The Morse Lemma generalizes to Morse-Bott functions as follows: assume that $\dim M = n$, $\dim N = d$ and that the index of N is i. Then, for every $m_0 \in N$, there exists a local chart of M at m_0 with coordinates $x^1, \dots, x^d$ and $y^1, \dots, y^{n-d}$ such that N is given locally by $\mathbf{y} = 0$ and the local representative of f has the form

$$f(\mathbf{x}, \mathbf{y}) = f(N) - \frac{1}{2}\left(y_1^2 + \cdots + y_i^2\right) + \frac{1}{2}\left(y_{i+1}^2 + \cdots + y_{n-d}^2\right). \tag{8.9.13}$$

By a diagonalization procedure similar to that used for Corollary 8.9.4, this will follow from

Theorem 8.9.12 (Morse-Bott Lemma) *Let M be a manifold of dimension n, let f be a Morse-Bott function on M and let N be a critical submanifold of f of dimension d. For every $m_0 \in N$, there exists a local chart of M at m_0 with coordinates $x^1, \dots, x^d$ and $y^1, \dots, y^{n-d}$, mapping points of N to $\mathbb{R}^d \times \{0\} \subset \mathbb{R}^n$, and a non-degenerate quadratic form Q on $\mathbb{R}^{n-d}$ such that the local representative of f is given by*

$$f(\mathbf{x}, \mathbf{y}) = f(N) + \mathsf{Q}(\mathbf{y}). \tag{8.9.14}$$

Proof We follow [32]. Denote $r = n - d$. Since N is embedded, we find a local chart (U, κ) of M at m_0 such that $\kappa(U \cap N) \subset \mathbb{R}^d \times \{0\} \subset \mathbb{R}^n$. Thus, without loss of generality, we may assume that $M = \mathbb{R}^n = \mathbb{R}^d \times \mathbb{R}^r$ and $N = \mathbb{R}^d$. We write $(\mathbf{x}, \mathbf{y})$ for the points of M. By replacing f by $f - f(0)$, we may also assume that $f(N) = 0$. With respect to the decomposition $\mathbb{R}^n = \mathbb{R}^d \oplus \mathbb{R}^r$ of the tangent spaces, the Hessian at $(\mathbf{x}, 0) \in N$ takes the form

$$\mathrm{Hess}_{(\mathbf{x},0)}(f) = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & \mathsf{Q}(\mathbf{x}) \end{array}\right],$$

where $\mathsf{Q}(\mathbf{x})$ is an invertible symmetric matrix of dimension r, depending smoothly on $\mathbf{x}$. We seek a diffeomorphism Φ defined on some open neighbourhood of the origin of $\mathbb{R}^n$ such that

$$\left(\Phi^* f\right)(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\mathsf{Q}(0)_{ij} y^i y^j$$

on that neighbourhood. In the first step, we define

$$h(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\mathsf{Q}(\mathbf{x})_{ij} y^i y^j, \qquad f_t = tf + (1-t)h.$$

We leave it to the reader to check that if we carry out the construction of the time-dependent vector field X from the proof of Theorem 8.9.3 for every fixed $\mathbf{x}$, we obtain an open neighbourhood U of N in M and a smooth vector field X on U

which is tangent to the affine subspaces $\{\mathbf{x}\} \times \mathbb{R}^r$ and whose flow $\Phi_t \equiv \Phi_{t,0}$ satisfies

$$\left(\Phi_1^* f\right)(\mathbf{x}, \mathbf{y}) = h(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \mathrm{Q}(\mathbf{x})_{ij} y^i y^j$$

for all $(\mathbf{x}, \mathbf{y}) \in U$. In the second step, we construct a family of linear transformations of $\mathbb{R}^r$ transforming $\mathrm{Q}(\mathbf{x})$ to $\mathrm{Q}(0)$. Since $\mathrm{Q}(0)$ is symmetric, up to a linear transformation of $\mathbb{R}^n$ we may assume that $\mathrm{Q}(0)$ is diagonal, with diagonal entries $q_1, \dots, q_r$. Since $\mathrm{Q}(0)$ is invertible, all q_i are nonzero. Thus, the diagonal entries of $\mathrm{Q}(\mathbf{x})$ remain nonzero and keep their sign in an open neighbourhood V of the origin in $\mathbb{R}^d$. There, we can define

$$T_1(\mathbf{x}) := \left[\begin{array}{c|ccc} \sqrt{\frac{|q_1|}{|\mathrm{Q}(\mathbf{x})_{11}|}} & -\frac{\mathrm{Q}(\mathbf{x})_{12}}{\mathrm{Q}(\mathbf{x})_{11}} & \cdots & -\frac{\mathrm{Q}(\mathbf{x})_{12}}{\mathrm{Q}(\mathbf{x})_{11}} \\ \hline 0 & & & \\ \vdots & & \mathbb{1}_{r-1} & \\ 0 & & & \end{array}\right]$$

and compute

$$T_1(\mathbf{x})^{\mathrm{T}} \mathrm{Q}(\mathbf{x}) T_1(\mathbf{x}) = \left[\begin{array}{c|ccc} q_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B_2(\mathbf{x}) & \\ 0 & & & \end{array}\right]$$

with a s ymmetric matrix $B_2(\mathbf{x})$ of dimension $r - 1$, depending smoothly on $\mathbf{x}$ and satisfying $B_2(0) = \mathrm{diag}(q_2, \dots, q_r)$. By possibly shrinking V, we may achieve that the diagonal entries of $B_2(\mathbf{x})$ are nonzero. Thus, we may apply the same procedure in one dimension less and with q_1 replaced by q_2. Iterating this, we obtain invertible matrices $T_1(\mathbf{x}), \dots, T_r(\mathbf{x})$ of dimension, respectively, $r, \dots, 1$, which depend smoothly on $\mathbf{x}$. Then,

$$T(\mathbf{x})^{\mathrm{T}} \mathrm{Q}(\mathbf{x}) T(\mathbf{x}) = \mathrm{Q}(0) \quad \text{with } T(\mathbf{x}) = T_1(\mathbf{x}) \left(\mathbb{1}_1 \oplus T_2(\mathbf{x})\right) \cdots \left(\mathbb{1}_{r-1} \oplus T_r(\mathbf{x})\right)$$

and hence

$$\left(\tilde{\Phi}^* h\right)(\mathbf{x}, \mathbf{y}) = \mathrm{Q}(0)_{ij} y^i y^j \quad \text{with } \tilde{\Phi}(\mathbf{x}, \mathbf{y}) = \left(\mathbf{x}, T(\mathbf{x})\mathbf{y}\right).$$

Thus, $\Phi = \Phi_1 \circ \tilde{\Phi}$ yields the desired local diffeomorphism (defined on some neighbourhood of the origin in $\mathbb{R}^n$, which can be determined from U and V). □

From the proof we can read off the following version of the Morse-Bott Lemma, to be used later on.

Corollary 8.9.13 *Let N be a manifold and let $f : N \times \mathbb{R}^r \to \mathbb{R}$ be a Morse-Bott function with critical submanifold $N \times \{0\}$ and $f(N \times \{0\}) = 0$. For every $m \in N$,* *there exist open neighbourhoods U of m in N and V of the origin in $\mathbb{R}^r$, a non-degenerate quadratic form Q on $\mathbb{R}^r$ and a diffeomorphism Φ from $U \times V$ into $N \times \mathbb{R}^r$ such that*

$$f \circ \Phi(m, \mathbf{y}) = \mathsf{Q}(\mathbf{y}), \qquad \mathrm{pr}_N \circ \Phi(m, \mathbf{y}) = m$$

for all $m \in U$ and $\mathbf{y} \in V$.

To conclude, let us add that, analogously to Morse homology, one can develop a homology theory based on Morse-Bott functions, see the Lecture Notes by Hutchings [146]. We will meet a further generalization of the classical Morse Lemma in the context of Morse families in Sect. 12.4.

Exercises

8.9.1 Prove Corollary 8.9.4 by diagonalization of $\mathrm{Hess}_0(f)$ and an appropriate scaling of basis vectors.

8.9.2 Find the critical sets of the functions

$$\begin{aligned} f(x) &= x^2, \\ f(x) &= x^3, \\ f(x, y) &= x^2, \\ f(x, y) &= x^2 y^2, \\ f(x, y) &= x^3 - 3x^2 y^2, \\ f(x, y) &= \cos(2\pi x) + \cos(2\pi y), \\ f(z_0 : \ldots : z_1) &= \frac{\sum_{i=0}^n i |z_i|^2}{\sum_{i=0}^n |z_i|^2}. \end{aligned}$$

Which of them are Morse functions? Find the indices for their critical points. *Hint*. The sixth function has to be understood as a function on the 2-dimensional torus and the last function has to be viewed as a function on the complex projective space $\mathbb{C}\mathrm{P}^n$, written down in homogeneous coordinates, cf. Example 1.1.15.

8.9.3 Show that the height function of an upright 2-torus T^2 as shown in Fig. 8.2 is a Morse function. Determine the critical points and their indices. Determine the homotopy type of the subsets M^a for all values of a. Convince yourself that at each critical point the homotopy type changes by attaching a cell whose dimension is given by the index.

8.9.4 Study the function (8.9.12) for the case $A = \mathbb{1}$. Show that the set of critical points of $f_{\mathbb{1}}$ coincides with the set of involutions $a^2 = \mathbb{1}$. Find a geometric interpretation of this set.
Hint. Recall the definition of the Graßmann manifolds $G_{\mathbb{K}}(k, n)$ in Example 5.7.6.

Chapter 9
Hamiltonian Systems

In this chapter we start discussing the theory of Hamiltonian systems. We begin with an introduction to the subject, including the Legendre transformation and a brief discussion of linear nonholonomic systems. Next, we present three classes of examples, which will play a role in the subsequent chapters: the geodesic flow, Hamiltonian systems on Lie group manifolds and Hamiltonian systems on coadjoint orbits. We close the elementary part of this chapter by presenting the time-dependent picture.

In Sect. 9.4 we investigate the structure of regular energy surfaces and discuss the problem of the existence of closed integral curves for autonomous systems. This leads us to the famous Weinstein conjecture and to some aspects of symplectic topology. We will see that a special type of symplectic invariants, called symplectic capacities, constitutes the most important technical tool for studying the Weinstein conjecture. In Sect. 9.5 we start to investigate the behaviour of a Hamiltonian system near a critical integral curve. We will see that, generically, there is a continuum of periodic integral curves nearby, constituting so called orbit cylinders. These cylinders can undergo bifurcations. We study one of these bifurcations, provided by the Lyapunov Centre Theorem, in detail. Next, we derive the so-called Birkhoff normal form both for symplectomorphisms in the neighbourhood of elliptic fixed points and for the Hamiltonian of a system near an equilibrium. This normal form implies a foliation of the phase space into invariant tori and, in the normal form approximation, the theory becomes integrable. According to the celebrated KAM theory many of the invariant tori persist the perturbation caused by taking into account the full symplectomorphism or the full Hamiltonian, respectively. Moreover, we use the Birkhoff Normal Form Theorem to prove the Birkhoff-Lewis Theorem, which yields the existence of infinitely many periodic points near a closed integral curve of a certain type on a given energy surface. Finally, we study some aspects of stability, with the main emphasis on systems with two degrees of freedom.

In the final two sections we study time-dependent Hamiltonian systems. In Sect. 9.8 we deal with the stability problem of time-periodic systems with emphasis on parametric resonance and in the final section we comment on the famous Arnold

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7_9, © Springer Science+Business Media Dordrecht 2013

conjecture about the existence of fixed points of Hamiltonian symplectomorphisms and its relation to the existence problem for closed integral curves.

9.1 Introduction

A Hamiltonian system is a triple (M, ω, H), where (M, ω) is a symplectic manifold, in the present context called the phase space, and H is a smooth function on M, called the Hamiltonian or the Hamiltonian function. In this section we restrict ourselves to autonomous[1] systems. Moreover, throughout the whole chapter we confine our attention to holonomic systems, except for a few remarks on the nonholonomic case in this section.[2] In this context, the above data have the following physical interpretation: every point $m \in M$ describes a (pure) state of the system. Smooth functions on M are observables. The result of a measurement of the observable $f \in C^\infty(M)$ in the state m is given by the value $f(m)$ of the function at m. The dynamics of the system is governed by the Hamiltonian: the time evolution of a state m is given by the integral curve through m of the Hamiltonian vector field X_H generated by H. Let us write down the equation for the integral curves of X_H in local Darboux coordinates q^i and p_i. By (8.2.2), we have

$$X_H = (\partial_{p_i} H)\partial_{q^i} - (\partial_{q^i} H)\partial_{p_i}. \tag{9.1.1}$$

Thus, the equations for the integral curves $t \mapsto (\mathbf{q}(t), \mathbf{p}(t))$ of X_H are given by

$$\dot{q}^i(t) = \frac{\partial H}{\partial p_i}\big(\mathbf{q}(t), \mathbf{p}(t)\big), \qquad \dot{p}_i(t) = -\frac{\partial H}{\partial q^i}\big(\mathbf{q}(t), \mathbf{p}(t)\big). \tag{9.1.2}$$

These are the Hamilton equations. We see that the whole information about the dynamics of the system is encoded in the flow of the Hamiltonian vector field generated by the Hamiltonian function, indeed.

In physics, a variety of phase space models occur. The most prominent one is that of a cotangent bundle,[3] that is, $M = \mathrm{T}^*Q$, see Sect. 8.3. The base manifold Q is called the configuration space of the system and $\dim Q$ is called the number of degrees of freedom. Given the great importance of this model, let us show how it is derived from the Lagrangian formulation, cf. Sect. 4.8, via the so-called Legendre transformation. For that purpose, let us consider a mechanical system with configuration space Q, described by a Lagrangian $L : \mathrm{T}Q \to \mathbb{R}$. Denote the canonical bundle projection by $\pi_\mathrm{T} : \mathrm{T}Q \to Q$ and extend L to a mapping

$$\tilde{L} : \mathrm{T}Q \to Q \times \mathbb{R}, \qquad \tilde{L}(X) := \big(\pi_\mathrm{T}(X), L(X)\big).$$

---

[1] That is, the Hamiltonian does not depend explicitly on time. We will show in Sect. 9.3 that the framework discussed here can be easily extended to the non-autonomous case.

[2] As mentioned in Sect. 4.8, in the theory of nonholonomic constraints many interesting branches are studied. Here, we limit ourselves to showing that the Lagrangian formulation presented there has a counterpart on the Hamiltonian level.

[3] Sometimes one is led to go beyond the cotangent bundle model though, notably in the study of systems with symmetries (Chap. 10) where coadjoint orbits of Lie groups play an important role as phase space models.

We treat $Q \times \mathbb{R}$ as a vector bundle over Q, denote its canonical projection by $\rho : Q \times \mathbb{R} \to Q$ and note that

$$\rho \circ \tilde{L} = \pi_{\mathrm{T}}. \tag{9.1.3}$$

Now, let us define the mapping

$$FL : \mathrm{T}Q \to \mathrm{Hom}(\mathrm{T}Q, Q \times \mathbb{R}) \cong \mathrm{T}^*Q, \qquad (FL)(X) := (\tilde{L}_{\pi_{\mathrm{T}}(X)})'_X. \tag{9.1.4}$$

Here $\tilde{L}_q := \tilde{L}_{\restriction \pi_{\mathrm{T}}^{-1}(q)}$ is the restriction of the mapping $\tilde{L}$ to the fibre over $q \in Q$ and $\mathrm{Hom}(\mathrm{T}Q, Q \times \mathbb{R})$ is the vector bundle defined in Remark 2.4.7. The mapping FL is called the fibre derivative[4] of L. It is obviously fibre preserving and smooth. We have

$$\big\langle (FL)(X), Y \big\rangle = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} L(X + tY)$$

for all X, Y in the same fibre of $\mathrm{T}Q$. For a typical Lagrangian, given by the difference of the kinetic and the potential energy,

$$L(X) = \frac{1}{2}\mathrm{g}(X, X) - V\big(\pi_{\mathrm{T}}(X)\big), \tag{9.1.5}$$

where g is a Riemannian metric on Q and V is a potential function on Q, the Legendre transform reduces to

$$\big\langle (FL)(X), Y \big\rangle = \mathrm{g}(X, Y). \tag{9.1.6}$$

Thus, it coincides with the natural vector bundle isomorphism $\mathrm{g} : \mathrm{T}Q \to \mathrm{T}^*Q$ induced by the metric. If the fibre derivative $FL : \mathrm{T}Q \to \mathrm{T}^*Q$ is a diffeomorphism, it is called the Legendre transformation induced by L and the Lagrangian function L is called hyperregular. A Lagrangian of the type given by (9.1.5) is always hyperregular. Finally, L is called regular iff FL is a local diffeomorphism.

Remark 9.1.1 Let us analyze the fibre derivative in local coordinates. Thus, let q^i be local coordinates on Q and let $q^i, \dot{q}^i$ and q^i, p_i be the induced local coordinates on $\mathrm{T}Q$ and T^*Q, respectively. The Lagrangian L is regular iff

$$\det\left(\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\right) \neq 0, \tag{9.1.7}$$

(Exercise 9.1.1). The fibre derivative FL takes the local form

$$(FL)(\mathbf{q}, \dot{\mathbf{q}}) = \frac{\partial L}{\partial \dot{q}^i}(\mathbf{q}, \dot{\mathbf{q}})\, \mathrm{d}q^i \tag{9.1.8}$$

(Exercise 9.1.2), that is, it is given by the mapping

$$(\mathbf{q}, \dot{\mathbf{q}}) \mapsto \left(\mathbf{q}, \mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}(\mathbf{q}, \dot{\mathbf{q}})\right).$$

[4] More generally, if we replace $\mathrm{T}Q$ and $Q \times \mathbb{R}$ by arbitrary vector bundles E and F over Q, respectively, then (9.1.4) yields the definition of the fibre derivative $Ff : E \to \mathrm{Hom}(E, F)$ for any smooth mapping $f : E \to F$ fulfilling condition (9.1.3).

If L is hyperregular, this is the local representative of the Legendre transformation known from classical mechanics.

Now assume that L is hyperregular. In the Lagrangian formulation, the dynamics of the system is governed by the Euler-Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0, \tag{9.1.9}$$

compare with (4.8.11). This system of equations can be cast into a coordinate-free form in different ways. For our purposes, it is convenient to proceed as follows. We take the pullback $\omega_L := (FL)^*\omega$ of the canonical symplectic structure on T^*Q, and thus endow $\mathrm{T}Q$ with a symplectic structure such that FL becomes a symplectomorphism. Next, we define the energy function

$$E: \mathrm{T}Q \to \mathbb{R}, \qquad E(X) := \big\langle (FL)(X), X\big\rangle - L(X), \tag{9.1.10}$$

and consider the Hamiltonian vector field X_E on $\mathrm{T}Q$ generated by E,

$$X_E \lrcorner \omega_L = -\mathrm{d}E. \tag{9.1.11}$$

We encourage the reader to check that in local coordinates q^i, $\dot{q}^i$ the equations for the integral curves of X_E are equivalent to the Euler-Lagrange equations (9.1.9) (Exercise 9.1.3). Now, the Hamiltonian of the system is defined by

$$H: \mathrm{T}^*Q \to \mathbb{R}, \qquad H := E \circ (FL)^{-1}. \tag{9.1.12}$$

From (9.1.11) we read off

$$\big((FL)^{-1}\big)^*(X_E \lrcorner \omega_L) = -\mathrm{d}H.$$

Moreover, Proposition 8.2.9 implies

$$(FL)_* X_E = X_H. \tag{9.1.13}$$

Thus, if L is hyperregular, the Lagrangian formulation and the Hamiltonian formulation are equivalent.

Remark 9.1.2

1. If L is hyperregular, the tuple $(\mathrm{T}Q, \omega_L, E)$ is a Hamiltonian system which is equivalent to $(\mathrm{T}^*Q, \omega, H)$. Usually, by the Hamiltonian formulation one means the latter setting. However, it may be convenient to use the former setting, see e.g. Example 9.2.1.
2. For the model class defined by (9.1.5) we obtain

$$H(\xi) = \frac{1}{2}\mathrm{g}^{-1}(\xi,\xi) + V\big(\pi(\xi)\big) = T(\xi) + V\big(\pi(\xi)\big), \tag{9.1.14}$$

with

$$\mathrm{g}^{-1}(\xi,\eta) := \mathrm{g}\big(\mathrm{g}^{-1}(\xi), \mathrm{g}^{-1}(\eta)\big)$$

being the metric induced on T^*Q via the vector bundle isomorphism

$$\mathrm{g} = FL: \mathrm{T}Q \to \mathrm{T}^*Q.$$

3. In local coordinates, we obtain

$$H(\mathbf{q},\mathbf{p}) = \frac{\partial L}{\partial \dot{q}^i}(\mathbf{q},\dot{\mathbf{q}})\dot{q}^i - L(\mathbf{q},\dot{\mathbf{q}}), \tag{9.1.15}$$

with $\dot{\mathbf{q}}$ obtained in terms of $(\mathbf{q},\mathbf{p})$ from solving $\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{q}}}$ for $\dot{\mathbf{q}}$. Thus, H is the Legendre transform of L known from classical mechanics.

Next, let us make a few remarks on linear nonholonomic systems. Recall from Sect. 4.8 that a linear nonholonomic constraint is given by a smooth non-integrable distribution $D \subset \mathrm{T}Q$ on the configuration space Q of the system and that the Euler-Lagrange equations for this case are given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = \lambda_a \mu_i^a. \tag{9.1.16}$$

Here, the λ_a are Lagrange multipliers and μ^a, $a = 1,\ldots,s$, is a system of local 1-forms spanning the annihilator D^0 of D, that is,

$$D_m := \left\{X \in \mathrm{T}_m Q : \left\langle \mu^a(m), X\right\rangle = 0 \text{ for all } a\right\}. \tag{9.1.17}$$

The solutions $t \mapsto \mathbf{q}(t)$ and the Lagrange multipliers λ_a are determined by the Euler-Lagrange equations (9.1.16) and the constraint equations

$$\mu_i^a\big(\mathbf{q}(t)\big)\dot{q}^i(t) = 0, \tag{9.1.18}$$

cf. (4.8.5). We wish to cast this system of equations into a coordinate-free Hamiltonian form. For that purpose, assume that L is hyperregular. Similarly to (9.1.11), the Euler-Lagrange equations (9.1.16) are equivalent to the equations for the integral curves of a vector field $\tilde{X}_E$, which, here, is defined by

$$\tilde{X}_E \lrcorner \omega_L = -\mathrm{d}E + \lambda_a \pi_{\mathrm{T}}^* \mu^a.$$

The coordinate-free form of the constraint equation (9.1.18) is

$$\left\langle \mu^a \circ \pi_{\mathrm{T}}, \pi_{\mathrm{T}}' \circ \tilde{X}_E\right\rangle = 0,$$

that is, $\pi_{\mathrm{T}}' \circ \tilde{X}_E : \mathrm{T}Q \to \mathrm{T}Q$ takes values in D. We apply the Legendre transformation to these equations. Denoting

$$\tilde{X}_H = (FL)_* \tilde{X}_E, \tag{9.1.19}$$

and using $\pi_{\mathrm{T}} = \pi \circ (FL)$, for the first equation we obtain

$$\tilde{X}_H \lrcorner \omega = -\mathrm{d}H + \lambda_a \pi^* \mu^a. \tag{9.1.20}$$

In canonical bundle coordinates, the equations for the integral curves of $\tilde{X}_H$ read

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i} + \lambda_a \mu_i^a. \tag{9.1.21}$$

These are the Hamilton equations for the case of linear nonholonomic constraints. For the constraint equation we obtain

$$\left\langle \mu^a \circ \pi, \pi' \circ \tilde{X}_H\right\rangle = 0, \tag{9.1.22}$$

that is, $\pi' \circ \tilde{X}_H : \mathrm{T}^*Q \to \mathrm{T}Q$ takes values in D. The image $M = (FL)(D) \subset \mathrm{T}^*Q$ of D under the Legendre transformation is called the constraint submanifold. We have

$$M = \left\{\xi \in \mathrm{T}^*Q : \left\langle \mu^a\left(\pi(\xi)\right), (FL)^{-1}(\xi)\right\rangle = 0 \text{ for all } a\right\}.$$

Next, we wish to implement the constraint equation (9.1.22) in an intrinsic manner, this way removing the Lagrange multipliers. We follow Bates and Śniatycki [35], see also [170]. Let us consider the distribution $\hat{D} = (\pi')^{-1}(D)$ on T^*Q. In terms of the local frame $\{\mu^a\}$ in D^0, it is given by

$$\hat{D} = \left\{X \in \mathrm{TT}^*Q : \left\langle \pi^*\mu^a, X\right\rangle = 0 \text{ for all } a\right\}. \tag{9.1.23}$$

Since the constraint equation is equivalent to $\pi' \circ \tilde{X}_H : \mathrm{T}^*Q \to \mathrm{T}Q$ taking values in D, $\tilde{X}_H$ must take values in $\hat{D}$. This implies, in addition, that $\tilde{X}_H$ is tangent to M (Exercise 9.1.4). Hence, $\tilde{X}_H$ takes values in

$$F := \hat{D} \cap \mathrm{T}M. \tag{9.1.24}$$

It is easy to show that $\hat{D} = \hat{D}^\omega \oplus F$, see Exercise 9.1.5. Therefore, F is a regular distribution on M, called the constraint distribution, and the fibrewise restriction ω_F of ω to F is non-degenerate. Since the restriction of $\tilde{X}_H$ to M, denoted by the same symbol, takes values in F, the fibrewise restriction of $\tilde{X}_H \lrcorner \omega$ to F coincides with $\tilde{X}_H \lrcorner \omega_F$. Since, in addition, $\pi^*\mu_a$ vanishes on F, restricting (9.1.20) to F we obtain

$$\tilde{X}_H \lrcorner \omega_F = -(\mathrm{d}H)_F. \tag{9.1.25}$$

Since ω_F is non-degenerate, this equation determines $\tilde{X}_H$ uniquely. This way, we have reduced the dynamics to the constraint submanifold M: it is given by the integral curves of the vector field $\tilde{X}_H$ on M which takes values in the constraint distribution F. We encourage the reader to work out the description of the above structures in the bundle coordinates q^i, p_i, $\dot{q}^i$ and $\dot{p}_i$ on TT^*Q induced from bundle coordinates q^i and p_i on T^*Q (Exercise 9.1.7).

Remark 9.1.3 The distribution D is spanned by the Hamiltonian vector fields Z_a on T^*Q generated by the functions

$$\xi \mapsto \left\langle \mu_a\left(\pi(\xi)\right), (FL)^{-1}(\xi)\right\rangle.$$

Evaluating the Hamilton equations (9.1.20) on Z^a, we obtain

$$0 = -\left\langle Z^a, \mathrm{d}H\right\rangle + \sum_b \lambda_b \left\langle \pi^*\mu^a, Z^b\right\rangle. \tag{9.1.26}$$

It is easy to show that for the class of models defined by (9.1.5), the matrix $\langle \pi^*\mu^a, Z^b\rangle$ is invertible (Exercise 9.1.6). Thus, this equation can be used to eliminate the Lagrange multipliers λ_b. The resulting equation reduces to (9.1.25).

Now, we discuss symplectomorphisms. In the context of physics, a symplectomorphism of the phase space is referred to as a canonical transformation. Canonical transformations play an enormous role in the study of Hamiltonian systems. The first simple but very important observation follows from Remark 8.2.5/1, which states that the flow of the Hamiltonian vector field X_H consists of locally defined symplectomorphisms. This means that the dynamics of a Hamiltonian system can be viewed as a time-dependent canonical transformation. This observation can be considered the starting point of Hamilton-Jacobi theory, to be presented in Chap. 12. Moreover, Proposition 8.1.3 implies

Theorem 9.1.4 (Liouville) *The phase space volume form Ω_ω of a Hamiltonian system (M, ω, H) is invariant under the flow of the Hamiltonian vector field X_H.*

The following corollary is a direct consequence of Proposition 8.2.9. It states that the Hamilton equations are invariant under canonical transformations.

Corollary 9.1.5 *Let (M, ω, H) be a Hamiltonian system and let $\Phi: M \to M$ be a canonical transformation. Then,*

$$\Phi_* X_H = X_{H \circ \Phi^{-1}}. \tag{9.1.27}$$

Remark 9.1.6

1. If the canonical transformation Φ is given in local Darboux coordinates by
$$(\mathbf{q}, \mathbf{p}) \mapsto \big(\bar{\mathbf{q}}(\mathbf{q}, \mathbf{p}), \bar{\mathbf{p}}(\mathbf{q}, \mathbf{p})\big),$$
then (9.1.27) yields the Hamilton equations in the variables $(\bar{\mathbf{q}}, \bar{\mathbf{p}})$, with Hamiltonian $\bar{H}(\bar{\mathbf{q}}, \bar{\mathbf{p}}) = (H \circ \Phi^{-1})(\bar{\mathbf{q}}, \bar{\mathbf{p}})$. We stress that there exist transformations which leave the Hamilton equations invariant but which are not canonical, e.g.
$$(\mathbf{q}, \mathbf{p}, H) \mapsto (\bar{\mathbf{q}} = \mathbf{q}, \bar{\mathbf{p}} = a\mathbf{p}, \bar{H} = aH)$$
with an arbitrary constant $a \neq 0$.
2. With a slight abuse of language, the view on Hamiltonian systems described above may be called the Schrödinger picture of classical mechanics, because the dynamics of the system is given in terms of the time evolution of the states. The time evolution can be shifted to the observables as follows. For $f \in C^\infty(M)$, define the family $t \mapsto f(t) := f \circ \Phi_t$, where Φ is the flow of X_H. Then,
$$\big(f(t)\big)(m) = f\big(\Phi_t(m)\big),$$
that is, the result of a measurement of $f \equiv f(0)$ in the state $\Phi_t(m)$ coincides with the result of the measurement of $f(t)$ in the state m. By (8.2.6), using the standard physics notation $\dot{f} \equiv \frac{\mathrm{d}}{\mathrm{d}t} f(t)$, we obtain the following equation of motion for $f(t)$:
$$\dot{f} = \{H, f\}. \tag{9.1.28}$$
In particular, the Hamilton equations can be rewritten as
$$\dot{q}^i = \big\{H, q^i\big\}, \quad \dot{p}_i = \{H, p_i\}. \tag{9.1.29}$$
This may be viewed as the Heisenberg picture of classical mechanics.

3. The notion of Hamiltonian system generalizes to Poisson manifolds. In this generalized sense, a Hamiltonian system is a triple $(M, \{\,,\,\}, H)$ where $(M, \{\,,\,\})$ is a Poisson manifold and $H \in C^\infty(M)$. The dynamics is given by $\dot{f} = \{H, f\}$. In local coordinates, this yields the equations of motion
$$\dot{x}^i = \left\{H, x^i\right\} = -\Pi^{ik} \frac{\partial H}{\partial x^k}.$$

Finally, let us turn to the important notion of constant of motion.[5]

Definition 9.1.7 Let (M, ω, H) be a Hamiltonian system. A function $f \in C^\infty(M)$ is called constant of motion if
$$\frac{\mathrm{d}}{\mathrm{d}t} f\big(\gamma(t)\big) = 0$$
for every integral curve $t \to \gamma(t)$ of the Hamiltonian vector field X_H.

Equivalently, f is a constant of motion iff its pullback $\Phi_t^* f$ under the flow Φ_t of X_H is independent of t.

Remark 9.1.8 By (9.1.2), the function p_i is a constant of motion iff H does not depend on the corresponding coordinate q^i. In this case, q^i is said to be cyclic.

Proposition 9.1.9 *The Hamiltonian H of a Hamiltonian system (M, ω, H) is a constant of motion.*

Proof For any integral curve $t \mapsto \gamma(t)$ of X_H, we have
$$\frac{\mathrm{d}}{\mathrm{d}t} H\big(\gamma(t)\big) = X_H(H)\big(\gamma(t)\big) = \langle \mathrm{d}H, X_H\rangle\big(\gamma(t)\big) = -\omega(X_H, X_H)\big(\gamma(t)\big) = 0. \qquad \square$$

This is, of course, the law of energy conservation for autonomous systems. The following proposition characterizes constants of motion in terms of the Poisson bracket.

Proposition 9.1.10 *Let (M, ω, H) be a Hamiltonian system.*

1. *A function $f \in C^\infty(M)$ is a constant of motion iff $\{f, H\} = 0$.*
2. *If $f, g \in C^\infty(M)$ are constants of motion, then $\{f, g\}$ is a constant of motion, too.*

Proof Point 1 follows from
$$\frac{\mathrm{d}}{\mathrm{d}t} f\big(\gamma(t)\big) = X_H(f)\big(\gamma(t)\big) = \{H, f\}\big(\gamma(t)\big)$$
and point 2 is a consequence of point 1 and the Jacobi identity. $\square$

[5] Also referred to as an integral of motion or a first integral.

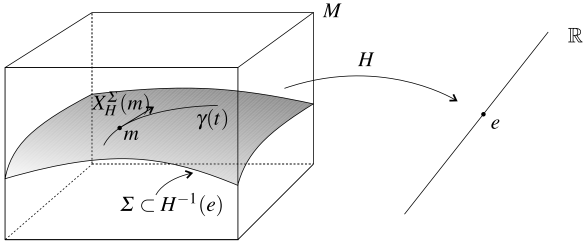


Fig. 9.1 The restriction of X_H to Σ is tangent to Σ

Now, assume that $e \in \mathbb{R}$ is a regular value of H, that is, $\mathrm{d}H(m) \neq 0$ for all $m \in H^{-1}(e)$. By the Level Set Theorem, $H^{-1}(e)$ is an embedded submanifold of M of dimension $2n-1$. The connected components of $H^{-1}(e)$ are referred to as regular energy surfaces. Let Σ be such a regular energy surface and let $\iota\colon \Sigma \to M$ be the natural inclusion mapping. By Proposition 9.1.9, Σ is invariant under the flow of X_H. Thus, X_H is tangent to Σ and, therefore, induces a vector field X_H^Σ on Σ which is ι-related to X_H,

$$\iota'\big(X_H^\Sigma(m)\big) = X_H\big(\iota(m)\big), \quad m \in \Sigma. \tag{9.1.30}$$

We draw the important conclusion that for every regular value of the Hamiltonian we get a reduction of the dynamics to a submanifold of codimension 1, see Fig. 9.1. Since $\mathrm{d}H$ vanishes nowhere on Σ, X_H^Σ does not have equilibria.

Let us consider $\omega_\Sigma := \iota^*\omega$. This is a closed 2-form on Σ, which is necessarily degenerate, because Σ has dimension $(2n-1)$. To calculate the rank, we use (7.2.3) and Proposition 7.2.4. This yields

$$\operatorname{rank}(\omega_\Sigma)_m = \dim(\mathrm{T}_m\Sigma) - \dim\big(\mathrm{T}_m\Sigma \cap (\mathrm{T}_m\Sigma)^\omega\big) = 2(n-1)$$

for all $m \in \Sigma$. Hence, the rank of ω_Σ is maximal and the characteristic distribution $D^{\omega_\Sigma} = \ker \omega_\Sigma$ of ω_Σ is regular of rank 1. Moreover, using (9.1.30), we obtain

$$\big(\iota^*\omega\big)_m\big(X_H^\Sigma(m), Y\big) = \omega_{\iota(m)}\big(X_H\big(\iota(m)\big), \iota'Y\big) = -\big\langle \mathrm{d}H\big(\iota(m)\big), \iota'Y\big\rangle = 0$$

for any $Y \in \mathrm{T}_m\Sigma$. Thus, $X_H^\Sigma \lrcorner \omega_\Sigma = 0$ and hence D^{ω_Σ} is spanned by X_H^Σ. In particular, the (images of the) integral curves of X_H^Σ coincide with the characteristics of Σ.[6] Moreover, as a vector bundle, D^{ω_Σ} is trivial.

The reduction to Σ, induced by a single constant of motion, can be generalized to several constants of motion $f_1, \dots, f_k$. Consider the mapping

$$F = (f_1, \dots, f_k)\colon M \to \mathbb{R}^k$$

and assume that $\mathbf{h} \in \mathbb{R}^k$ is a regular value of F. Then, by the Level Set Theorem, $F^{-1}(\mathbf{h})$ is a $(2n-k)$-dimensional embedded submanifold of M. This submanifold is invariant under the flow of X_H. Chapters 10 and 11 are devoted to the study of the following specific reductions of this type.

[6]More precisely, they are equivalent as submanifolds of M, because they are integral manifolds of the same integrable distribution, cf. Proposition 3.5.15 and Remark 1.6.13/5.

(a) If the reduction is induced by a symmetry of the system, the constants of motion f_i arise as the components of a so-called momentum mapping. There exists a general symmetry reduction procedure, called Marsden-Weinstein reduction. This theory will be presented in detail in Chap. 10.

(b) If the constants of motion f_i fulfil an additional global integrability condition, namely, if all their pairwise Poisson brackets vanish, one says that they are in involution. If there exist n constants of motion in involution, the system is said to be integrable. All the exactly solvable models usually taught in a course on classical mechanics are of this type. This class of systems will be studied in Chap. 11.

We note that, locally, there are always $(2n-1)$ functionally independent constants of motion. Indeed, Proposition 3.2.17 provides local charts (U,κ) such that $(X_H)_{\restriction U} = \partial_1^\kappa$. The corresponding coordinate functions $(\kappa^2,\dots,\kappa^n)$ are constants of motion on the open subset U. However, in general these local constants of motion cannot be extended to the whole phase space in a functionally independent way.

As a matter of fact, many examples belong both to classes (a) and (b).

We conclude this section with a simple example which illustrates that a reduction may allow to draw conclusions about the qualitative behaviour of the dynamics of the full system.

Example 9.1.11 Consider the Hamiltonian system with Hamiltonian

$$H = \frac{1}{2}\big(p_1^2 + p_2^2 + \omega_1^2 q_1^2 + \omega_2^2 q_2^2\big)$$

on the phase space $M = \mathrm{T}^*(\mathbb{R}\setminus\{0\}) \times \mathrm{T}^*(\mathbb{R}\setminus\{0\})$. The functions

$$H_i := \frac{1}{2}\big(p_i^2 + \omega_i^2 q_i^2\big), \quad i = 1, 2,$$

are constants of motion. Every $\mathbf{h} = (E_1, E_2) \in \mathbb{R}^2$ fulfilling $E_1, E_2 \neq 0$ is a regular value of (H_1, H_2). The corresponding level set Σ is a two-dimensional torus. In angle coordinates ϕ_1, ϕ_2 defined by

$$q_i = \frac{\sqrt{2E_i}}{\omega_i}\sin\phi_i, \qquad p_i = \sqrt{2E_i}\cos\phi_i,$$

X_H^Σ is given by

$$X_H^\Sigma = \omega_1\partial_{\phi_1} + \omega_2\partial_{\phi_2} \tag{9.1.31}$$

(Exercise 9.1.8). If ω_1/ω_2 is rational, the integral curves are closed and hence isomorphic to S^1. If ω_1/ω_2 is irrational, every integral curve is dense in Σ.

Exercises

9.1.1 Prove that a Lagrangian is regular iff it satisfies (9.1.7).

9.1.2 Prove Formula (9.1.8).

9.1.3 Prove that in bundle coordinates q^i and $\dot{q}^i$ on $\mathrm{T}Q$, the equations for the integral curves of the vector field X_E defined by (9.1.11) are equivalent to the Euler-Lagrange equations (9.1.9).

9.1.4 Show that if the vector field $\tilde{X}_H$ defined by (9.1.19) takes values in $\hat{D}$, it is tangent to the constraint manifold $M = FL(D)$.
Hint. Show that, in bundle coordinates q^i and $\dot{q}^i$ on $\mathrm{T}Q$, the integral curves $t \mapsto (\mathbf{q}(t), \dot{\mathbf{q}}(t))$ of $\tilde{X}_E$ satisfy $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{q}(t) = \dot{\mathbf{q}}(t)$.

9.1.5 Prove that the distribution F defined by (9.1.23) satisfies $\hat{D}^\omega \oplus F = \hat{D}$.
Hint. Show that the distribution $\hat{D}$ is coisotropic and that $\hat{D}^\omega \cap \mathrm{T}M = 0$.

9.1.6 Show that the matrix $\langle \pi^* \mu^b, Z_a \rangle$ in Eq. (9.1.26) is non-degenerate.

9.1.7 Analyze the Hamiltonian description of linear nonholonomic constraints, given in this section, in terms of the bundle coordinates q^i, p_i, $\dot{q}^i$ and $\dot{p}_i$ on $\mathrm{T}\mathrm{T}^*Q$ induced by coordinates q^i on Q.

9.1.8 Prove Formula (9.1.31).

9.1.9 Let (M, ω, H) be a Hamiltonian system and let m_0 be a regular point of H. Show that there exists a local Darboux chart (U, κ) at m_0 with coordinates $\mathbf{x} = (\mathbf{q}, \mathbf{p})$ such that

$$\kappa \circ \Phi_t \circ \kappa^{-1}(\mathbf{x}) = \mathbf{x} + (t, 0, \ldots, 0), \qquad p_1 = H_{\restriction U}$$

for all $(t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^{2n}$ for which the left hand side of the first equation is defined. From the first equation we read off

$$\kappa_1\big(\Phi_t(m)\big) = \kappa_1(m) + t, \qquad \kappa_i\big(\Phi_t(m)\big) = \kappa_i(m), \quad 1 < i \leq 2n,$$

that is, q^1 is the local coordinate along the flow line. A chart of this type is called a Hamiltonian flow box chart.

9.2 Examples

In this section we discuss three classes of examples, which are important both in mathematics and in physics. Each of them will be taken up again later on.

Example 9.2.1 (Geodesic flow) Let (M, g) be a Riemannian manifold. Consider the cotangent bundle T^*M, endowed with the natural symplectic structure $\omega = \mathrm{d}\theta$. Using the isomorphism $\mathrm{g}\colon \mathrm{T}M \to \mathrm{T}^*M$, we can transport this structure to the tangent bundle:[7]

$$\theta_L := \mathrm{g}^*\theta, \qquad \omega_L := \mathrm{d}\theta_L = \mathrm{g}^*\omega.$$

This way, $\mathrm{T}M$ becomes a symplectic manifold. As the Hamiltonian, we choose the energy function

$$E\colon \mathrm{T}M \to \mathbb{R}, \qquad E(X) := \frac{1}{2}\mathrm{g}(X, X). \tag{9.2.1}$$

[7] Cf. Remark 9.1.2/1 and recall that g is the Legendre transformation induced by the Lagrangian $L(X) = \frac{1}{2}\mathrm{g}(X, X)$.

The Hamiltonian vector field of E is referred to as the geodesic vector field of (M, g). Let us analyze the corresponding Hamilton equations in bundle coordinates q^i and v^i on $\mathrm{T}M$ induced by local coordinates q^i on M. Let q^i and p_i be the corresponding bundle coordinates on T^*M. In these coordinates, g is given by

$$(\mathbf{q}, \mathbf{v}) \mapsto \big(\mathbf{q}, \mathbf{p} = \mathrm{g}(\mathbf{v})\big)$$

and the tangent mapping is given by

$$\mathrm{g}'_{(\mathbf{q},\mathbf{v})}\big(X^i\partial_{q^i} + Y^i\partial_{v^i}\big) = X^i\partial_{q^i} + \big(\mathrm{g}_{ij}(\mathbf{q})Y^i + \mathrm{g}_{ij,k}(\mathbf{q})v^iX^k\big)\partial_{p_j},$$

where $\mathrm{g}_{ij,k} = \frac{\partial \mathrm{g}_{ij}}{\partial q^k}$. Therefore, using $\theta = p_k\,\mathrm{d}q^k$, we obtain

$$(\theta_L)_{(\mathbf{q},\mathbf{v})}\big(X^i\partial_{q^i} + Y^i\partial_{v^i}\big) = \theta\big(\mathrm{g}'_{(\mathbf{q},\mathbf{v})}\big(X^i\partial_{q^i} + Y^i\partial_{v^i}\big)\big) = \mathrm{g}_{kl}(\mathbf{q})v^lX^k.$$

Thus,

$$\theta_L = \mathrm{g}_{kl}v^l\,\mathrm{d}q^k, \qquad \omega_L = \mathrm{g}_{kl}\,\mathrm{d}v^l \wedge \mathrm{d}q^k + \mathrm{g}_{kl,j}v^l\,\mathrm{d}q^j \wedge \mathrm{d}q^k. \tag{9.2.2}$$

In these coordinates, the Hamiltonian takes the form

$$E(\mathbf{q}, \mathbf{v}) = \frac{1}{2}\mathrm{g}_{ij}(\mathbf{q})v^iv^j. \tag{9.2.3}$$

Thus,

$$\mathrm{d}E = \frac{1}{2}\mathrm{g}_{ij,k}v^iv^j\,\mathrm{d}q^k + \mathrm{g}_{ik}v^i\,\mathrm{d}v^k. \tag{9.2.4}$$

Using (9.2.4) and (9.2.3), from $X_E \lrcorner \omega_L = -\mathrm{d}E$ one reads off the following local formula for the Hamiltonian vector field (Exercise 9.2.1):

$$X_E = v^k\partial_{q^k} - v^iv^j\Gamma^k_{ij}\partial_{v^k}. \tag{9.2.5}$$

Here,

$$\Gamma^m_{ij} := \frac{1}{2}\mathrm{g}^{mk}(\mathrm{g}_{ik,j} + \mathrm{g}_{jk,i} - \mathrm{g}_{ij,k})$$

are the Christoffel symbols of the Levi-Civita connection[8] defined by the metric g. Thus, the Hamilton equations take the form

$$\dot{q}^k = v^k, \qquad \dot{v}^k = -v^iv^j\Gamma^k_{ij}.$$

The corresponding Euler-Lagrange equations are

$$\ddot{q}^k + \dot{q}^i\dot{q}^j\Gamma^k_{ij} = 0. \tag{9.2.6}$$

This is the geodesic equation on (M, g). Thus, the projections of the integral curves of X_E to M are geodesics of the Riemannian structure. A detailed discussion of models of this type can be found in [1], Sect. 3.7. We will come back to the geodesic flow in Example 10.6.1.

[8]In this example we use some facts from Riemannian geometry which go beyond the material of Sect. 4.4. In part II of this book we will give a concise treatment of the theory of Riemannian manifolds. For the time being, we refer to standard textbooks, see e.g. [166] or [76].

Example 9.2.2 (Hamiltonian systems on Lie groups) We take up Example 8.3.4 and consider a Hamiltonian system $(\mathrm{T}^*G, \omega, H)$, where G is a Lie group with Lie algebra $\mathfrak{g}$, ω is the canonical symplectic form on T^*G and $H \in C^\infty(\mathrm{T}^*G)$. We determine the Hamiltonian vector field X_H in the (inverse) left trivialization $G \times \mathfrak{g} \to \mathrm{T}^*G$ given by (8.3.6). For that purpose, for $(a, \mu) \in G \times \mathfrak{g}^*$ we make the ansatz

$$X_H(a, \mu) = \big(\mathrm{L}'_a A, \rho\big)$$

with $A \in \mathfrak{g}$ and $\rho \in \mathfrak{g}^*$ and analyze the defining equation

$$\omega\big(X_H(a, \mu), Y\big) = -\mathrm{d}H(Y)$$

for arbitrary tangent vectors $Y = (\mathrm{L}'_a B, \sigma)$ at (a, μ). By (8.3.8), the left hand side reads

$$\omega\big(X_H(a, \mu), Y\big) = \langle \rho, B\rangle - \langle \sigma, A\rangle - \big\langle \mu, [A, B]\big\rangle.$$

For the right hand side, we compute

$$\mathrm{d}H(Y) = \frac{\mathrm{d}}{\mathrm{d}t}{\restriction_0} H\big(a \exp(tB), \mu + t\sigma\big) = (H_a)'_\mu(\sigma) + (H_\mu)'_a \circ \mathrm{L}'_a(B)$$

with the induced functions $H_a : \mathfrak{g}^* \to \mathbb{R}$ and $H_\mu : G \to \mathbb{R}$. Viewing the linear mappings $(H_a)'_\mu : \mathfrak{g}^* \to \mathbb{R}$ and $(H_\mu)'_a \circ \mathrm{L}'_a : \mathfrak{g} \to \mathbb{R}$ as elements of $\mathfrak{g}$ and $\mathfrak{g}^*$, respectively, we obtain

$$\mathrm{d}H(Y) = \big\langle \sigma, (H_a)'_\mu\big\rangle + \big\langle (H_\mu)'_a \circ \mathrm{L}'_a, B\big\rangle.$$

We read off

$$A = (H_a)'_\mu, \qquad \rho = -(H_\mu)'_a \circ \mathrm{L}'_a - \mathrm{ad}^*\big((H_a)'_\mu\big)\mu. \tag{9.2.7}$$

As a consequence, the Hamilton equations are

$$\dot a = \mathrm{L}'_a\big((H_a)'_\mu\big), \qquad \dot\mu = -(H_\mu)'_a \circ \mathrm{L}'_a - \mathrm{ad}^*\big((H_a)'_\mu\big)\mu. \tag{9.2.8}$$

In the special case where H is invariant under the action of G by left translation, $H'_\mu = 0$ and the Hamilton equations read

$$\dot a = \mathrm{L}'_a\big((H_a)'_\mu\big), \qquad \dot\mu = -\mathrm{ad}^*\big((H_a)'_\mu\big)\mu. \tag{9.2.9}$$

Finally, using (8.3.8) and (9.2.7), for the Poisson bracket of $f, g \in C^\infty(\mathrm{T}^*G)$ we obtain

$$\begin{aligned}\{f, g\}(a, \mu) &= \omega\big(X_f(a, \mu), X_g(a, \mu)\big)\\ &= \big\langle -(f_\mu)'_a \circ \mathrm{L}'_a - \mathrm{ad}^*\big((f_a)'_\mu\big)\mu, (g_a)'_\mu\big\rangle\\ &\quad - \big\langle -(g_\mu)'_a \circ \mathrm{L}'_a - \mathrm{ad}^*\big((g_a)'_\mu\big)\mu, (f_a)'_\mu\big\rangle - \big\langle \mu, \big[(f_a)'_\mu, (g_a)'_\mu\big]\big\rangle,\end{aligned}$$

that is,

$$\begin{aligned}&\{f, g\}(a, \mu)\\ &\quad = \big\langle (g_\mu)'_a \circ \mathrm{L}'_a, (f_a)'_\mu\big\rangle - \big\langle (f_\mu)'_a \circ \mathrm{L}'_a, (g_a)'_\mu\big\rangle + \big\langle \mu, \big[(f_a)'_\mu, (g_a)'_\mu\big]\big\rangle.\end{aligned} \tag{9.2.10}$$

Models of this type are relevant, for example, in the theory of the top.

Example 9.2.3 (Hamiltonian systems on coadjoint orbits) Let G be a Lie group and let $\mathfrak{g}$ be the Lie algebra of G. By Example 8.2.18/3, the dual vector space $\mathfrak{g}^*$ carries a natural linear Poisson structure, which according to (8.2.18) is given by

$$\{f, h\}(\mu) = \langle \mu, \big[\mathrm{d}f(\mu), \mathrm{d}h(\mu)\big]\rangle$$

with $\mu \in \mathfrak{g}^*$ and $f, h \in C^\infty(\mathfrak{g}^*)$. Let $H \in C^\infty(\mathfrak{g}^*)$. As discussed in Sect. 8.2, the Hamiltonian vector field X_H on the Poisson manifold $\mathfrak{g}^*$ is defined to be the vector field corresponding to the derivation $\{H, \cdot\}$ of $C^\infty(\mathfrak{g}^*)$, that is, $X_H(f) = \{H, f\}$ for every $f \in C^\infty(\mathfrak{g}^*)$. Since

$$\big(X_H(f)\big)(\mu) = \{H, f\}(\mu) = \langle \mu, \big[\mathrm{d}H(\mu), \mathrm{d}f(\mu)\big]\rangle = -\langle \mathrm{ad}^*\big(\mathrm{d}H(\mu)\big)\mu, \mathrm{d}f(\mu)\rangle,$$

we read off

$$X_H(\mu) = -\mathrm{ad}^*\big(\mathrm{d}H(\mu)\big)\mu. \tag{9.2.11}$$

Consequently, the Hamilton equations associated with the Hamiltonian system $(\mathfrak{g}^*, \{,\}, H)$ are

$$\dot{\mu} = -\mathrm{ad}^*\big(\mathrm{d}H(\mu)\big)\mu. \tag{9.2.12}$$

In coordinates μ_i with respect to a basis in $\mathfrak{g}^*$, this reads

$$\dot{\mu}_j = -c^l_{jk}\mu_l \frac{\partial H}{\partial \mu_k},$$

where c^l_{jk} are the structure constants with respect to the dual basis in $\mathfrak{g}$. As a Hamiltonian vector field, X_H is tangent[9] to the symplectic leaves of $\mathfrak{g}^*$, cf. the remark after Theorem 8.2.20. By Proposition 8.4.3, the leaves coincide with the coadjoint orbits. Thus, the dynamics of the (Poisson) Hamiltonian system $(\mathfrak{g}^*, \{,\}, H)$ reduces to the coadjoint orbits, hence inducing a (symplectic) Hamiltonian system on each of them. As a consequence of this observation, every Ad^*-invariant function on $\mathfrak{g}^*$ is a constant of motion. Let us add that if $\mathfrak{g}$ admits an Ad-invariant scalar product, according to Remark 8.4.4, we can identify $\mathfrak{g}^*$ with $\mathfrak{g}$ and the adjoint representation with the coadjoint representation. Under this identification, the Hamilton equations take the form

$$\dot{\mu} = [\mu, \mathrm{d}H]. \tag{9.2.13}$$

One says that $(\mu, \mathrm{d}H)$ constitutes a Lax pair. In Chap. 11 we will discuss Lax pairs in some detail.

Exercises

9.2.1 Prove Formula (9.2.5).

[9]This also follows by direct inspection: from (6.2.3) and (9.2.11) we read off that $X_H(\mu)$ coincides—up to the sign—with the value at μ of the Killing vector field of the coadjoint action generated by $\mathrm{d}H(\mu)$.

9.2.2 Analyze the Poisson structure of Example 9.2.3 for the case of the Lie group SO(3). Compare with the symplectic structures on the coadjoint orbits found in Example 8.4.5/3 and write down the Hamilton equations for the quadratic Hamiltonian

$$H = \frac{1}{2}\sum_i a_i \mu_i^2, \qquad a_i > 0.$$

Which physical model is described by this Hamiltonian?

9.2.3 Consider the Euclidean group E(3), defined in Exercise 8.4.2. Show that the Poisson structure on the dual space $\mathfrak{g}^* \cong \mathbb{R}^6$ of E(3) is given in standard coordinates (μ_i, λ_i), $i = 1, 2, 3$, by the Poisson brackets

$$\{\mu_i, \mu_j\} = \varepsilon_{ij}{}^k \mu_k, \qquad \{\mu_i, \lambda_j\} = \varepsilon_{ij}{}^k \lambda_k, \qquad \{\lambda_i, \lambda_j\} = 0.$$

9.3 The Time-Dependent Picture

In this section we develop a calculus where the time variable is naturally integrated into the phase space of the system. One of the motivations for doing so is to deal with explicitly time-dependent Hamiltonians.

Thus, let (M, ω) be the phase space and let $H = \{H_t : t \in \mathbb{R}\}$ be a smooth family of Hamiltonian functions. To this smooth family there corresponds a smooth family of Hamiltonian vector fields, that is, a time-dependent vector field X_H, given by

$$(X_H)(t, m) = X_{H_t}(m).$$

Such Hamiltonian systems are called time-dependent or non-autonomous. The special case of an autonomous system is of course included. The geometric structure relevant for the description of such systems is the extended phase space

$$\tilde{M} = \mathrm{T}^*\mathbb{R} \times M, \qquad \tilde{\omega} := \mathrm{pr}_2^* \omega - \mathrm{pr}_1^*(\mathrm{d}E \wedge \mathrm{d}t), \tag{9.3.1}$$

where t and E denote the standard coordinates on $\mathrm{T}^*\mathbb{R} \cong \mathbb{R} \times \mathbb{R}$ and

$$\mathrm{pr}_1 : \tilde{M} \to \mathrm{T}^*\mathbb{R}, \qquad \mathrm{pr}_2 : \tilde{M} \to M,$$

are the natural projections. In what follows, the latter will usually be omitted. In local Darboux coordinates q^i and p_i on M, we have

$$\tilde{\omega} = \mathrm{d}p_i \wedge \mathrm{d}q^i - \mathrm{d}E \wedge \mathrm{d}t.$$

Points of $\tilde{M} \cong \mathbb{R} \times \mathbb{R} \times M$ will be denoted by (t, E, m).

Remark 9.3.1 For $M = \mathrm{T}^* Q$, the extended phase space $(\tilde{M}, \tilde{\omega})$ is isomorphic to the cotangent bundle of the extended configuration space $\tilde{Q} = \mathbb{R} \times Q$.

Now, let us define the extended[10] Hamiltonian function

$$\tilde{H} : \tilde{M} \to \mathbb{R}, \qquad \tilde{H}(t, E, m) := H(t, m) - E,$$

[10] Sometimes also called the suspended Hamiltonian.

and consider the extended Hamiltonian system $(\tilde{M}, \tilde{\omega}, \tilde{H})$. This is an ordinary (autonomous) Hamiltonian system. Thus, the Hamiltonian vector field $X_{\tilde{H}}$ of $\tilde{H}$ is defined by

$$X_{\tilde{H}} \lrcorner \tilde{\omega} = -\mathrm{d}\tilde{H}, \tag{9.3.2}$$

and we have $\mathscr{L}_{X_{\tilde{H}}} \tilde{\omega} = 0$, that is, the flow of $X_{\tilde{H}}$ is a local symplectomorphism. The relation between $X_{\tilde{H}}$ and the time-dependent vector field X_H is given by

$$X_{\tilde{H}} = \frac{\partial}{\partial t} + \frac{\partial H}{\partial t} \frac{\partial}{\partial E} + X_H \tag{9.3.3}$$

(Exercise 9.3.1). Here, X_H is trivially extended to $\tilde{M}$. Thus, the equations for the integral curves $s \mapsto (t(s), E(s), \gamma(s))$ of $X_{\tilde{H}}$ are

$$\frac{\mathrm{d}}{\mathrm{d}s} t(s) = 1, \qquad \frac{\mathrm{d}}{\mathrm{d}s} E(s) = \frac{\partial H}{\partial t}\big(t(s), \gamma(s)\big), \qquad \frac{\mathrm{d}}{\mathrm{d}s} \gamma(s) = X_H\big(t(s), \gamma(s)\big). \tag{9.3.4}$$

From the first of these equations we read off $t(s) = t_0 + s$, that is, up to a constant, the parameter s coincides with the time variable t. Moreover, an analogous calculation as in the proof of Proposition 9.1.9 yields

$$\frac{\mathrm{d}}{\mathrm{d}s} \tilde{H}\big(t(s), E(s), \gamma(s)\big) = 0,$$

that is, $\tilde{H}$ is a constant of motion. Therefore,

$$\tilde{H}\big(t_0 + s, E(s), \gamma(s)\big) = \tilde{H}\big(t_0, E(0), \gamma(0)\big)$$

and thus

$$E(s) = E(0) + H\big(t_0 + s, \gamma(s)\big) - H\big(t_0, \gamma(0)\big).$$

Consequently, the flows $\tilde{\Phi}$ of $X_{\tilde{H}}$ and Φ of X_H are related by

$$\tilde{\Phi}\big(s, (t, E, m)\big) = \big(t + s, E + H\big(t + s, \Phi_{t+s,t}(m)\big) - H(t, m), \Phi_{t+s,t}(m)\big). \tag{9.3.5}$$

Remark 9.3.2 In local Darboux coordinates q^i and p_i on M, the system of Eqs. (9.3.4) reads

$$\frac{\mathrm{d}t}{\mathrm{d}s}(s) = 1 \qquad\qquad \frac{\mathrm{d}E}{\mathrm{d}s}(s) = \frac{\partial H}{\partial t}\big(t(s), \mathbf{q}(s), \mathbf{p}(s)\big), \tag{9.3.6}$$

$$\dot{q}^i(s) = \frac{\partial H}{\partial p_i}\big(t(s), \mathbf{q}(s), \mathbf{p}(s)\big), \qquad \dot{p}_i(s) = -\frac{\partial H}{\partial q^i}\big(t(s), \mathbf{q}(s), \mathbf{p}(s)\big). \tag{9.3.7}$$

These are the Hamilton equations on the extended phase space. We note that, in contrast to $\tilde{H}$, H is not a constant of motion.

Now, let us consider the level sets of $\tilde{H}$. Since $\frac{\partial \tilde{H}}{\partial E} = -1$, every value of $\tilde{H}$ is regular and hence every level set is an embedded submanifold of $\tilde{M}$. Since any two

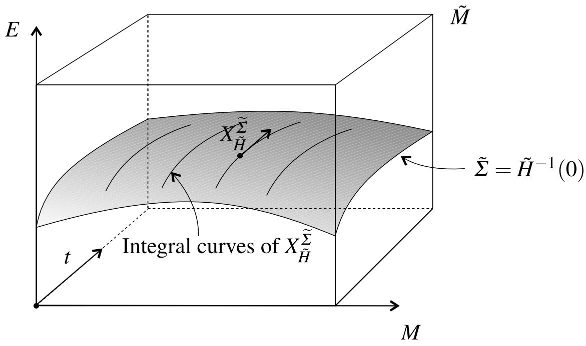


Fig. 9.2 The induced vector field $X_{\tilde{H}}^{\tilde{\Sigma}}$ spans the characteristic distribution $D^{\omega_{\tilde{\Sigma}}}$ on $\tilde{\Sigma}$

level sets are mapped onto one another by a shift in the variable E, it suffices to consider the level set

$$\tilde{\Sigma} := \tilde{H}^{-1}(0).$$

Let $\iota : \tilde{\Sigma} \to \tilde{M}$ denote the natural inclusion mapping. Since $\tilde{\Sigma}$ is defined by the equation $E = H(t, m)$, it is mapped to the graph of H under the reordering

$$\tilde{M} \to (\mathbb{R} \times M) \times \mathbb{R}, \qquad (t, E, m) \mapsto \big((t, m), E\big).$$

Suppressing this reordering, we will interpret the natural mapping

$$\mathrm{gr}_H : \mathbb{R} \times M \to \tilde{M}, \qquad (t, m) \mapsto \big(t, H(t, m), m\big), \tag{9.3.8}$$

as the graph embedding of H (therefore the notation). By Remarks 1.6.12/2 and 1.6.13/5, gr_H is a diffeomorphism onto $\tilde{\Sigma}$. Thus, $\tilde{\Sigma}$ is connected iff so is M. Since $\tilde{H}$ is a constant of motion, $X_{\tilde{H}}$ is tangent to $\tilde{\Sigma}$ and hence induces a vector field $X_{\tilde{H}}^{\tilde{\Sigma}}$ on $\tilde{\Sigma}$ which is ι-related to $X_{\tilde{H}}$. We have

$$X_{\tilde{H}}^{\tilde{\Sigma}} = (\mathrm{gr}_H)_*\left(\frac{\partial}{\partial t} + X_H\right), \tag{9.3.9}$$

where gr_H is viewed as a mapping to $\tilde{\Sigma}$ and both $\frac{\partial}{\partial t}$ and X_H are viewed as vector fields on $\mathbb{R} \times M$ (Exercise 9.3.2). Note that their sum coincides with the extended vector field $\overline{X_H}$, cf. Sect. 3.4. As shown in Sect. 9.1, the induced 2-form

$$\omega_{\tilde{\Sigma}} := \iota^* \tilde{\omega}$$

on $\tilde{\Sigma}$ has maximal rank and the characteristic distribution $D^{\omega_{\tilde{\Sigma}}} = \ker \omega_{\tilde{\Sigma}}$ is spanned by $X_{\tilde{H}}^{\tilde{\Sigma}}$, see Fig. 9.2.

Definition 9.3.3 (Time-dependent canonical transformation) A symplectomorphism $\tilde{\Phi}$ of the extended phase space $\tilde{M}$ is called a time-dependent canonical transformation of $\tilde{M}$ if $\tilde{\Phi}(\tilde{\Sigma}) \subset \tilde{\Sigma}$ and $\tilde{\Phi}^* t = t$.

Time-dependent canonical transformations of $\tilde{M}$ induce time-dependent canonical transformations of M:

Proposition 9.3.4 *Every time-dependent canonical transformation $\tilde{\Phi}$ of $\tilde{M}$ induces a smooth mapping $\Phi : \mathbb{R} \times M \to M$ by*

$$\tilde{\Phi}\big(t, H(t,m), m\big) = \big(t, H \circ \Phi(t,m), \Phi(t,m)\big).$$

For every fixed t, $\Phi(t,\cdot)$ is a canonical transformation of M.

The mapping Φ is called a time-dependent canonical transformation of M.

Proof Since $\tilde{\Phi}$ leaves $\tilde{\Sigma}$ invariant and since gr_H is an embedding, $\tilde{\Phi}$ induces a diffeomorphism $\hat{\Phi}$ of $\mathbb{R} \times M$ by

$$\mathrm{gr}_H \circ \hat{\Phi} = \tilde{\Phi} \circ \mathrm{gr}_H . \tag{9.3.10}$$

We set $\Phi := \mathrm{pr}_M \circ \hat{\Phi}$. Then, $\Phi_t := \Phi(t,\cdot)$ is a diffeomorphism of M for every t. To see that Φ_t is symplectic, define $\iota_t : M \to \mathbb{R} \times M$ by $\iota_t(m) := (t,m)$. Using (9.3.10) and $\hat{\Phi} \circ \iota_t = \iota_t \circ \Phi_t$, we obtain

$$\tilde{\Phi} \circ (\mathrm{gr}_H \circ \iota_t) = (\mathrm{gr}_H \circ \iota_t) \circ \Phi_t .$$

A straightforward computation shows that $(\mathrm{gr}_H \circ \iota_t)^* \tilde{\omega} = \omega$. Thus, on the one hand,

$$\big((\mathrm{gr}_H \circ \iota_t) \circ \Phi_t\big)^* \tilde{\omega} = \Phi_t^* \omega,$$

whereas on the other hand,

$$\big((\mathrm{gr}_H \circ \iota_t) \circ \Phi_t\big)^* \tilde{\omega} = (\mathrm{gr}_H \circ \iota_t)^* \circ \tilde{\Phi}^* \tilde{\omega} = \omega.$$

This yields the assertion. □

Next, we characterize time-dependent canonical transformations in terms of generating functions. Thus, let $\tilde{\Phi}$ be a time-dependent canonical transformation of $\tilde{M}$ and let $\Gamma_{\tilde{\Phi}} \subset \tilde{M} \times \tilde{M}$ be its graph. Let q^i, p_i and $\bar{q}^i$, $\bar{p}_i$ be Darboux coordinates on the first and the second copy of M, respectively. For $f \in C^\infty(\tilde{M})$, we write $\bar{f} = f \circ \tilde{\Phi}$. According to Sect. 8.8, the defining relation for a generating function $\tilde{S} = \tilde{S}(\mathbf{q}, \bar{\mathbf{q}}, t, \bar{t})$ of the first kind for $\tilde{\Phi}$ reads

$$\big(\bar{p}_i \, \mathrm{d}\bar{q}^i - \bar{E} \, \mathrm{d}\bar{t}\big) - \big(p_i \, \mathrm{d}q^i - E \, \mathrm{d}t\big) = -\mathrm{d}\tilde{S}. \tag{9.3.11}$$

Since $\bar{t} = t$ and since on $\tilde{\Sigma}$ we have $E = H$ and $\bar{E} = \bar{H}$, the restriction S of $\tilde{S}$ to $\Gamma_{\tilde{\Phi}} \cap (\tilde{\Sigma} \times \tilde{\Sigma})$ satisfies

$$\big(\bar{p}_i \, \mathrm{d}\bar{q}^i - \bar{H} \, \mathrm{d}t\big) - \big(p_i \, \mathrm{d}q^i - H \, \mathrm{d}t\big) = -\mathrm{d}S. \tag{9.3.12}$$

By comparison of coefficients, we obtain the relations

$$\bar{p}_i = -\frac{\partial S}{\partial \bar{q}^i}, \qquad p_i = \frac{\partial S}{\partial q^i}, \qquad \bar{H} = H + \frac{\partial S}{\partial t}. \tag{9.3.13}$$

Thus, in particular, for every fixed t, the function $(\mathbf{q}, \bar{\mathbf{q}}) \mapsto S(\mathbf{q}, \bar{\mathbf{q}}, t)$ is a generating function of the first kind for the canonical transformation $\Phi(t,\cdot)$ of M. The corresponding equations for the generating functions of the second, third and fourth kind are obtained by replacing the defining relations (9.3.13) by, respectively, (8.8.4), (8.8.5) and (8.8.6). The relation for the Hamiltonian holds for every choice of S.

Finally, let us study the geometry of $\tilde{\Sigma}$ under the assumption that $M = \mathrm{T}^*Q$ and $\omega = \mathrm{d}\theta$, where θ is the canonical 1-form on T^*Q.[11] In this situation, according to Remark 9.3.1, the extended phase space is given by $\tilde{M} = \mathrm{T}^*(\mathbb{R} \times Q)$. Moreover,

$$\tilde{\theta} = \mathrm{pr}_2^*\theta - \mathrm{pr}_1^*(E\,\mathrm{d}t) \tag{9.3.14}$$

is the canonical 1-form on $\tilde{M}$ and we have $\tilde{\omega} = \mathrm{d}\tilde{\theta}$.

Proposition 9.3.5 *The 1-form $\iota^*\tilde{\theta}$ is relatively invariant with respect to the vector field $X_{\tilde{H}}^{\tilde{\Sigma}}$. It endows $\tilde{\Sigma}$ with a (strict) contact structure iff the function $\iota^*(\theta(X_H) - H)$ vanishes nowhere on $\tilde{\Sigma}$.*

Proof The first assertion is obvious, because $\mathrm{d}(\iota^*\tilde{\theta}) = \omega_{\tilde{\Sigma}}$ and $X_{\tilde{H}}^{\tilde{\Sigma}}$ spans $D^{\omega_{\tilde{\Sigma}}}$. We prove the second assertion. Since $X_{\tilde{H}}^{\tilde{\Sigma}}$ and $X_{\tilde{H}}$ are ι-related, for every $m \in \tilde{\Sigma}$ we have

$$\big(\iota^*\tilde{\theta}\big)_m\big(X_{\tilde{H}}^{\tilde{\Sigma}}\big) = \tilde{\theta}_{\iota(m)}(X_{\tilde{H}}) = \big(\theta(X_H) - H\big)\big(\iota(m)\big).$$

Since $X_{\tilde{H}}$ vanishes nowhere, we conclude that $\iota^*\tilde{\theta}$ is non-degenerate on $D^{\omega_{\tilde{\Sigma}}}$ iff the function $\iota^*(\theta(X_H) - H)$ vanishes nowhere. In this case, $\ker(\iota^*\tilde{\theta})$ is a hyperplane distribution on $\tilde{\Sigma}$. To prove that it is a contact structure on $\tilde{\Sigma}$, we show that $(\iota^*\tilde{\theta}) \wedge (\mathrm{d}(\iota^*\tilde{\theta}))^n$ is a volume form on $\tilde{\Sigma}$ and apply Proposition 8.5.12. Since $\iota^*\tilde{\theta}$ is non-degenerate on $D^{\omega_{\tilde{\Sigma}}}$, the distribution $\ker(\iota^*\tilde{\theta})$ is complementary to $D^{\omega_{\tilde{\Sigma}}}$. Thus, it suffices to show that $(\mathrm{d}(\iota^*\tilde{\theta}))^n = \omega_{\tilde{\Sigma}}^n$ is non-degenerate on $\ker(\iota^*\tilde{\theta})$. Since $\tilde{\omega}^n$ is non-degenerate, the kernel of $\omega_{\tilde{\Sigma}}^n = \iota^*\tilde{\omega}^n$ has dimension at most 1. Since it contains $\ker\omega_{\tilde{\Sigma}} = D^{\omega_{\tilde{\Sigma}}}$, it coincides with the latter. Thus, $(\mathrm{d}(\iota^*\tilde{\theta}))^n$ is non-degenerate on $\ker(\iota^*\tilde{\theta})$, indeed. □

Remark 9.3.6

1. Obviously, the forms $\tilde{\omega}, \tilde{\omega}^2, \dots, \tilde{\omega}^n$ are integral invariants of the extended Hamiltonian vector field $X_{\tilde{H}}$ and their pull-backs $\iota^*\tilde{\omega}, \iota^*\tilde{\omega}^2, \dots, \iota^*\tilde{\omega}^n$ to $\tilde{\Sigma}$ are integral invariants of the induced vector field $X_{\tilde{H}}^{\tilde{\Sigma}}$.
2. The 1-form $\iota^*\tilde{\theta}$ is called the Poincaré-Cartan integral invariant. Since on $\tilde{\Sigma}$ we have $E = H(t, m)$, in Darboux coordinates it is given by
$$\iota^*\tilde{\theta} = p_i\,\mathrm{d}q^i - H\,\mathrm{d}t.$$
3. Since the characteristic distribution $D^{\omega_{\tilde{\Sigma}}}$ of $\omega_{\tilde{\Sigma}}$ is spanned by $X_{\tilde{H}}^{\tilde{\Sigma}}$, it is trivial as a vector bundle and hence orientable. Thus, under the additional assumption

[11] More generally, we could assume that the symplectic form ω is exact.

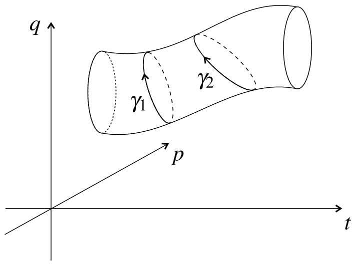


Fig. 9.3 Closed curves γ_1 and γ_2 in $\tilde{\Sigma}$ which have the same flow cylinder

that $\tilde{\Sigma}$ be compact, it is a hypersurface of contact type. The reader should compare Proposition 9.3.5 with Example 8.5.18/5, which states that $T^*Q \times \mathbb{R}$ is a strict contact manifold without any further assumptions. The reader should also convince himself that, for the class of models defined by (9.1.14), the mysterious function $\theta(X_H) - H$ is nothing but the Legendre transform $L \circ (FL)^{-1}$ of the Lagrangian L. If it does not vanish, $(\theta(X_H) - H)^{-1} X^{\tilde{\Sigma}}_{\tilde{H}}$ is the Reeb vector field of the contact form $\iota^* \tilde{\theta}$, cf. Remark 8.5.17.

4. The reader who feels uncomfortable with the distinguished role which the global time variable t plays in this section should consult Sect. 16 of Chap. V in the book of Libermann and Marle [181] for a more intrinsic treatment.

The Poincaré-Cartan integral invariant has the following interesting property. To formulate it, we observe that, under the flow of $X^{\tilde{\Sigma}}_{\tilde{H}}$, every closed (oriented) curve γ in $\tilde{\Sigma}$ whose tangent vectors are nowhere parallel to $X^{\tilde{\Sigma}}_{\tilde{H}}$ generates a two-dimensional oriented submanifold diffeomorphic to $S^1 \times \mathbb{R}$, called the flow cylinder of γ. Here, the orientation of $\mathbb{R}$ is induced by the time evolution. Proposition 4.2.16 implies

Corollary 9.3.7 *If γ_1 and γ_2 are closed curves in $\tilde{\Sigma}$ whose flow cylinders coincide as oriented submanifolds, see Fig. 9.3, then $\int_{\gamma_1} \iota^* \tilde{\theta} = \int_{\gamma_2} \iota^* \tilde{\theta}$.*

For a further discussion of the Poincaré-Cartan integral invariant we refer to the book of Arnold[12] [18].

Exercises

9.3.1 Prove Formula (9.3.3).

9.3.2 Prove Formula (9.3.9).

[12] In this book, Arnold develops an almost philosophical attitude towards this integral invariant by showing that it can be taken as a starting point for building Hamiltonian mechanics. Then, also symplectic geometry is a derived structure. From a more modern point of view, the reader can find a lot of interesting thoughts about the unifying power of symplectic and especially of contact geometry in another paper by Arnold [22], which ends with the statement that "contact geometry is all geometry".

Hint. Show that the mapping gr_H, given by (9.3.8), satisfies

$$\mathrm{gr}'_H \circ \frac{\partial}{\partial t} = \left(\frac{\partial}{\partial t} + \frac{\partial H}{\partial t}\frac{\partial}{\partial E}\right) \circ \mathrm{gr}_H .$$

9.4 Regular Energy Surfaces and Symplectic Capacities

In this section we study the geometry of regular energy surfaces. From this, we derive some basic statements about the existence of periodic integral curves in autonomous systems.

Thus, let (M, ω, H) be a Hamiltonian system of dimension $2n$ and let Φ denote the flow of the Hamiltonian vector field X_H. For a regular energy surface Σ of H, let $\iota : \Sigma \to M$ be the natural inclusion mapping and X_H^Σ the vector field on Σ induced by X_H. As before, denote $\omega_\Sigma := \iota^*\omega$ and let $D^{\omega_\Sigma} := \ker \omega_\Sigma$ be the characteristic distribution of ω_Σ, cf. Sect. 9.1. Recall that D^{ω_Σ} is spanned by X_H^Σ and that integral curves of sections of D^{ω_Σ} are called characteristics of D^{ω_Σ}.

On a regular energy surface, $\mathrm{d}H$ does not vanish. This implies

Proposition 9.4.1 *Let (M, ω, H) be a Hamiltonian system and let Σ be a regular energy surface of H. The symplectic form ω induces a natural volume form μ on Σ. This form is invariant under the flow of the induced vector field X_H^Σ.*

Proof Let Ω_ω be the canonical volume form on M induced by ω, cf. (8.1.3). Since $\mathrm{d}H$ vanishes nowhere on Σ, it vanishes nowhere on some open neighbourhood $U \subset M$ of Σ as well. Hence, there exists a $(2n-1)$-form σ on U such that $\Omega_\omega = \mathrm{d}H \wedge \sigma$ on U. We put $\mu := \iota^*\sigma$. This form does not depend on the choice of σ: let σ' be a second local $(2n-1)$-form fulfilling $\Omega_\omega = \mathrm{d}H \wedge \sigma'$ on some open neighbourhood U' of Σ. Then, $(\sigma - \sigma') \wedge \mathrm{d}H = 0$ on $U \cap U'$, so that there exists a $2(n-1)$-form ρ on $U \cap U'$ satisfying $\sigma - \sigma' = \mathrm{d}H \wedge \rho$. But $\iota^*(\sigma - \sigma') = \iota^*(\mathrm{d}H \wedge \rho) = 0$, because $\iota^* H$ is constant. It remains to show that μ is X_H^Σ-invariant. Since

$$0 = \mathscr{L}_{X_H} \Omega_\omega = \mathrm{d}H \wedge \mathscr{L}_{X_H}\sigma,$$

there exists a $2(n-1)$-form τ on U such that $\mathscr{L}_{X_H}\sigma = \mathrm{d}H \wedge \tau$. Thus,

$$\mathscr{L}_{X_H^\Sigma}(\mu) = \iota^*\mathscr{L}_{X_H}\sigma = \iota^*(\mathrm{d}H \wedge \tau) = 0.$$ □

The existence of an invariant volume form on a regular energy surface has the following important consequence.

Theorem 9.4.2 (Recurrence Theorem of Poincaré) *Let (M, ω, H) be a Hamiltonian system and let Σ be a compact regular energy surface. With respect to the measure on Σ defined by the induced volume form μ, almost every point $m \in \Sigma$ is a recurrent point, that is, there exists a sequence $\{t_j\}$ with $t_j \to \infty$ such that the flow Φ of X_H fulfils*

$$\lim_{j\to\infty} \Phi_{t_j}(m) = m.$$

Proof We denote the measure on Σ defined by the induced volume form μ by the same symbol. Since Σ is compact, the flow of X_H is complete. Let $\tilde{\Sigma}$ denote the subset of Σ of points which are not recurrent. We show that $\mu(\tilde{\Sigma}) = 0$. Let $\{B_\alpha : \alpha = 1, 2, 3, \ldots\}$ be a basis of the topology of Σ. For every α, let $\tilde{B}_\alpha$ denote the subset of B_α of points m for which there exists t_0 such that $\Phi_t(m) \notin B_\alpha$ for all $t \geq t_0$. We claim that

$$\tilde{\Sigma} = \bigcup_{\alpha=1}^{\infty} \tilde{B}_\alpha. \tag{9.4.1}$$

The inclusion $\supset$ is obvious. Conversely, assume that $m \notin \bigcup_{\alpha=1}^{\infty} \tilde{B}_\alpha$. The numbers α such that $m \in B_\alpha$ form a sequence $\{\alpha_j : j = 1, 2, 3, \ldots\}$. Since $m \in B_{\alpha_j} \setminus \tilde{B}_{\alpha_j}$ for every j, we can construct a sequence $\{t_j\}$ such that $t_j \to \infty$ and $\Phi_{t_j}(m) \in B_{\alpha_j}$. Since $\{B_{\alpha_j}\}$ is a neighbourhood basis for m, this implies that $\lim_{j\to\infty} \Phi_{t_j}(m) = m$, that is, m is recurrent. Thus, (9.4.1) holds, indeed, and it suffices to show that $\mu(\tilde{B}_\alpha) = 0$ or, equivalently,

$$\mu(B_\alpha \setminus \tilde{B}_\alpha) = \mu(B_\alpha) \tag{9.4.2}$$

for all α. For that purpose, let α be fixed and consider the sequence of subsets

$$B_\alpha^k := \bigcup_{j \geq k} \Phi_{-j}(B_\alpha), \quad k = 0, 1, 2, \ldots.$$

Since $\Phi_k(B_\alpha^k) = B_\alpha^0$ and since Φ preserves the measure, we get $\mu(B_\alpha^k) = \mu(B_\alpha^0)$. Then, by compactness of Σ, we have $\mu(B_\alpha^0) < \infty$ and hence $\mu(B_\alpha^0 \setminus B_\alpha^k) = 0$ for all k. Since $B_\alpha \subset B_\alpha^0$, this implies $\mu(B_\alpha \setminus B_\alpha^k) = 0$ for all k and hence

$$\mu\left(\bigcup_{k\geq 0}\left(B_\alpha \setminus B_\alpha^k\right)\right) = 0.$$

In view of the disjoint decomposition $B_\alpha = (B_\alpha \bigcap_{k\geq 0} B_\alpha^k) \cup (\bigcup_{k\geq 0}(B_\alpha \setminus B_\alpha^k))$, we conclude that

$$\mu(B_\alpha) = \mu\left(B_\alpha \bigcap_{k\geq 0} B_\alpha^k\right).$$

Using $B_\alpha \bigcap_{k\geq 0} B_\alpha^k = B_\alpha \setminus \tilde{B}_\alpha$, we obtain (9.4.2). This proves the theorem. □

From the proof of Theorem 9.4.2 it is clear that instead of requiring Σ to be compact, it is enough to assume that X_H^Σ be complete and that $\int_\Sigma \mu < \infty$. It is also clear that a similar statement holds for the Liouville measure on M and any measurable subset of M which is invariant under the Hamiltonian flow.

Obviously, the Recurrence Theorem does not imply the existence of closed integral curves. From the point of view of physics, the problem whether a Hamiltonian system possesses periodic integral curves is of great importance.[13] This turns out to

[13]This problem has its origin in celestial mechanics. In particular, it is interesting to ask whether our planetary systems admits periodic orbits, that is, whether there exist initial conditions to which the planets would return after a finite time.

be a deep and difficult question, which has been intensively studied. For a comprehensive presentation we refer to the books of Hofer and Zehnder [139] and McDuff and Salamon [206]. Here, we give an introduction to the subject.

We start with the crucial observations that the existence of periodic integral curves on a regular energy surface does not depend on the choice of the Hamiltonian.

Lemma 9.4.3

1. *Let H and F be Hamiltonian functions on M and let Σ be a regular energy surface of both H and F. Then, there exists a nowhere-vanishing function $\lambda : \Sigma \to \mathbb{R}$ such that $X_F^\Sigma = \lambda X_H^\Sigma$. Up to reparameterizations, X_H^Σ and X_F^Σ have the same integral curves.*
2. *Let H be a Hamiltonian function on M and let Σ be a regular energy surface of H. Up to a reparameterization, every non-constant integral curve of a vector field on M with values in $D^{\omega\Sigma}$ coincides with an integral curve of X_H.*

Proof 1. Under the assumptions made, $\mathrm{d}H_{\restriction\Sigma}$ and $\mathrm{d}F_{\restriction\Sigma}$ are nowhere-vanishing sections of the annihilator of $\mathrm{T}\Sigma$ in $\mathrm{T}M_{\restriction\Sigma}$. Since the annihilator has dimension 1, there exists a nowhere-vanishing smooth function λ on Σ such that $\mathrm{d}F = \lambda\,\mathrm{d}H$ and hence $X_F^\Sigma = \lambda X_H^\Sigma$. Let Φ^H and Φ^F denote the flows of X_H^Σ and X_F^Σ with domains $\mathscr{D}^H$ and $\mathscr{D}^F$, respectively. Making the ansatz $\Phi_t^F(m) = \Phi^H_{\tau(t,m)}(m)$ for some smooth function $\tau : \mathscr{D}^F \to \mathbb{R}$ satisfying $\tau(0,m) = 0$, we find

$$X_F^\Sigma\big(\Phi_t^F(m)\big) = \frac{\mathrm{d}}{\mathrm{d}t}\Phi_t^F(m) = \frac{\mathrm{d}}{\mathrm{d}t}\Phi^H_{\tau(t,m)}(m) = \frac{\mathrm{d}}{\mathrm{d}s}_{\restriction t}\tau(s,m) X_H^\Sigma\big(\Phi_t^F(m)\big)$$

and hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\tau(t,m) = \lambda\big(\Phi_t^F(m)\big),$$

which yields

$$\tau(t,m) = \int_0^t \lambda\big(\Phi_s^F(m)\big)\,\mathrm{d}s.$$

Since λ is nowhere vanishing on Σ we may interchange F and H in this argument. This shows that the mapping $(t,m) \mapsto (\tau(t,m),m)$ is a bijection (in fact, a diffeomorphism) from $\mathscr{D}^F$ onto $\mathscr{D}^H$.

2. Let X be a section of $D^{\omega\Sigma}$ and let γ be a non-constant integral curve of X. Then, X does not have equilibria in a neighbourhood of γ. By possibly modifying X outside this neighbourhood, we may assume that it vanishes nowhere on Σ. Since $D^{\omega\Sigma}$ is spanned by X_H^Σ, there exists a nowhere-vanishing smooth function λ on Σ such that $X = \lambda X_H^\Sigma$. Now, by replacing X_F^Σ by X in the proof of point 1 we obtain the assertion. □

Since the integral curves of X_H coincide with the characteristics of $D^{\omega\Sigma}$, cf. Sect. 9.1, the question

Does a regular energy surface Σ admit a periodic solution of X_H?

can be formulated purely geometrically:

Does a regular energy surface Σ admit a closed characteristic of D^{ω_Σ}?

An important class of energy surfaces for which this question can be successfully addressed is that of energy surfaces which are hypersurfaces of contact type, cf. Definition 8.5.23. The following proposition shows that in the typical Hamiltonian systems met in physics, such energy surfaces appear frequently.

Proposition 9.4.4 *Let* $(\mathrm{T}^*Q, \mathrm{d}\theta, H)$ *be a Hamiltonian system with* H *being of the form*

$$H(\xi) = T(\xi) + V\big(\pi(\xi)\big), \quad \xi \in \mathrm{T}^*Q,$$

cf. (9.1.14), *and let* Σ *be a regular energy surface which does not intersect the zero section of* T^*Q. *Then,* Σ *is a strict contact manifold with contact form* $\iota^*\theta$. *If, additionally,* Σ *is compact, then it is a hypersurface of contact type.*

Proof By assumption, $\theta(X_H) = 2T$ vanishes nowhere on Σ. Hence, $\iota^*\theta(X_H^\Sigma)$ is nowhere-vanishing function on Σ and the proof that Σ is a strict contact manifold with contact form $\iota^*\theta$ is completely analogous to that of the if-direction of Proposition 9.3.5. It is, therefore, left to the reader (Exercise 9.4.1). The rest is obvious. □

Now, let Σ be a hypersurface of contact type. Concerning our question, the first fundamental results were obtained in 1978 by Rabinowitz [249] and Weinstein [306]. They showed that under some additional topological assumptions, a hypersurface of contact type in $\mathbb{R}^{2n}$ admits a closed characteristic. In 1979, Weinstein invented the general definition of a hypersurface of contact type, cf. [307], and formulated the

Weinstein conjecture *Every hypersurface* Σ *of contact type fulfilling* $H^1(\Sigma) = 0$ *carries a closed characteristic.*

The next milestone is due to Viterbo [299]. He proved that every hypersurface of contact type in $\mathbb{R}^{2n}$ carries a closed characteristic. Here, the assumption that $H^1(\Sigma) = 0$ is not necessary. A breakthrough was made by Ekeland, Hofer and Zehnder who observed that a special type of symplectic invariants, called capacities [84, 138], can be used to tackle the Weinstein conjecture. Let us give an introduction to this approach. For a detailed presentation we refer again to the books of Hofer and Zehnder [139] and McDuff and Salamon [206]. Let

$$B_{2n}(r) = \big\{(\mathbf{q}, \mathbf{p}) \in \mathbb{R}^{2n} : \| \mathbf{q} \|^2 + \| \mathbf{p} \|^2 < r^2\big\},$$
$$Z_{2n}(r) = \big\{(\mathbf{q}, \mathbf{p}) \in \mathbb{R}^{2n} : \big(q^1\big)^2 + (p_1)^2 < r^2\big\}$$

denote, respectively, the open ball and the open cylinder of radius r in $\mathbb{R}^{2n}$, with the symplectic structure induced from the canonical symplectic structure ω_0 on $\mathbb{R}^{2n}$.

Definition 9.4.5 A symplectic capacity is a mapping c from the class of all symplectic manifolds, possibly with boundary, of a fixed dimension $2n$ to the positive real numbers (including infinity) fulfilling the following axioms:

1. Monotonicity: if there exists a symplectic embedding $\varphi : (M, \omega) \to (N, \sigma)$, then
$$c(M, \omega) \leq c(N, \sigma).$$
2. Conformality: $c(M, a\omega) = |a| c(M, \omega)$ for all $a \in \mathbb{R}$, $a \neq 0$.
3. Nontriviality: $c(B_{2n}(1), \omega_0) = \pi = c(Z_{2n}(1), \omega_0)$.

Every capacity is a symplectic invariant. Indeed, if $\varphi : (M, \omega) \to (N, \sigma)$ is a symplectomorphism, then φ^{-1} is a symplectomorphism, too. Application of Axiom 1 to both of these mappings yields $c(M, \omega) = c(N, \sigma)$. While the symplectic volume is a capacity for $n = 1$ (Exercise 9.4.2), Axiom 3 implies that it cannot be a capacity for $n > 1$, because it is infinite for the cylinder. Thus, provided capacities exist, they yield new symplectic invariants, different from the volume. For two open subsets U and V of (M, ω) fulfilling $U \subset V$, Axiom 1 implies $c(U) \leq c(V)$. To extend c to arbitrary subsets $A \subset M$, one puts
$$c(A) = \inf\{c(U) : A \subset U \text{ open}\}.$$
From the axioms we read off the capacities of balls and cylinders of radius r in $(\mathbb{R}^{2n}, \omega_0)$ (Exercise 9.4.3):
$$c\big(B_{2n}(r)\big) = \pi r^2 = c\big(Z_{2n}(r)\big). \tag{9.4.3}$$
This implies in particular that the closure of the open ball of radius r has capacity πr^2, too.

Remark 9.4.6 Assuming that a capacity exists, as an immediate consequence of (9.4.3) one obtains the famous Nonsqueezing Theorem of Gromov, which states that the ball $B_{2n}(1)$ cannot be embedded symplectically into the cylinder $Z_{2n}(r)$ unless $r \geq 1$.

Now, let us introduce the capacity of Hofer and Zehnder: consider a symplectic manifold (M, ω, H), possibly with boundary. Let $\mathscr{H}(M, \omega)$ denote the set of Hamiltonian functions H satisfying the following.

(a) There exists a compact set $K \subset \mathrm{Int}(M)$ such that H is constant on $M \setminus K$ and $0 \leq H(m) \leq H(M \setminus K)$ for all $m \in M$.
(b) There is a nonempty open subset $U \subset M$ on which H vanishes.

Under these assumptions, $H(M \setminus K) = \max(H)$ and X_H is complete. We call a Hamiltonian function $H \in \mathscr{H}(M, \omega)$ admissible iff it does not have a periodic integral curve of period $T \leq 1$. Let us denote the set of admissible Hamiltonians by $\mathscr{H}_{ad}(M, \omega)$. The Hofer-Zehnder capacity is defined by
$$c_{HZ}(M, \omega) := \sup\{\max(H) : H \in \mathscr{H}_{ad}(M, \omega)\}. \tag{9.4.4}$$
Note that c_{HZ} can be expressed in terms of the Hofer norm (8.8.14),
$$c_{HZ}(M, \omega) = \sup\{\| H \| : H \in \mathscr{H}_{ad}(M, \omega)\}.$$

Theorem 9.4.7 (Hofer-Zehnder) *The function c_{HZ} is a symplectic capacity.*

While it is easy to show that c_{HZ} fulfils the axioms of monotonicity and conformality (Exercise 9.4.4), it is hard to prove that it fulfils the axiom of nontriviality, see [139]. The proof requires techniques from the calculus of variations. We also refer to [139], Sect. 3.5, for examples where c_{HZ} can be calculated explicitly.

Let us derive a criterion for the existence of closed integral curves in terms of the Hofer-Zehnder capacity.

Lemma 9.4.8 *Let (M, ω, H) be a Hamiltonian system and let $\Sigma \subset H^{-1}(e)$ be a compact regular energy surface. There exist $\varepsilon > 0$, an open neighbourhood U of Σ and a diffeomorphism $\varphi : (e - \varepsilon, e + \varepsilon) \times \Sigma \to U$ mapping $\{E\} \times \Sigma$ to a regular energy surface Σ_E of H of energy E.*

Proof Choose a Riemannian metric on M and denote the corresponding gradient vector field of H by ∇H. There exists an open neighbourhood V of Σ where ∇H is nowhere vanishing. There, we can define the normalized gradient vector field

$$\hat{X} = \frac{\nabla H}{\| \nabla H \|^2}.$$

As in the proof of Proposition 8.9.6, we can extend $\hat{X}$ from some smaller open neighbourhood $W \subset V$ of Σ to a vector field on M with compact support whose flow Φ satisfies

$$H\big(\Phi_t(m)\big) = e + t$$

for all $m \in \Sigma$ and t such that $\Phi_t(m) \in W$. It follows that there exists $\varepsilon > 0$ and an open neighbourhood $U \subset W$ of Σ such that Φ defines a diffeomorphism

$$\varphi : (e - \varepsilon, e + \varepsilon) \times \Sigma \to U, \qquad \varphi(E, m) = \Phi_{E-e}(m).$$

By construction, $\varphi(\{E\} \times \Sigma)$ is a regular energy surface of H with energy E. □

As a consequence of the Lemma, U is foliated by the energy surfaces Σ_E,

$$U = \bigcup_{E \in (e-\varepsilon, e+\varepsilon)} \Sigma_E.$$

Theorem 9.4.9 (Hofer-Zehnder) *Let (M, ω, H) be a Hamiltonian system and let $\Sigma \subset H^{-1}(e)$ be a compact regular energy surface. Assume that there exists an open neighbourhood U of Σ with finite Hofer-Zehnder capacity. Then, there exists $\varepsilon > 0$ such that for a dense set of parameters $E \in (e - \varepsilon, e + \varepsilon)$ the energy surface defined by E contains a periodic integral curve of X_H.*

Note that the periodic solutions provided by this theorem in general do not lie on Σ but only nearby.

Proof We follow [139, §4.1, Thm. 1]. According to Lemma 9.4.8, by possibly shrinking U we may assume that there exists $\varepsilon > 0$ and a diffeomorphism $\varphi : (e-\varepsilon, e+\varepsilon) \times \Sigma \to U$ satisfying $\varphi(\{E\} \times \Sigma) = \Sigma_E$. Moreover, by Lemma 9.4.3/1, we have the freedom to choose a convenient Hamiltonian function which has Σ as an energy surface. Thus, let us choose a real number δ fulfilling $0 < \delta < \varepsilon$ and a smooth function $f\colon (e-\varepsilon, e+\varepsilon) \to \mathbb{R}$ fulfilling

$$\begin{aligned} &f(E) = c_{HZ}(U,\omega) + 1 && \text{for } E \le e-\delta \text{ or } E \ge e+\delta \\ &f(E) = 0 && \text{for } e - \frac{\delta}{2} \le E \le e + \frac{\delta}{2} \\ &f'(E) < 0 && \text{for } e - \delta < E < e - \frac{\delta}{2} \\ &f'(E) > 0 && \text{for } e + \frac{\delta}{2} < E < e + \delta \end{aligned}$$

and let us consider the function $F = f \circ H\colon M \to \mathbb{R}$. Since $F \in \mathscr{H}(M,\omega)$ and $\max(F) > c_{HZ}(U,\omega)$, X_F possesses a periodic integral curve γ (with period $T \le 1$). This curve must be contained in the open subset

$$\tilde{U} := \bigcup_{E \in (e-\delta, e-\frac{\delta}{2}) \cup (e+\frac{\delta}{2}, e+\delta)} \Sigma_E$$

of U, because X_F vanishes outside. On the other hand, since $\mathrm{d}F = (f' \circ H)\,\mathrm{d}H$ and since $f' \circ H$ is nowhere vanishing on $\tilde{U}$, F and H have the same energy surfaces in $\tilde{U}$. Now, Lemma 9.4.3/1 yields that, up to a reparameterization, γ is an integral curve of X_H. Thus, for arbitrarily small $0 < \delta < \varepsilon$, we find a periodic integral curve of X_H with energy between $e-\delta$ and $e+\delta$. In the above argument, e can be replaced by any value $e' \in (e-\varepsilon, e+\varepsilon)$. Then, δ has to be chosen so that

$$0 < \delta < \min\{e' - e + \varepsilon, e + \varepsilon - e'\}.$$

This yields the assertion. □

Now, let us turn back to hypersurfaces of contact type. Corollary 8.5.26 implies

Corollary 9.4.10 *Let (M,ω) be a symplectic manifold and let Σ be a hypersurface of contact type. If Σ admits an open neighbourhood U with finite Hofer-Zehnder capacity, then the characteristic distribution D^{ω_Σ} of Σ possesses a closed characteristic.*

Proof By Proposition 8.5.25 and by the monotonicity of c_{HZ}, we may assume that on U there exists a Liouville vector field Z transversal to Σ. Then, by Corollary 8.5.26, there exists $\varepsilon > 0$ such that the flow Φ of Z induces a diffeomorphism

$$\varphi : (-\varepsilon, \varepsilon) \times \Sigma \to U, \qquad \varphi(t,m) = \Phi_t(m).$$

Define $H : U \to \mathbb{R}$ by $H := \mathrm{pr}_1 \circ \varphi^{-1}$ and consider the Hamiltonian system (U, ω, H). By construction, the hypersurfaces $\varphi(\{E\} \times \Sigma)$ are the energy surfaces of H. By Theorem 9.4.9, X_H possesses a periodic integral curve γ on the hypersurface $\Sigma_E := \varphi(\{E\} \times \Sigma)$ for some $E \in (-\varepsilon, \varepsilon)$. Since Σ_E is an energy surface of H, the induced vector field $X_H^{\Sigma_E}$ takes values in the characteristic distribution of ω_{Σ_E} and γ is a characteristic of the latter. By point 2 of Corollary 8.5.26, then $\mathrm{pr}_2 \circ \varphi^{-1} \circ \gamma$ is a characteristic of D^{ω_Σ}. □

Remark 9.4.11

1. There are examples of hypersurfaces in $\mathbb{R}^{2n}$, $n \geq 3$, which do not admit any closed characteristic [104, 105]. Thus, the condition on Σ to be of contact type cannot be removed.
2. Using the Hofer-Zehnder capacity, one can give an alternative proof of the result of Viterbo [299] stating that every hypersurface of contact type in $\mathbb{R}^{2n}$ admits a closed characteristic, see [206, Thm. 12.32] or Sect. 4.3 of [139].
3. For a discussion of the special case $M = \mathrm{T}^*Q$, endowed with the canonical symplectic structure, and H belonging to the class defined by (9.1.14) we refer to [139]. In this case, some nice geometrical ideas can be applied. In particular, for $E > \max V$ the problem of finding closed characteristics reduces to the problem of finding closed geodesics for a special metric on T^*Q, called the Jacobi metric, which is constructed from the Hamiltonian.

More generally, one can consider the case of an abstract contact manifold, that is, one can discard the Hamiltonian and the ambient symplectic manifold. Then the Weinstein conjecture reads as follows. Does a Reeb vector field on a compact contact manifold admit a periodic integral curve? There are partial results concerning this question, see [135] and [136]. For a survey on the Weinstein conjecture and related aspects, in particular those of nearby and so-called almost existence theorems, we refer to [106]. Some aspects concerning the behaviour of a Hamiltonian system near a critical integral curve will be presented in the next section.

Exercises

9.4.1 Carry over the proof of Proposition 9.3.5 to Proposition 9.4.4.

9.4.2 Show that for any 2-dimensional manifold (M, ω), the total area $|\int_M \omega|$ is a symplectic capacity.

9.4.3 Prove Eq. (9.4.3).

9.4.4 Prove that the Hofer-Zehnder capacity, defined by (9.4.4), fulfils the axioms of monotonicity and conformality.

9.5 The Poincaré Mapping and Orbit Cylinders

The aim of this section is to study the dynamics of a Hamiltonian system near a periodic integral curve. An important tool for this is the Poincaré mapping introduced in Sect. 3.7. As in that section, depending on the context, by an integral curve we mean the curve itself or the corresponding submanifold, cf. Proposition 3.2.11. If γ is periodic, all of its points are regular with respect to H. Therefore, when studying the behaviour of the flow in the vicinity of γ, by modifying H outside of some neighbourhood of γ we may assume H to be a submersion if necessary. In particular, there exists an open neighbourhood of γ in $H^{-1}(H(\gamma))$ which is an embedded submanifold of M. We will refer to such a neighbourhood as a local energy surface of γ.

Our first aim is to show that in the case of a Hamiltonian system the decomposition of the phase space into energy surfaces induces a decomposition of every Poincaré mapping into a family of symplectomorphisms.

Proposition 9.5.1 *Let (M, ω, H) be a Hamiltonian system and let γ be a periodic integral curve of X_H. Let $m_0 \in \gamma$ and let $(\mathscr{P}, \mathscr{W}, \Theta)$ be a Poincaré mapping of γ at m_0.*

1. *$\mathscr{P}$ can be shrunk so that it is foliated by embedded submanifolds*
$$\mathscr{P}_E := \mathscr{P} \cap H^{-1}(E)$$
with E in some open interval I containing $E_0 = H(m_0)$.
2. *For every $E \in I$, the submanifolds $\mathscr{P}_E$ and $\mathscr{W}_E := \mathscr{W} \cap \mathscr{P}_E$ are symplectic submanifolds of (M, ω) and the mapping $\Theta_E : \mathscr{W}_E \to \mathscr{P}_E$, induced by restriction of Θ, is a symplectomorphism onto its image.*

The triple $(\mathscr{P}_{E_0}, \mathscr{W}_{E_0}, \Theta_{E_0})$ is a Poincaré mapping for γ with respect to the flow induced on a local energy surface of γ. Accordingly, it is called an isoenergetic Poincaré mapping for γ.

Proof 1. Since X_H is transversal to $\mathscr{P}$, all points of $\mathscr{P}$ are regular with respect to H. Hence, $H_{\restriction \mathscr{P}}$ is a submersion and the Constant Rank Theorem yields a local chart (U, κ) on $\mathscr{P}$ at m_0 with image $B \times I \subset \mathbb{R}^{2n-1}$, where B is some open ball in $\mathbb{R}^{2n-2}$ and I is some open interval containing $H(m_0)$ such that $\kappa(H^{-1}(E) \cap U) = B \times \{E\}$. If we replace $\mathscr{P}$ by $\mathscr{P} \cap U$, we obtain the assertion.

2. Let $E \in I$ and let $j : \mathscr{P}_E \to M$ denote the natural inclusion mapping. Without loss of generality, we may assume that H is a submersion on M. Then, $\mathscr{P}_E$ is a submanifold of some regular energy surface $\Sigma_E \subset H^{-1}(E)$. Since the characteristic distribution of ω_{Σ_E} is spanned by the induced vector field $X_H^{\Sigma_E}$, the kernel of $(\omega_m)_{\restriction \mathrm{T}_m \Sigma_E}$ is spanned by $X_H(m)$. Since $X_H(m) \notin \mathrm{T}_m \mathscr{P}_E$ and since the codimension of $\mathrm{T}_m \mathscr{P}_E$ in $\mathrm{T}_m \Sigma_E$ is 1, we conclude that $(\omega_m)_{\restriction \mathrm{T}_m \mathscr{P}_E} = (j^* \omega)_m$ is symplectic. This shows that $\mathscr{P}_E$ is a symplectic submanifold of M. Next, since Θ is defined by the flow of X_H and H is a constant of motion, Θ restricts to a mapping $\Theta_E : \mathscr{W}_E \to \mathscr{P}_E$. Since Θ is a diffeomorphism onto its image and since $\mathscr{W}_E$ and $\mathscr{P}_E$ are embedded submanifolds of $\mathscr{W}$ and $\mathscr{P}$, respectively, Θ_E is a diffeomorphism onto its image, too. It remains to prove that Θ_E is symplectic, that is,
$$\Theta_E^*\big(j^* \omega\big) = j^* \omega.$$
By Formula (3.7.2), for $m \in \mathscr{W}_E$, we have
$$j \circ \Theta_E(m) = \Phi_{\tau(m)}\big(j(m)\big),$$

where τ denotes the first return time function of $\mathscr{P}$. To calculate the tangent mapping of $j \circ \Theta_E$ at m, let $X \in \mathrm{T}_m \mathscr{W}_E$ and let $t \to \delta(t)$ be a curve representing X. Then,

$$\begin{aligned}(j \circ \Theta_E)'_m(X) &= \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \Phi_{\tau(\delta(t))}\big(j(\delta(t))\big)\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \Phi_{\tau(\delta(t))}\big(j(m)\big) + \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \Phi_{\tau(m)}\big(j(\delta(t))\big)\\ &= \tau'(X) X_H\big(\Phi_{\tau(m)}\big(j(m)\big)\big) + (\Phi_{\tau(m)})' \circ j'(X)\\ &= (\Phi_{\tau(m)})'\big\{\tau'(X) X_H\big(j(m)\big) + j'(X)\big\},\end{aligned}$$

where $\tau'(X) \in \mathbb{R}$. Using this and $(\Phi_t)^*\omega = \omega$, for $X, Y \in \mathrm{T}_m \mathscr{W}_E$ we obtain

$$\begin{aligned}\big(\Theta_E^*\big(j^*\omega\big)\big)_m(X,Y) &= \omega_{j(m)}\big(\tau'(X)X_H\big(j(m)\big) + j'(X), \tau'(Y)X_H\big(j(m)\big) + j'(Y)\big)\\ &= \tau'(X)\omega_{j(m)}\big(X_H\big(j(m)\big), j'(Y)\big)\\ &\quad + \tau'(Y)\omega_{j(m)}\big(j'(X), X_H\big(j(m)\big)\big) + \big(j^*\omega\big)_m(X,Y).\end{aligned}$$

Up to a factor, the first term yields $\mathrm{d}H(j'(Y)) = Y(H \circ j) = 0$, because $H \circ j = E$. In the same way, the second term vanishes. Thus, Θ_E is symplectic, indeed. □

Remark 9.5.2

1. The Uniqueness Theorem 3.7.5 implies a uniqueness theorem for the isoenergetic Poincaré mapping (Exercise 9.5.1).
2. Periodic points of Θ correspond to periodic integral curves of X_H. In particular, $\Theta_{E_0}(m_0) = m_0$, where $H(m_0) = E_0$. We show that the eigenvalues of the tangent mapping

$$(\Theta_{E_0})'_{m_0}: \mathrm{T}_{m_0}\mathscr{P}_{E_0} \to \mathrm{T}_{m_0}\mathscr{P}_{E_0} \tag{9.5.1}$$

coincide with the Floquet multipliers of γ, cf. Remark 8.2.11/1. Since $(\Theta_{E_0})'_{m_0}$ is given by the restriction of Θ'_{m_0} to the invariant subspace $\mathrm{T}_{m_0}\mathscr{P}_{E_0} \subset \mathrm{T}_{m_0}\mathscr{P}$ and since the codimension of this subspace is 1, the spectra of $(\Theta_{E_0})'_{m_0}$ and Θ'_{m_0} differ by a single eigenvalue λ. On the other hand, by Proposition 3.7.6, the spectra of Θ'_{m_0} and $(\Phi_T)'_{m_0}$ differ by the single eigenvalue 1. Hence, the spectra of $(\Theta_{E_0})'_{m_0}$ and $(\Phi_T)'_{m_0}$ differ by the two eigenvalues λ and 1. Since both mappings are symplectomorphisms, by Proposition 7.4.1/4, they have unit determinant. Thus, $\lambda = 1$ and hence the eigenvalues of $(\Theta_{E_0})'_{m_0}$ coincide with the Floquet multipliers, indeed.

Next, we will prove that periodic integral curves of Hamiltonian systems are not isolated. Rather, they appear in 1-parameter families called orbit cylinders.

Definition 9.5.3 (Orbit cylinder) Let (M, ω, H) be a Hamiltonian system. An orbit cylinder of (M, ω, H) is an embedding $\zeta : \mathrm{S}^1 \times I \to M$, where I is some open interval, such that $\zeta(\mathrm{S}^1 \times \{c\})$ is the image of a periodic integral curve of X_H for every $c \in I$. If $H \circ \zeta(\mathrm{S}^1 \times \{c\}) = c$, then ζ is called regular.

For a regular orbit cylinder, the parameter labelling the periodic integral curves coincides with the energy. Therefore, a regular orbit cylinder intersects every energy surface transversally. This implies, in particular, that regular orbit cylinders are symplectic submanifolds: since their tangent spaces are pointwise spanned by X_H and some vector Y satisfying $\mathrm{d}H(Y) \neq 0$, the 2-form $\zeta^*\omega$ vanishes nowhere.

Theorem 9.5.4 (Existence of orbit cylinders) *Let (M, ω, H) be a Hamiltonian system and let γ be a periodic integral curve of X_H. If 1 is not a Floquet multiplier of γ, then there exists an open interval I containing $E_0 = H(\gamma)$ and a regular orbit cylinder $\zeta : \mathrm{S}^1 \times I \to M$ such that $\gamma_{E_0} = \gamma$.*

Proof Choose $m_0 \in \gamma$. According to Proposition 9.5.1, there exists a Poincaré mapping $(\mathscr{P}, \mathscr{W}, \Theta)$ for γ at m_0 such that $\mathscr{P}$ is foliated by embedded submanifolds $\mathscr{P}_E = \mathscr{P} \cap H^{-1}(E)$. We choose a local chart (U, κ) on $\mathscr{P}$ at m_0, adapted to this foliation. Without loss of generality, we may assume that $\kappa(m_0) = (0, E_0)$ and $\kappa(U) = \mathbb{R}^{2n-2} \times I$ with some open interval I containing E_0. Denote $\tilde{\mathscr{W}}_{E_0} := \mathrm{pr}_1 \circ \kappa(\mathscr{W}_{E_0}) \subset \mathbb{R}^{2n-2}$ and let

$$\tilde{\Theta} = \kappa \circ \Theta \circ \kappa^{-1} : \tilde{\mathscr{W}}_{E_0} \times I \to \mathbb{R}^{2n-2} \times I$$

be the local representative of Θ. Consider the mapping

$$\varphi : \tilde{\mathscr{W}}_{E_0} \times I \to \mathbb{R}^{2n-2}, \qquad \varphi(\mathbf{x}, E) := \mathrm{pr}_1\big(\tilde{\Theta}(\mathbf{x}, E) - \mathbf{x}\big).$$

Since $\Theta(m_0) = m_0$, this mapping vanishes at the point $(0, E_0)$. We claim that φ fulfils the assumptions of the Implicit Function Theorem of classical calculus at that point. To see this, let φ_{E_0} denote the induced mapping $\mathbf{x} \mapsto \varphi_{E_0}(\mathbf{x}) := \varphi(\mathbf{x}, E_0)$ and let $\mathbf{Y} \in \mathbb{R}^{2n-2}$ be such that $(\varphi_{E_0})'_0(\mathbf{Y}) = 0$. Then, $\tilde{\Theta}'_{(0,E_0)}(\mathbf{Y}, 0) = (\mathbf{Y}, 0)$, that is, $(\mathbf{Y}, 0)$ is an eigenvector of $\tilde{\Theta}'_{(0,E_0)}$ with eigenvalue 1. Since the eigenvalues of the restriction of $\tilde{\Theta}'_{(0,E_0)}$ to the invariant subspace $\mathbb{R}^{2n-2}$ coincide with the Floquet multipliers of γ and since, by assumption, 1 is not among them, we conclude $\mathbf{Y} = 0$. Therefore, $(\varphi_{E_0})'_0$ is bijective. Now, the Implicit Function Theorem yields that I can be shrunk so that there exists a differentiable mapping $h : I \to \tilde{\mathscr{W}}_{E_0}$ satisfying $\varphi(h(E), E) = 0$ and hence $\tilde{\Theta}(h(E), E) = (h(E), E)$ for all $E \in I$. Let us denote $\psi(E) = \kappa^{-1}(h(E), E)$. By construction, ψ yields a fixed point of Θ, that is, a point on a periodic integral curve, for every energy value $E \in I$. Finally, using ψ and the flow Φ of X_H, we construct the desired orbit cylinder: let τ be the first return time function of $\mathscr{P}$. Consider the mapping

$$\zeta_0 : \mathbb{R} \times I \to M, \qquad \zeta_0(t, E) := \Phi_{t\tau(\psi(E))}\big(\psi(E)\big).$$

Since $\Phi_{\tau(\psi(E))}(\psi(E)) = \Theta(\psi(E)) = \psi(E)$, this mapping is 1-periodic,

$$\zeta_0(t+1, E) = \Phi_{(t+1)\tau(\psi(E))}\big(\psi(E)\big) = \Phi_{t\tau(\psi(E))}\big(\psi(E)\big) = \zeta_0(t, E).$$

Thus, it induces a smooth mapping $\zeta : \mathrm{S}^1 \times I \to M$. Obviously, ζ is an injective immersion. By construction, $\zeta(\mathrm{S}^1 \times \{E\})$ is the image of a periodic integral curve of X_H and $\zeta(\mathrm{S}^1 \times \{E_0\}) = \gamma$. It remains to prove that ζ is an embedding, that is, that it is open onto its image. By shrinking I once again we may assume that ζ extends to a continuous mapping $\bar{\zeta}$ on the closure $\mathrm{S}^1 \times \overline{I}$. Since $\mathrm{S}^1 \times \overline{I}$ is compact, we can apply the argument of Remark 1.6.13/2, which showed that every compact submanifold is embedded, to prove that $\bar{\zeta}$ is open onto its image. Then, so is ζ. □

Remark 9.5.5 The orbit cylinder at γ provided by Theorem 9.5.4 is unique in the following sense. If $\tilde{\zeta} : \mathrm{S}^1 \times \tilde{I} \to M$ is another orbit cylinder containing γ, there exists an open interval $\hat{I} \subset I \cap \tilde{I}$ containing E_0 and a smooth family of diffeomorphisms $\varphi_E : \mathrm{S}^1 \to \mathrm{S}^1$, $E \in \hat{I}$, such that $\tilde{\zeta}(x, E) = \zeta(\varphi_E(x), E)$ for all $x \in \mathrm{S}^1$ and $E \in \hat{I}$.

Now, we combine orbit cylinders with the Poincaré mapping.

Proposition 9.5.6 *Let $\zeta : \mathrm{S}^1 \times I \to M$ be a regular orbit cylinder, let $E_0 \in I$ and let $(\mathscr{P}, \mathscr{W}, \Theta)$ be a Poincaré mapping for γ_{E_0}. Assume that $\mathscr{P}$ is foliated by level sets $\mathscr{P}_E$ of H according to Proposition 9.5.1, with E in some open interval J containing E_0.*

1. *There exists a unique embedding $\iota : I \cap J \to M$ such that $\mathscr{P} \cap \gamma_E = \iota(E)$ for all $E \in I \cap J$.*
2. *For all $E \in I \cap J$, $(\mathscr{P}, \mathscr{W}, \Theta)$ is a Poincaré mapping for γ_E and $(\mathscr{P}_E, \mathscr{W}_E, \Theta_E)$ is an isoenergetic Poincaré mapping for γ_E.*

Proof The embedded submanifolds $\mathscr{P}$ and $\zeta(\mathrm{S}^1 \times I)$ are transversal, because X_H is transversal to the former and tangent to the latter. Thus, by the Transversal Mapping Theorem 1.8.2, the intersection is an embedded submanifold of M. Obviously, the intersection can be parametrized by $E \in I \cap J$. The proof of point 2 is left to the reader (Exercise 9.5.2). □

Since Proposition 9.5.6 implies that the isoenergetic Poincaré mappings Θ_E depend smoothly on E, we conclude

Corollary 9.5.7 *The Floquet multipliers of the periodic integral curves γ_E of a regular orbit cylinder depend smoothly on E.*

In general, a regular orbit cylinder cannot be extended to all $E \in \mathbb{R}$. Rather, it will meet periodic integral curves having 1 among their Floquet multipliers or it may degenerate to an equilibrium. Under certain conditions, the orbit cylinder can be continued nonetheless, though in a nonregular way, see e.g. [1, Thm. 8.2.4]. Usually, however, bifurcation phenomena like the degeneration to an equilibrium already mentioned or a splitting into several new orbit cylinders will occur. A discussion of the types of bifurcations present in Hamiltonian systems with two degrees of freedom can be found in [1, §8.6]. The following classical result of Lyapunov [188] which dates back to 1907 describes bifurcation of an orbit cylinder from a linearly stable equilibrium under certain non-resonance conditions. The periodic solutions can be viewed as nonlinear continuations of the linear normal modes of the system. They have periods close to the periods of the linearized system.

Theorem 9.5.8 (Lyapunov Centre Theorem) *Let (M, ω, H) be a Hamiltonian system, let m_0 be an equilibrium of X_H of energy $E_0 = H(m_0)$ and let $\{\lambda_1, \ldots, \lambda_n\}$ be a basis set*[14] *of characteristic exponents for m_0. Assume that $\lambda_1 = \mathrm{i}\alpha$ with $\alpha > 0$ and that λ_j is not an integer multiple of λ_1 for $j = 2, \ldots, n$. Then, there exists $\delta > 0$ and a regular orbit cylinder with energy interval $(E_0, E_0 + \delta)$ such that in the limit $E \to E_0$, the integral curves γ_E approach m_0 and the period of γ_E approaches $\frac{2\pi}{\alpha}$. Moreover, the set of Floquet multipliers of γ_E, counted with multiplicities, tends to $\{\mathrm{e}^{\frac{2\pi}{\alpha}\lambda_2}, \mathrm{e}^{\frac{2\pi}{\alpha}\overline{\lambda_2}}, \ldots, \mathrm{e}^{\frac{2\pi}{\alpha}\lambda_n}, \mathrm{e}^{\frac{2\pi}{\alpha}\overline{\lambda_n}}\}$.*

Proof Since the statement is of local character, we may assume (M, ω) to be given by $\mathbb{R}^{2n}$ with the standard symplectic structure. Moreover, we can put $m_0 = 0$ and $E_0 = H(m_0) = 0$. We denote $\mathsf{Q} := \frac{1}{2} H''(0)$ and view it as a quadratic form on $\mathbb{R}^{2n}$. Moreover, we denote $A := \mathrm{Hess}_{m_0}(X_H)$. One can show that the Hamiltonian vector field X_{Q} coincides with the linear vector field on $\mathbb{R}^{2n}$ represented by A (Exercise 9.5.4). The basic idea of the proof is to rescale[15] the Hamiltonian as

$$H_\varepsilon(\mathbf{x}) = \frac{1}{\varepsilon^2} H(\varepsilon \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^{2n},$$

with $\varepsilon \in \mathbb{R}$. Then, $H_0 = \mathsf{Q} = \frac{1}{2} H_\varepsilon''(0)$. A brief computation reveals

$$\varepsilon X_{H_\varepsilon}(\mathbf{x}) = X_H(\varepsilon \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^{2n}.$$

Thus, if $\varepsilon \neq 0$ and if γ is a periodic integral curve of X_{H_ε}, then $\varepsilon\gamma$ is a periodic integral curve of X_H with the same period. Hence, we may proceed as follows. First, we determine a periodic integral curve γ_0 of $X_{H_0} = X_{\mathsf{Q}} = A$. Then, we use the Implicit Function Theorem to generate from γ_0 a family of closed curves γ_ε for small values of ε such that γ_ε is an integral curve of X_{H_ε}. Finally, we pass to the curves $\varepsilon\gamma_\varepsilon$ and express ε in terms of the energy E, which yields the desired integral curves γ_E.

To find a periodic integral curve γ_0 of A, we use the fact that A has a pair of eigenvalues given by $\pm\mathrm{i}\alpha$. Accordingly, there exists a symplectic basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_{2n}\}$ such that A has the form

$$A = \left[\begin{array}{cc|c} 0 & \alpha & 0 \\ -\alpha & 0 & 0 \\ \hline 0 & 0 & A_F \end{array}\right]$$

[14] See Remark 8.2.11/2.

[15] This is a standard tool from bifurcation theory also called the blowing up technique, see [1] for historal references.

with respect to the decomposition $\mathbb{R}^{2n} = F_1 \oplus F_2 \oplus F$, where $F_1 = \mathbb{R}\mathbf{e}_1$, $F_2 = \mathbb{R}\mathbf{e}_{n+1}$, and F is spanned by $\mathbf{e}_2, \dots, \mathbf{e}_n, \mathbf{e}_{n+2}, \dots, \mathbf{e}_{2n}$. Then,

$$\Phi_t^{\mathsf{Q}} = \mathrm{e}^{tA} = \left[\begin{array}{cc|c} \cos(\alpha t) & \sin(\alpha t) & 0 \\ -\sin(\alpha t) & \cos(\alpha t) & 0 \\ \hline 0 & 0 & \mathrm{e}^{tA_F} \end{array}\right],$$

hence all points of $F_1 \oplus F_2$ are periodic with period $T_0 := \frac{\alpha}{2\pi}$. Thus, for γ_0 we may choose the integral curve of A through the point $\mathbf{x}_0 = \mathbf{e}_1$. To compute $\mathsf{Q}(\mathbf{x}_0)$, we observe that $A_{\restriction F_1 \oplus F_2}$ is the Hamiltonian vector field of the quadratic form $(q, p) \mapsto \frac{1}{2}\alpha(q^2 + p^2)$ on the symplectic vector space $F_1 \oplus F_2$. Since a quadratic form is uniquely determined by its first partial derivatives, and hence by its Hamiltonian vector field (in case it is defined on a symplectic vector space), we conclude that $\mathsf{Q}_{\restriction F_1 \oplus F_2}(q, p) = \frac{\alpha}{2}(q^2 + p^2)$. Thus, $\mathsf{Q}(\mathbf{x}_0) = \frac{\alpha}{2}$.

Next, to generate the curves γ_ε from γ_0, we use the same idea as in the proof of Theorem 9.5.4, that is, we look for solutions of the equation

$$\Phi_t^{H_\varepsilon}(\mathbf{x}) - \mathbf{x} = 0 \tag{9.5.2}$$

in the variables $\mathbf{x}$ close to $\mathbf{x}_0$ and t close to T_0 for small ε. Now, for $\varepsilon = 0$, the set of solutions $(t, \mathbf{x})$ of this equation is given by the continuum $t = T_0$ and $\mathbf{x} \in F_1 \oplus F_2$. In order to obtain $(T_0, \mathbf{x}_0)$ as an isolated solution for $\varepsilon = 0$, one must impose appropriate constraints on the set of solutions of (9.5.2). First, we require $\mathbf{x} \in F_1 \oplus F$, because this subspace is transversal to γ_0. Second, we require

$$H_\varepsilon(\mathbf{x}) = \frac{\alpha}{2}. \tag{9.5.3}$$

To implement this condition, write $\mathbf{x} = q\mathbf{e}_1 + \mathbf{y} \in F_1 \oplus F$ as a pair $(q, \mathbf{y})$ with $q \in \mathbb{R}$ and $\mathbf{y} \in F$, and consider (9.5.3) as an equation in the indeterminate q. Since for $\varepsilon = 0$ and $\mathbf{y} = 0$ we have the unique solution $q = 1$ and since

$$\left(\frac{\partial H_\varepsilon(q, \mathbf{y})}{\partial q}\right)_{\restriction(q=1, \mathbf{y}=0, \varepsilon=0)} = \alpha \neq 0,$$

the Implicit Function Theorem yields a smooth function $(\mathbf{y}, \varepsilon) \mapsto q(\mathbf{y}, \varepsilon)$, defined for $\mathbf{y}$ in a neighbourhood of 0 in F and for small ε, such that $H_\varepsilon(q(\mathbf{y}, \varepsilon), \mathbf{y}) = \frac{\alpha}{2}$. Finally, since H_ε is invariant under Φ^{H_ε}, it suffices to consider Eq. (9.5.2) on the energy surface defined by (9.5.3). This can be achieved by applying the orthogonal projection $\mathrm{pr} : \mathbb{R}^{2n} \to F_2 \oplus F$, because the latter is bijective in a neighbourhood of the point $(q(0, \varepsilon), 0)$ in that energy surface. Thus, we arrive at the equation

$$\varphi(\mathbf{y}, t; \varepsilon) := \mathrm{pr}\big(\Phi_t^{H_\varepsilon}\big(q(\mathbf{y}, \varepsilon), \mathbf{y}\big) - \big(q(\mathbf{y}, \varepsilon), \mathbf{y}\big)\big) = 0$$

in the variables $\mathbf{y} \in F$ and $t \in \mathbb{R}$ close to $\mathbf{y} = 0$ and $t = T_0$. For the matrix of the partial derivatives of $\varphi = (\varphi^{F_2}, \varphi^F)$ with respect to $\mathbf{y}$ and t at the solution $\mathbf{y} = 0$, $t = T_0$ and $\varepsilon = 0$ we find

$$\frac{\partial(\varphi^{F_2}, \varphi^F)}{\partial(t, \mathbf{y})}_{\restriction \mathbf{y}=0, t=T_0, \varepsilon=0} = \begin{bmatrix} \alpha & 0 \\ 0 & \mathrm{e}^{T_0 A_F} - \mathbb{1}_F \end{bmatrix}. \tag{9.5.4}$$

Since, by assumption, none of the eigenvalues of A_F is an integer multiple of α and since $T_0 = \frac{2\pi}{\alpha}$, the linear mapping $e^{T_0 A_F}$ does not have 1 as an eigenvalue. Therefore, (9.5.4) is bijective. Now, the Implicit Function Theorem yields a mapping $\varepsilon \mapsto (\mathbf{y}(\varepsilon), T(\varepsilon))$ for small ε such that $\varphi(\mathbf{y}(\varepsilon), T(\varepsilon), \varepsilon) = 0$. Thus, we obtain a mapping $\varepsilon \mapsto \mathbf{x}(\varepsilon) := (q(\mathbf{y}(\varepsilon), \varepsilon), \mathbf{y}(\varepsilon)) \in \mathbb{R}^{2n}$ such that $\mathbf{x}(\varepsilon)$ is a periodic point of X_{H_ε} with period $T(\varepsilon)$. Let γ_ε be the corresponding periodic integral curve of X_{H_ε}. As explained above, then $\varepsilon\gamma_\varepsilon$ is a periodic integral curve of X_H with period $T(\varepsilon)$. Obviously, for $\varepsilon \to 0$, these integral curves approach m_0 and the period $T(\varepsilon)$ approaches $T_0 = \frac{2\pi}{\alpha}$. Moreover, $(\Phi^{H_\varepsilon}_{T(\varepsilon)})'_{\mathbf{x}(\varepsilon)}$ approaches $(\Phi^{\mathrm{Q}}_{T_0})'_0 = e^{T_0 A}$. Taking the affine subspace $\varepsilon\mathbf{x}(\varepsilon) + F$ as a Poincaré section for the integral curve $\varepsilon\gamma_\varepsilon$, it is easy to see that for $\varepsilon \to 0$, the Floquet multipliers of this integral curve tend to the eigenvalues of $e^{T_0 A_F}$, that is, to

$$e^{T_0\lambda_2}, e^{T_0\overline{\lambda_2}}, \dots, e^{T_0\lambda_n}, e^{T_0\overline{\lambda_n}}.$$

It remains to express ε for $\varepsilon > 0$ by the energy E with respect to H: since

$$E = H(\varepsilon\gamma_\varepsilon) = \varepsilon^2 H_\varepsilon(\gamma_\varepsilon) = \frac{1}{2}\varepsilon^2\alpha,$$

we obtain $\varepsilon(E) = \sqrt{\frac{2E}{\alpha}}$ and thus $\gamma_E := \varepsilon(E)\gamma_{\varepsilon(E)}$ is a periodic integral curve of X_H of energy E with period $T(\varepsilon(E))$. Now, the construction of the orbit cylinder from the curves γ_E is completely analogous to that in the proof of Theorem 9.5.4. □

Remark 9.5.9

1. From the above proof it is clear that the point m_0 complements the orbit cylinder to a two-dimensional embedded topological submanifold. One can show that if H is of class C^{r+2}, the embedding is of class C^r and if H is analytic, the embedding is analytic [273]. Moreover, this submanifold is symplectic, because orbit cylinders are symplectic submanifolds and the tangent space at m_0 coincides with the eigenspace of $\mathrm{Hess}_{m_0}(X_H)$ corresponding to the pair of eigenvalues $\pm i\alpha$.
2. If the non-resonance condition in the Lyapunov Centre Theorem is violated, no periodic solutions need exist, see Example 9.5.10 below. However, under certain additional assumptions, remarkable generalizations have been found. If one additionally assumes that the Hessian of H is positive or negative definite, that is, $H''(0) > 0$ or $H''(0) < 0$, then at least n geometrically distinct periodic solutions exist. This result belongs to Weinstein [304] and Moser [224]. Moreover, Faddell and Rabinowitz showed the existence of periodic solutions under the weaker assumption $\mathrm{sign}(H''(0)) \neq 0$ [88].
3. One can replace the equilibrium $m_0 \in M$ by a non-degenerate Morse-Bott minimum along a closed symplectic submanifold $N \subset M$ and one can ask whether every level set of H near N carries at least one periodic integral curve. The so-called generalized Weinstein-Moser conjecture states that this is true. For some particular cases, this conjecture has been proven, but up to our knowledge, in the general case the problem is still open. We refer to [106] for a thorough discussion.

Example 9.5.10 Following Moser [226], we consider the Hamiltonian

$$H(\mathbf{z}, \bar{\mathbf{z}}) = \frac{1}{2}\big(|z_2|^2 - |z_1|^2\big) + \big(|z_2|^2 + |z_1|^2\big)\,\mathrm{Re}(z_1 z_2) \tag{9.5.5}$$

on $\mathbb{R}^4$, where $z_i = q_i + \mathrm{i}p_i$. One can show the following (Exercise 9.5.5). The point $z_1 = z_2 = 0$ is an equilibrium. There are no periodic solutions. The eigenvalues of the linearized system are $\pm\mathrm{i}$ and the non-resonance condition is not satisfied. Moreover, the Hessian of H has signature 0.

By the above theorems, for a given periodic integral curve γ, generically there is a continuum of periodic integral curves nearby. However, each of these curves lies on a different energy surface. It is interesting to ask whether there exist periodic integral curves near γ lying on the same energy surface. Thus, our next task will be the study of the behaviour of our system on a given energy surface near a critical integral curve. This will be one of the topics of the subsequent section.

Exercises

9.5.1 Prove the statements of Remark 9.5.2.

9.5.2 Prove point 2 of Proposition 9.5.6.

9.5.3 Prove Remark 9.5.5.

9.5.4 Complete the proof of Theorem 9.5.8 by showing that the Hamiltonian vector field X_{Q} coincides with the linear vector field $\mathrm{Hess}_0(X_H)$.
Hint. Show that in the standard Darboux coordinates q^i, p_i on $\mathbb{R}^{2n}$, both linear vector fields are represented by the matrix

$$\begin{bmatrix} \frac{\partial^2 H}{\partial q^i \partial p_j} & \frac{\partial^2 H}{\partial p_i \partial p_j} \\ -\frac{\partial^2 H}{\partial q^i \partial q^j} & -\frac{\partial^2 H}{\partial p_i \partial q^j} \end{bmatrix}.$$

9.5.5 Prove the statements of Example 9.5.10.

9.6 Birkhoff Normal Form and Invariant Tori

The basic idea for our further analysis is to bring the Hamiltonian system to a normal form. This idea goes back to Poincaré, see [240]. It has been deeply analyzed and further developed by Birkhoff, see [47]. First, we will prove the normal form theorem for a symplectomorphism in a neighbourhood of a fixed point. Then, we will state the analogous theorem for a Hamiltonian near an equilibrium. In both cases, the unperturbed normal forms give rise to a foliation of the phase space into invariant tori, whose physical meaning will be discussed in some detail. Next we comment on KAM theory, which guarantees the persistence of many of these tori under the perturbation which is caused by considering the full mapping or the full Hamiltonian. Moreover, using the normal form, we will prove the Birkhoff-Lewis Theorem, which states the existence of infinitely many periodic points in the neighbourhood of a given periodic integral curve lying on the same energy surface.

Definition 9.6.1 (r-Elementarity)

1. A fixed point m_0 of a symplectomorphism is called r-elementary if some (and hence any) basis set[16] $\{\lambda_1, \dots, \lambda_n\}$ of characteristic multipliers of m_0 has the property that
$$\lambda_1^{k_1} \cdots \lambda_n^{k_n} \neq 1$$
for all integers $k_1, \dots, k_n$ with $0 < \sum_{i=1}^n |k_i| \le r$. Analogously, a periodic integral curve γ of a Hamiltonian vector field is called r-elementary if some basis set of Floquet multipliers has this property.
2. An equilibrium of a Hamiltonian vector field is called r-elementary if some basis set $(\mu_1, \dots, \mu_n)$ of characteristic exponents of m_0 satisfies
$$k_1\mu_1 + \cdots + k_n\mu_n \neq 0$$
for all integers $k_1, \dots, k_n$ with $0 < \sum_{i=1}^n |k_i| \le r$.

Theorem 9.6.2 (Birkhoff normal form) *Let (M, ω) be a symplectic manifold of dimension $2n$ and let Ψ be a local symplectomorphism*[17] *of open subsets of M. Let m_0 be an elliptic and 4-elementary fixed point of Ψ. Then, in a neighbourhood of m_0 there exist Darboux coordinates such that Ψ is given by*
$$\bar{q}_k = \cos(\alpha_k)q_k - \sin(\alpha_k)p_k + f_k, \qquad \bar{p}_k = \sin(\alpha_k)q_k + \cos(\alpha_k)p_k + g_k \quad (9.6.1)$$
with
$$\alpha_k = a_k + \sum\nolimits_{l=1}^n \beta_{kl} I_l, \qquad I_l = \frac{1}{2}\big(q_l^2 + p_l^2\big). \quad (9.6.2)$$
Here, $\{e^{ia_1}, \dots, e^{ia_n}\}$ is a basis set for the characteristic exponents of m_0, β_{kl} is a real $(n \times n)$-matrix and f_k, g_k are C^∞-functions of the variables q_i and p_i such that all their partial derivatives vanish at the origin up to order 3.

Denoting $\mathbf{x} = (\mathbf{q}, \mathbf{p})^{\mathrm{T}}$, $X = \mathrm{diag}(a_1, \dots, a_n, a_1, \dots, a_n)$ and
$$B = \mathrm{diag}(b_1, \dots, b_n, b_1, \dots, b_n), \quad b_k = \sum_{l=1}^n \beta_{kl} I_l,$$
up to fourth order, Formula (9.6.1) can be rewritten as
$$\mathbf{x} \mapsto e^{-J(X+B)}\mathbf{x} + \cdots. \quad (9.6.3)$$

The following proof is along the lines of Theorem 7 in [225]. The reader should recall the notion of a generating function of a symplectomorphism, see Sect. 8.8.

Proof Since the statement of the theorem is local, we may assume $M = \mathbb{R}^{2n}$, endowed with the canonical symplectic form given by J. Moreover, we can assume

[16] Cf. Remark 8.2.11.

[17] It is enough to assume the Ψ is of class C^3 [225].

that the fixed point coincides with the origin of $\mathbb{R}^{2n}$. Since m_0 is elliptic, the eigenvalues λ_k of $A := \Psi'(0)$ lie on the unit circle. By 4-elementarity, they are all distinct. Hence, Proposition 7.4.5 yields a symplectic basis such that

$$A = e^{-JX}. \tag{9.6.4}$$

Let us decompose $\Psi = A \circ \phi$. Then, $\phi'(0) = \mathrm{id}$. In what follows, for a function F we denote $F_{\mathbf{x}} := (F'(\mathbf{x}))^{\mathrm{T}}$.

Lemma 9.6.3 *There exists a homogeneous polynomial $\mathscr{P}$ of order $\nu + 1 \geq 3$ such that the Taylor expansion to order ν of ϕ at the origin is given by*

$$\phi(\mathbf{x}) = \mathbf{x} + J\mathscr{P}_{\mathbf{x}}(\mathbf{x}) + \cdots. \tag{9.6.5}$$

Proof of Lemma 9.6.3 Let us denote $\phi(\mathbf{x}) = (\bar{\mathbf{q}}, \bar{\mathbf{p}})^T$. Since $\phi'(0) = \mathrm{id}$, ϕ has a generating function of the second kind of the form $(\mathbf{q}, \bar{\mathbf{p}}) \mapsto \mathbf{q}^{\mathrm{T}}\bar{\mathbf{p}} + S(\mathbf{q}, \bar{\mathbf{p}})$, where S has vanishing partial derivatives at the origin up to order 2. Then,

$$\bar{q}_i = q_i + \frac{\partial}{\partial \bar{p}_i} S(\mathbf{q}, \bar{\mathbf{p}}), \qquad \bar{p}_i = p_i - \frac{\partial}{\partial q_i} S(\mathbf{q}, \bar{\mathbf{p}}). \tag{9.6.6}$$

Expanding $S(\mathbf{q}, \bar{\mathbf{p}})$ and ϕ into a Taylor series at the origin and comparing coefficients in (9.6.6), one finds that $\mathscr{P}$ is given by the first non-vanishing Taylor term of S.

Lemma 9.6.4 *Let $F(\mathbf{q}, \bar{\mathbf{p}})$ be a homogeneous polynomial of order $\nu + 1$ (the same order as $\mathscr{P}$) and let τ be the canonical transformation defined by the generating function of the second kind $(\mathbf{q}, \bar{\mathbf{p}}) \mapsto \mathbf{q}^{\mathrm{T}}\bar{\mathbf{p}} + F(\mathbf{q}, \bar{\mathbf{p}})$. Then, the canonical transformation*

$$\tilde{\Psi} = \tau^{-1} \circ \Psi \circ \tau \tag{9.6.7}$$

can be represented by $\tilde{\Psi} = A \circ \tilde{\phi}$, where the Taylor expansion to order ν of $\tilde{\phi}$ at the origin is given by (9.6.5) with $\mathscr{P}$ replaced by

$$\tilde{\mathscr{P}}(\mathbf{x}) = \mathscr{P}(\mathbf{x}) - F(A\mathbf{x}) + F(\mathbf{x}). \tag{9.6.8}$$

Proof of Lemma 9.6.4 We rewrite (9.6.7) as $\tau \circ \tilde{\Psi} = \Psi \circ \tau$ and compare the terms of order ν in the Taylor expansion of both sides. The right hand side yields

$$\begin{aligned}\Psi \circ \tau(\mathbf{x}) &= A \circ \phi \circ \tau(\mathbf{x}) \\ &= A \circ \phi\big(\mathbf{x} + JF_{\mathbf{x}}(\mathbf{x})\big) \\ &= A\big(\mathbf{x} + JF_{\mathbf{x}}(\mathbf{x}) + J\mathscr{P}_{\mathbf{x}}\big(\mathbf{x} + JF_{\mathbf{x}}(\mathbf{x})\big)\big)\end{aligned}$$

and the left hand side yields

$$\begin{aligned}\tau \circ \tilde{\Psi}(\mathbf{x}) &= \tau \circ A \circ \tilde{\phi}(\mathbf{x}) \\ &= \tau(A\mathbf{x} + AJ\tilde{\mathscr{P}}_{\mathbf{x}}(\mathbf{x})) \\ &= A\mathbf{x} + AJ\tilde{\mathscr{P}}_{\mathbf{x}}(\mathbf{x}) + JF_{\mathbf{x}}\big(A\mathbf{x} + AJ\tilde{\mathscr{P}}_{\mathbf{x}}(\mathbf{x})\big).\end{aligned}$$

Thus, in order ν we obtain the relation

$$AJ\big(F_{\mathbf{x}}(\mathbf{x})+\mathscr{P}_{\mathbf{x}}(\mathbf{x})\big)=AJ\tilde{\mathscr{P}}_{\mathbf{x}}(\mathbf{x})+JF_{\mathbf{x}}(A\mathbf{x}).$$

Multiplying from the left by $J^{-1}A^{-1}$ and using that A is symplectic, $A^TJA=J$, we obtain

$$\big(\tilde{\mathscr{P}}(\mathbf{x})-\mathscr{P}(\mathbf{x})+F(A\mathbf{x})-F(\mathbf{x})\big)_{\mathbf{x}}=0.$$

Since all terms in the bracket are homogeneous of degree $\nu+1$, the assertion follows.

Proof of Theorem 9.6.2 (continued) Now, we will analyze (9.6.8) in order 3 and 4. For that purpose, it is convenient to pass to complex variables $z_k=q_k+ip_k$. In these variables, the homogeneous polynomials F, $\mathscr{P}$ and $\tilde{\mathscr{P}}$ take the form

$$F=\sum F_{\rho\sigma}z^\rho\bar{z}^\sigma,\qquad \mathscr{P}=\sum \mathscr{P}_{\rho\sigma}z^\rho\bar{z}^\sigma,\qquad \tilde{\mathscr{P}}=\sum \tilde{\mathscr{P}}_{\rho\sigma}z^\rho\bar{z}^\sigma,$$

where

$$z^\rho=\prod_{k=1}^n z_k^{\rho_k},\qquad \sum_k(\rho_k+\sigma_k)=\nu+1.$$

Moreover, by formula (9.6.4), in these variables the mapping $\mathbf{x}\mapsto A\mathbf{x}$ is given by

$$(z_k,\bar{z}_k)\mapsto\big(\lambda_k z_k,\lambda_k^{-1}\bar{z}_k\big).$$

Thus, Eq. (9.6.8) takes the form

$$\big(\lambda^{\rho-\sigma}-1\big)F_{\rho\sigma}=\mathscr{P}_{\rho\sigma}-\tilde{\mathscr{P}}_{\rho\sigma},\tag{9.6.9}$$

where $\lambda^{\rho-\sigma}=\prod_{k=1}^n\lambda_k^{\rho_k-\sigma_k}$. The factor $\lambda^{\rho-\sigma}-1$ vanishes iff

$$\sum_{k=1}^n(\rho_k-\sigma_k)a_k=2\pi m\tag{9.6.10}$$

for some integer m. The order we have to start with is $\nu=2$. Here, $\sum(\rho_k+\sigma_k)=3$. This implies, in particular, that $\sum\rho_k\neq\sum\sigma_k$. On the other hand, since the fixed point is 4-elementary, the only solution of (9.6.10) is $\rho=\sigma$. Thus, $\lambda^{\rho-\sigma}-1$ does not vanish and we can choose $F_{\rho\sigma}$ such that $\tilde{\mathscr{P}}_{\rho\sigma}$ vanishes. By performing the canonical transformation τ defined by F we obtain a new canonical transformation for which $\nu=3$. Here, a similar analysis applies, but the solution $\rho=\sigma$ is allowed now, so that $\tilde{\mathscr{P}}_{\rho\rho}=\mathscr{P}_{\rho\rho}$ and $F_{\rho\rho}$ remains undetermined. For $\rho\neq\sigma$, we choose $F_{\rho\sigma}$ such that $\tilde{\mathscr{P}}_{\rho\sigma}$ vanishes and we fix F by setting $F_{\rho\rho}=0$. Thus, $\tilde{\mathscr{P}}$ is a homogeneous polynomial of order 2 in $z_k\bar{z}_k=2I_k$:

$$\tilde{\mathscr{P}}(\mathbf{q},\mathbf{p})=-\sum_{k,l}\beta_{kl}I_kI_l,\quad I_k=\frac{1}{2}z_k\bar{z}_k=\frac{1}{2}\big(q_k^2+p_k^2\big).\tag{9.6.11}$$

To summarize, by a transformation of the form (9.6.7) we have brought Ψ to the form $\tilde{\Psi}=A\circ\tilde{\phi}$, with $\tilde{\phi}$ given by

$$\mathbf{x}\mapsto\mathbf{x}+J\mathscr{P}_{\mathbf{x}}+\cdots=\mathbf{x}-JB\mathbf{x}+\cdots.$$

Using (9.6.4) we conclude that $\tilde{\Psi}$ is given by

$$\mathbf{x} \mapsto e^{-JX}(\mathbf{x} - JB\mathbf{x} + \cdots) = e^{-J(X+B)}\mathbf{x} + \cdots,$$

because X, B and J commute. This is the normal form up to order 3 stated in the theorem. □

Remark 9.6.5

1. Formula (9.6.1) is called the Birkhoff normal form of the symplectomorphism Ψ in the neighbourhood of the fixed point m_0. The coefficients a_k and β_{kl} are uniquely determined by Ψ. Therefore, they are called the Birkhoff invariants of Ψ.
2. It is obvious from the above proof that the procedure described can be iterated to arbitrary order.
3. Theorem 9.6.2 applies in particular to the isoenergetic Poincaré mapping of an r-elementary periodic integral curve of a Hamiltonian system (but of course not to the period mapping). Along the corresponding orbit cylinder, the Birkhoff invariants depend smoothly on the energy E. The argument is the same as for the Floquet multipliers, cf. Corollary 9.5.7.

Definition 9.6.6 Let Ψ be a local symplectomorphism of a symplectic manifold (M, ω). Let $m_0 \in M$ be an elliptic 4-elementary fixed point of Ψ given in normal form (9.6.1). Then, m_0 is called non-degenerate if $\det \beta \neq 0$.

Now, for a moment, let us ignore the higher order terms f_k and g_k in (9.6.1) and let us consider the symplectomorphism

$$\hat{\Psi}(\mathbf{x}) = e^{-J(X+B)}\mathbf{x},$$

cf. (9.6.3). Moreover, let us assume that the fixed point m_0 is non-degenerate. It is clear that the functions $(I_1, I_2, \ldots, I_n)$ define a foliation of the open submanifold

$$\mathbb{R}^{2n} \setminus \left\{(\mathbf{q}, \mathbf{p}) \in \mathbb{R}^{2n} : I_k = 0 \text{ for some } k\right\} \cong \mathrm{T}^n \times (\mathbb{R}_+)^n$$

into n-dimensional tori. By formula (9.6.2), the assumption that m_0 be non-degenerate can be rewritten as

$$\det\left(\frac{\partial \alpha_k}{\partial I_l}\right) \neq 0,$$

that is, the tori can be also labelled by the set $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of angles of rotation. By this labelling, these tori fall into two classes: a torus is called non-resonant, if the angles α_k are rationally independent, that is, if for $\mathbf{k} \in \mathbb{Z}^n$ we have

$$\sum_{i=1}^{n} k_i \alpha_i \notin 2\pi\mathbb{Z}. \tag{9.6.12}$$

Otherwise, the torus is called resonant. Both the resonant and the non-resonant tori constitute dense subsets. Moreover, the set of non-resonant tori has full measure in phase space. Note that the set of resonant tori further decomposes into subsets consisting of tori, for which there are $n-1, n-2, \ldots, 1$ rationally independent angles α_k. All these subsets are also dense and for the case of one independent angle every point on the corresponding torus is necessarily periodic. From this geometric picture the following important conclusions can be drawn:

(a) Every neighbourhood of the fixed point contains a periodic orbit of $\hat{\Psi}$. Indeed, since the set of resonant tori with one independent angle is dense, every neighbourhood of the fixed point contains such a torus. On this torus, every point m is periodic of some period N. Then, $\{\hat{\Psi}^k(m),\ k = 1, \ldots, N-1\}$ is the corresponding periodic orbit.
(b) Every neighbourhood of m_0 contains uncountably many invariant tori with the property that every orbit of $\hat{\Psi}$ is dense in its torus. Indeed, the set of non-resonant tori contained in a given neighbourhood is dense in this neighbourhood. For each torus of this type, every orbit of $\hat{\Psi}$ is dense. Density statements of this type can be shown using arguments from ergodic theory, see e.g. §51 of [18].

Next, let us consider a Hamiltonian system (M, ω, H) and let us assume that m_0 is an elliptic equilibrium of X_H, fulfilling a similar non-resonance condition. Then, by a completely analogous procedure as in the proof of Theorem 9.6.2, in a neighbourhood of m_0 one can bring the Hamiltonian to the following normal form.

Theorem 9.6.7 *Let (M, ω, H) be a Hamiltonian system and let $m_0 \in M$ be a 4-elementary elliptic equilibrium of X_H. There exist Darboux coordinates such that*

$$H(\mathbf{q}, \mathbf{p}) = \sum_l a_l I_l + \frac{1}{2} \sum_{kl} \beta_{kl} I_k I_l + \cdots, \tag{9.6.13}$$

with I_l given by (9.6.11) and β_{kl} being a uniquely determined $n \times n$ matrix.

Proof See Theorem 1.8.11 in [139]. □

Let us discuss the dynamics near the equilibrium by ignoring the higher order terms in the normal form. The resulting Hamiltonian will be denoted by

$$\hat{H} = \sum_l a_l I_l + \frac{1}{2} \sum_{k.l} \beta_{kl} I_k I_l .$$

We complement the I_k to symplectic polar coordinates by choosing angle coordinates ϑ_k on the tori:

$$p_k = \sqrt{2I_k} \sin \vartheta_k, \qquad q_k = \sqrt{2I_k} \cos \vartheta_k .$$

Then,

$$\sum_k dp_k \wedge dq_k = \sum_k dI_k \wedge d\vartheta_k \tag{9.6.14}$$

and the Hamilton equations take the form

$$\dot{\vartheta}_k = \frac{\partial \hat{H}}{\partial I_k}, \qquad \dot{I}_k = 0. \tag{9.6.15}$$

Thus, the flow of $X_{\hat{H}}$ leaves the tori $\mathrm{T}^n \times \{I\}$ invariant and on each torus we have a quasiperiodic motion

$$p_k = \sqrt{2I_k}\sin(\vartheta_k + \omega_k t), \qquad q_k = \sqrt{2I_k}\cos(\vartheta_k + \omega_k t), \tag{9.6.16}$$

with characteristic frequencies

$$\omega_k = \frac{\partial \hat{H}}{\partial I_k} = a_k + \sum_l \beta_{kl} I_l. \tag{9.6.17}$$

As above, let us assume that the system is non-degenerate, that is, $\det \beta \neq 0$. Then, since

$$\beta_{kl} = \frac{\partial \omega_k}{\partial I_l},$$

the invariant tori can be labelled by the frequencies and we have a completely analogous picture of resonant and non-resonant tori as described after the proof of Theorem 9.6.2. If a torus Σ_I is non-resonant, that is, if the frequencies ω on this torus are rationally independent,

$$\sum_{i=1}^{n} \omega^i k_i \neq 0 \quad \text{for all } k_j \in \mathbb{Z}, \tag{9.6.18}$$

each integral curve is dense in Σ_I. This can be proved using elementary arguments from ergodic theory, see §51 of [18]. Since the resonant tori are also dense, for any neighbourhood of the equilibrium there exists a resonant torus on which we have a periodic solution. Since, by (9.6.14), the variables I_k Poisson-commute, they constitute a system of n commuting constants of motion, that is, the system defined by $\hat{H}$ is integrable. This notion was already introduced in Sect. 9.1 and will be analyzed in detail in Chap. 11. In this context, the coordinates (I_k, ϑ_k) are called action and angle variables. To summarize, the Birkhoff normal form procedure yields a method for transforming a Hamiltonian system near an equilibrium into an integrable system up to a higher order perturbation.

Remark 9.6.8 (KAM Theory) One can pass to the full Hamiltonian by considering the higher order terms in (9.6.13) as a small perturbation of $\hat{H}$. Applying the celebrated KAM[18] theory, one can show that many[19] among the non-resonant tori survive[20] the perturbation procedure, that is, for these tori the quasiperiodic motion

[18]KAM refers to original work of Kolmogorov [168, 169], Arnold [11, 12] and Moser [217, 219, 220].

[19]Those which fulfil an additional, stronger non-resonance condition.

[20]They are only getting slightly deformed. More precisely, in each step of the perturbation procedure the variables (I_K, ϑ_k) are getting modified.

persists. For a nice and transparent qualitative discussion we refer the reader to the book of Arnold [18], paragraph 4 of Appendix 8. There, one can find 6 versions of the theorem on invariant tori, including the two cases (behaviour near an equilibrium of an autonomous Hamiltonian and behaviour near a fixed point of a symplectomorphism) discussed above. We also recommend the classical textbooks [273], §36, and [26], Chap. 5. For a more recent overview containing a lot of further references we refer to [63]. For an abstract discussion, which in particular goes beyond the Hamiltonian setting, we refer to [270]. We also recommend the pedagogical presentation by Pöschel [246], which contains a complete proof of the KAM Theorem in its classical form. In this version, small time-independent perturbations of an autonomous integrable system are considered. We will discuss this case in some detail in Chap. 11. In all these variants, normal forms of either a symplectomorphism or a Hamiltonian function play a basic role. There is a beautiful paper by Douady [75], where the relation between these two aspects is discussed.

Remark 9.6.9 If the characteristic multipliers are r-elementary, by iterating the procedure of Theorem 9.6.7, one obtains a normal form of order s in the variables I_k, with s denoting the integer part of $\frac{r}{2}$. On the other hand, if the characteristic exponents satisfy certain resonance conditions, the normal form becomes more complicated. Roughly speaking, it is then impossible to express all of the contributions to the Hamiltonian up to a given order in terms of variables I_k only. There is a huge literature on this subject, see [18], Sects. 4, 5 and 6 of Appendix 7 and [26], Sect. 3 of Chap. 7, where the reader can also find a lot of further references. In the case of 2 degrees of freedom, the phase portraits in the neighbourhood of resonant equilibria can be analyzed in detail, see Sect. 3.2 of Chap. 7 in [26]. An example can be found in Exercise 9.6.1.

Another important consequence of the Birkhoff Normal Form Theorem is the following.

Theorem 9.6.10 (Birkhoff-Lewis) *Let (M, ω) be a symplectic manifold and let Ψ be a local symplectomorphism of M. Let m_0 be an elliptic,*[21] *non-degenerate and* 4*-elementary fixed point of Ψ. Then, every neighbourhood of m_0 contains infinitely many periodic orbits of Ψ. The number of periodic orbits with period smaller than a given constant is finite.*

We give a proof of this theorem, except for an analytic estimate (Lemma 9.6.11) which can be found in [225] or in the appendix of Sect. 3.3. in [165] (written by Moser). The original work of Birkhoff and Lewis is contained in [48]. We note that this theorem can be viewed as a generalization of the classical Poincaré-Birkhoff Theorem (Theorem 9.9.3) which will be discussed in connection with the Arnold conjecture in Sect. 9.9.

[21] This assumption can be weakened. It is enough to assume that at least one of the Floquet multipliers lies on the unit circle and is different from 1 [225, Thm. 1].

Proof By Theorem 9.6.2, we may assume that m_0 is the origin of $M = \mathbb{R}^{2n}$, the latter endowed with the canonical symplectic structure, and that Ψ is given by (9.6.1). For given $\varepsilon > 0$, we define rescaled symplectic polar coordinates I_k, ϑ_k by

$$q_k = \sqrt{2\varepsilon I_k}\cos\vartheta_k, \qquad p_k = \sqrt{2\varepsilon I_k}\sin\vartheta_k.$$

In the sequel, we use the notation β for the matrix with entries β_{kl} and

$$\mathbf{I} = (I_1, \dots, I_n), \qquad \boldsymbol{\vartheta} = (\vartheta_1, \dots, \vartheta_n),$$
$$\mathbf{a} = (a_1, \dots, a_n), \qquad \boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n).$$

Choose $0 < r < \frac{1}{2n}$, and let

$$B_r := \left\{ \mathbf{I} \in \mathbb{R}^n : \sum_{k=1}^{n} \left(2I_k - \frac{1}{2n}\right)^2 < r^2 \right\}$$

be the open ball with radius $\frac{r}{2}$ centred at the point $\mathbf{e}^{(0)} := (\frac{1}{2n}, \dots, \frac{1}{2n})$. Due to

$$\frac{1}{2n} - r < 2I_k < \frac{1}{2n} + r, \qquad 0 < |\mathbf{q}|^2 + |\mathbf{p}|^2 = 2\varepsilon \sum_k I_k < \varepsilon\left(\frac{1}{2} + nr\right) < \varepsilon,$$

$\mathrm{T}^n \times B_r$ is an open subset of $\mathrm{T}^n \times \mathbb{R}^n_+$ contained in the open ε-ball in $\mathbb{R}^{2n}$ centred at the origin. To begin with, one has to prove the following estimate for the iterates Ψ^j of Ψ. Let $\hat{\Psi}$ denote the mapping (9.6.1) with I_k replaced by εI_k and the higher order terms f_k and g_k ignored.

Lemma 9.6.11 *For every $0 < r' < r$ and every $c_1 > 0$, there exists $\varepsilon_0 > 0$ such that*

$$\Psi^j(\mathrm{T} \times B_{r'}) \subset \mathrm{T} \times B_r$$

for all $1 \le j \le \frac{c_1}{\varepsilon}$ and all $0 < \varepsilon < \varepsilon_0$. Moreover,

$$\Psi^j - \hat{\Psi}^j = \varepsilon^{-1} o_1(\mathbf{I}, \boldsymbol{\vartheta}, \varepsilon),$$

where o_1 is some smooth function on $\mathrm{T}^n \times B_{r'}$ such that $\varepsilon^{-1} o_1$ and its first partial derivatives with respect to I_k and ϑ_k tend uniformly to zero with ε tending to zero.

For the proof, see [225, Lemma 1].

Next, let $0 < r'' < r' < r < \frac{1}{2n}$. We show that $\hat{\Psi}$ has a periodic point in $\mathrm{T}^n \times B_{r''}$. That is, there exist $\tilde{\mathbf{I}} \in B_{r''}$ and some positive integer N such that

$$(2\pi)^{-1} N \alpha(\varepsilon\tilde{\mathbf{I}}) \in \mathbb{Z}^n, \tag{9.6.19}$$

where we view $\boldsymbol{\alpha} = \alpha(\mathbf{I})$, cf. Formula (9.6.2). Then, every point of the torus $\mathrm{T}^n \times \{\tilde{\mathbf{I}}\}$ is a fixed point of $\hat{\Psi}^N$ and hence a periodic point of $\hat{\Psi}$. To prove this, consider the open ball $B_{r''}$ centred at $\mathbf{e}^{(0)}$. Using $\|\beta x\| \ge \|\beta^{-1}\|^{-1}\|x\|$ for all $\mathbf{x} \in \mathbb{R}^n$, where $\|\beta^{-1}\|$ denotes the operator norm, it is easy to see that the image of this ball under the mapping $\mathbf{I} \mapsto \alpha(\varepsilon\mathbf{I})$ contains a ball of radius $\varepsilon \parallel \beta^{-1} \parallel^{-1} r''$, centred at $\mathbf{a} +$

$\varepsilon\beta\mathbf{e}^{(0)}$. Thus, the image of $B_{r''}$ under the mapping $\mathbf{I} \mapsto (2\pi)^{-1} N\alpha(\varepsilon\mathbf{I})$ contains a ball of radius

$$R = (2\pi)^{-1} \left\|\beta^{-1}\right\|^{-1} N\varepsilon r'',$$

and hence a cube of side length $2n^{-\frac{1}{2}}R$. Thus, given r'' and ε, we can choose N so that

$$\varepsilon N r'' > \pi n^{\frac{1}{2}} \left\|\beta^{-1}\right\|.$$

Then, the side length satisfies $2n^{-\frac{1}{2}}R > 1$, so that the cube must contain a point of the lattice $\mathbb{Z}^n$. This proves the existence of periodic points of $\hat{\Psi}$ in $\mathrm{T}^n \times B_{r''}$.

Finally, we prove that there exists a fixed point for some iterate of the full mapping Ψ. With the shorthand notation $(\boldsymbol{\vartheta}^{(N)}, \mathbf{I}^{(N)}) = \Psi^N(\boldsymbol{\vartheta}, \mathbf{I})$, we must show that the system of equations

$$\mathbf{I}^{(N)} = \mathbf{I}, \qquad \boldsymbol{\vartheta}^{(N)} = \boldsymbol{\vartheta} + 2\pi\mathbf{h}, \quad \mathbf{h} \in \mathbb{Z}^n, \tag{9.6.20}$$

has a solution. For that purpose, first we show that there exists a torus which by an iterate Ψ^N is mapped radially, that is, in such a way that the second equation in (9.6.20) holds. By Lemma 9.6.11,

$$\boldsymbol{\vartheta}^{(N)} - \boldsymbol{\vartheta} = N\alpha(\varepsilon\mathbf{I}) + \varepsilon^{-1}o_1(\mathbf{I}, \boldsymbol{\vartheta}, \varepsilon).$$

On the other hand, as we have just shown, there exists $\tilde{\mathbf{I}} \in B_{r''}$, a positive integer N and $\mathbf{h} \in \mathbb{Z}^n$ such that

$$N\alpha(\varepsilon\tilde{\mathbf{I}}) = 2\pi\mathbf{h}.$$

Using the estimate of Lemma 9.6.11 once again, as well as a version of the Implicit Function Theorem[22] which provides estimates for the domain of the solution, one obtains that there exists a solution $\mathbf{I}_0 = \tilde{\mathbf{I}} + \varepsilon^{-1}o_1(\varepsilon) \in B_{r'}$ solving the equation

$$\boldsymbol{\vartheta}^{(N)} - \boldsymbol{\vartheta} = N\alpha(\varepsilon\mathbf{I}) + \varepsilon^{-1}o_1(\mathbf{I}, \boldsymbol{\vartheta}, \varepsilon) = 2\pi\mathbf{h}.$$

Let us denote the torus corresponding to $\mathbf{I}_0$ by T^n_0. It remains to show that the first equation in (9.6.20) has a solution on T^n_0. By (9.6.14), we have

$$\mathrm{d}\left(\sum_k \varepsilon I_k \,\mathrm{d}\vartheta_k\right) = \omega.$$

Hence, for any two-dimensional submanifold Σ of M with boundary $\gamma \subset \mathrm{T}^n_0$, Stokes' Theorem yields

$$\int_{\Psi(\gamma)} \sum_k \varepsilon I_k \,\mathrm{d}\vartheta_k = \int_\gamma \Psi^*\left(\sum_k \varepsilon I_k \,\mathrm{d}\vartheta_k\right) = \int_\Sigma \Psi^*\omega = \int_\Sigma \omega = \int_\gamma \sum_k \varepsilon I_k \,\mathrm{d}\vartheta_k.$$

Since (9.6.20) implies $\mathrm{d}\boldsymbol{\vartheta}^{(N)} = \mathrm{d}\boldsymbol{\vartheta}$, we conclude that

$$\int_\gamma \sum_k \left(I_k^{(N)} - I_k\right) \mathrm{d}\vartheta_k = 0$$

[22]See Lemma 3 in Sect. 2 of [225].

for any closed curve γ in T^n_0. Thus, there exists a smooth potential $S = S(\boldsymbol{\vartheta}, \varepsilon)$ on T^n_0 such that

$$I_k^{(N)} - I_k = \frac{\partial S}{\partial \vartheta_k}.$$

Since every smooth function on a compact manifold has a critical point, there exists $\mathbf{I} = \mathbf{I}(\boldsymbol{\vartheta}, \varepsilon)$ satisfying $\mathbf{I}^{(N)} = \mathbf{I}$. This completes the proof of the Theorem. □

Remark 9.6.12 Let γ be an elliptic, r-elementary and non-degenerate periodic integral curve of energy E of a Hamiltonian system (M, ω, H). Application of the Birkhoff-Lewis Theorem to an isoenergetic Poincaré mapping $(\mathscr{P}_E, \mathscr{W}_E, \Theta_E)$ at $m_0 \in \gamma$ yields that every neighbourhood of m_0 in $\mathscr{W}_E$ contains infinitely many periodic points of Θ_E, where the number of periodic points with period smaller than a given constant is finite. Consequently, every neighbourhood of γ in the energy surface Σ_E contains infinitely many points located on periodic integral curves of X_H. Note that this does not imply that the neighbourhood contains these periodic integral curves.

Example 9.6.13 (Two degrees of freedom) The geometric picture following from Theorems 9.6.7 and 9.6.10 becomes especially transparent for the case of two degrees of freedom, which has been analyzed in detail by Siegel and Moser [273, §§32–35]. In this case, every Poincaré section $\mathscr{P}_E$ is 2-dimensional, the invariant tori are circles with radius $2I$ enclosing the fixed point $(0, 0) \in \mathbb{R}^2$ and the I-dependent angle of rotation is given by $\alpha(I) = a + \beta I$. Here, non-degeneracy means $\beta \neq 0$. One can show that the concentric circles whose radii fulfil the condition

$$\left| \frac{\alpha(I)}{2\pi} k - l \right| \geq \frac{C}{k^\mu} \tag{9.6.21}$$

for all $k, l \in \mathbb{Z}$ with $k \geq 1$ survive the perturbation caused by passing to the full mapping Ψ. Here, C and μ are positive numbers, which do not depend on k and l. Condition (9.6.21) is the strong non-resonance condition, also called Diophantine condition, which we alluded to in Remark 9.6.8.

It is instructive to try to construct the invariant curves for the full mapping by using a formal power series expansion for f and g, see §32 in [273]. This way one gets insight into the role played by the Diophantine condition (9.6.21). Performing this procedure one gets a formal power series for the invariant curve, indeed. However, it is impossible to verify the convergence of these series.[23] Therefore, a stronger tool must be used—the rapidly convergent iteration scheme of Kolmogorov. As a consequence, for the tori fulfilling inequality (9.6.21) one gets convergence, indeed.

[23]This problem is due to the occurrence of "small divisors" in these power series. We met such divisors in the proof of Theorem 9.6.2, cf. Eq. (9.6.9). There, in order to solve for the coefficients $F_{\rho\sigma}$ one has to divide by the number $\lambda^{\rho-\sigma} - 1$, which is different from zero under the non-resonance assumption. However, for the convergence of the above power series, non-vanishing of such numbers is not sufficient. They have to be sufficiently large.

These tori constitute a Cantor set with non-vanishing measure in $\mathbb{R}_+$. In particular, there exist uncountably many such curves.

The circles for which α is resonant do not survive the perturbation.[24] However, from the Birkhoff-Lewis Theorem we know that we still have periodic integral curves for the full system. These periodic integral curves cannot lie on the surviving tori, they must be located between them. In the next section we will draw conclusions on stability from this fact. In general, among the periodic integral curves provided by the Birkhoff-Lewis Theorem there are curves of hyperbolic and elliptic type. Among the elliptic ones in general we can find again integral curves which are non-degenerate and 4-elementary. Around them, the above picture is being reproduced on a smaller scale. The dynamics in the neighbourhood of hyperbolic integral curves turns out to be more complicated. For an illustration of these phenomena we refer to Figs. 8.3-2 and 8.3-3 in [1], which go back to Arnold [12].

Exercises

9.6.1 Analyze the flow of the Hamiltonian

$$H = \frac{1}{2}\varepsilon\big(q^2 + p^2\big) + \big(q^3 - 3qp^2\big).$$

Hint. Consider the cases $\varepsilon < 0$, $\varepsilon = 0$ and $\varepsilon > 0$ separately.
Note. This is an example of a resonance of third order taken from Sect. 4 of [18].

9.7 Stability

For a Hamiltonian vector field, the problem of stability is a rather difficult task. To understand this, first recall that to every characteristic exponent of an equilibrium γ with negative real part there corresponds a characteristic exponent with positive real part, cf. Proposition 8.2.10. Accordingly, to every characteristic multiplier of a periodic integral curve γ inside the unit circle there corresponds a characteristic multiplier outside. Hence, γ cannot be asymptotically stable,[25] and if it is stable, then all of its characteristic exponents have zero real part (in case γ is an equilibrium) or all of its characteristic multipliers have modulus 1 (in case γ is periodic). In this case, the stable and unstable manifolds $S^-(\gamma)$ and $S^+(\gamma)$ coincide with γ itself and the whole manifold is a centre manifold. As a consequence, the usual stability theory for hyperbolic critical integral curves is pointless for Hamiltonian systems, and little is known in the general case.

In this section, we will limit our attention to autonomous systems and discuss three special situations. First, we will show that under a special assumption the

---

[24] This was already observed by Poincaré.

[25] Of course, this also follows from the Liouville Theorem 9.1.4.

Hamiltonian H possesses the properties of a Lyapunov function. Then, we will analyze the case of an equilibrium with purely imaginary characteristic exponents. Finally, we will discuss periodic integral curves in two degrees of freedom.

Thus, let (M, ω, H) be a Hamiltonian system and let γ be a critical integral curve. To find a criterion under which H is a Lyapunov function for γ, consider the second derivative H''_m of H at a point $m \in \gamma$, viewed as a symmetric bilinear form on $\mathrm{T}_m M$. Since H is constant along γ we have $H''_m(X, Y) = 0$ whenever X or Y is tangent to γ. Consequently, H''_m induces a mapping

$$\mathrm{N}_m\gamma \times \mathrm{N}_m\gamma \to \mathbb{R}, \qquad \big([X], [Y]\big) \mapsto H''_m(X, Y), \tag{9.7.1}$$

called the normal second derivative. If γ is an equilibrium, then $\mathrm{N}_m\gamma = \mathrm{T}_m M$ and the normal second derivative coincides with H''_m.

Proposition 9.7.1 *Let (M, ω, H) be a Hamiltonian system and let γ be a critical integral curve. If for all $m \in \gamma$ the second normal derivative of H at m is positive definite, γ is stable.*

Let us add that this assumption also implies that H is a Morse-Bott function in some neighbourhood of γ.

Proof Denote $c := H(\gamma)$. Choose a covering of γ by flow box charts (U_i, κ_i). Since the chart κ_i maps $\gamma \cap U_i$ onto an open interval of the x^1-axis, the normal second derivative of H at a point $m \in \gamma \cap U_i$ is given in this chart by the second derivative of the restriction of $H \circ \kappa_i^{-1}$ to the hyperplane defined by $x^1 = \kappa_i^1(m)$. Since, by assumption, the normal second derivative is positive definite, the point $\kappa_i(m)$ is a local mininum of the restriction. Since this is true for all $m \in \gamma \cap U_i$ we conclude that there exists a neighbourhood V_i of $\gamma \cap U_i$ in U_i where $H(m) > c$ for all $m \in V_i \setminus \gamma$. Then, $V := \bigcup_i V_i$ is a neighbourhood of γ in M with $H(m) > c$ for all $m \in V \setminus \gamma$. Thus, the function $H_{\restriction V} - c$ is a Lyapunov function for γ and Theorem 3.8.16 yields the assertion. □

Example 9.7.2 Let $M = \mathbb{R}^{2n}$ with the canonical symplectic structure and the standard Darboux coordinates q^i and p_i. Consider a Hamiltonian of the form

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2}\mathbf{p}^{\mathrm{T}} M^{-1}\mathbf{p} + V(\mathbf{q}),$$

where M is some positive symmetric invertible matrix and V is an arbitrary smooth function on $\mathbb{R}^n$. Obviously, $(\mathbf{q}_0, \mathbf{p}_0)$ is an equilibrium of X_H iff $\mathbf{p}_0 = 0$ and $\mathbf{q}_0$ is a critical point of V. The second derivative at $(\mathbf{q}_0, \mathbf{p}_0 = 0)$ is

$$H''(\mathbf{q}_0, 0) = \begin{bmatrix} M & 0 \\ 0 & \frac{\partial^2 V}{\partial q^i \partial q^j}(\mathbf{q}_0) \end{bmatrix}.$$

It is positive definite iff so is the second derivative of V at $\mathbf{q}_0$. Thus, Proposition 9.7.1 yields that if $\mathbf{q}_0$ is a local minimum of V_0, the equilibrium $(\mathbf{q}_0, 0)$ is stable under the flow of X_H.

Next, we analyze the case of an equilibrium whose characteristic exponents are purely imaginary. The following theorem is a direct consequence of the Lyapunov Centre Theorem 9.5.8 and the fact that the orbit cylinder provided by this theorem combines with the equilibrium point to a two-dimensional invariant embedded submanifold of M, cf. Remark 9.5.9/1.

Theorem 9.7.3 (Lyapunov Subcentre Theorem) *Let m be an equilibrium of a Hamiltonian system (M, ω, H), satisfying the following assumptions.*

1. *Zero is not a characteristic exponent.*
2. *If λ is a purely imaginary characteristic exponent, it has multiplicity* 1.
3. *If λ_1 and λ_2 are purely imaginary characteristic exponents, they are rationally independent.*

Then, for every purely imaginary characteristic exponent $\mathrm{i}\alpha$ there exists a two-dimensional X_H-invariant embedded symplectic submanifold C_α of M containing m with the following properties:

1. $\mathrm{T}_m C_\alpha$ *is the eigenspace of the characteristic exponents $i\alpha$ and $-i\alpha$.*
2. *$C_\alpha \setminus \{m\}$ is the image of a regular orbit cylinder whose integral curves γ_E approach m and whose periods T_E approach $\frac{2\pi}{\alpha}$ in the limit $E \to H(m)$. Thus, C_α is a union of periodic integral curves and $\{m\}$, that is, it is diffeomorphic to a* 2*-dimensional disc.*

The submanifold C_α is referred to as the subcentre manifold associated with the pair of characteristic exponents $\pm\mathrm{i}\alpha$.

Besides the fact that the Lyapunov Subcentre Theorem provides insight into the behaviour of the flow near m whenever characteristic exponents satisfying the assumptions exist, it yields the following information about stability: if m is elliptic and if the characteristic exponents, counted with multiplicities, are pairwise rationally independent, then m is stable in each subcentre manifold C_α. Indeed, given a (contractible) neighbourhood U of m in C_α we find an energy E such that γ_E is contained in U. The subset of C_α enclosed by γ_E is invariant under the flow and hence it stays in U for all times. To be stable in each subcentre is, of course, only a necessary condition for m to be stable in M. Whether or under which conditions it is sufficient is a hard (and open) question, which remains to be studied.

Example 9.7.4 (Subcentres in two degrees of freedom) Assume that $\dim M = 4$ and that m fulfils the assumptions of the Lyapunov Subcentre Theorem. Then, characteristic exponents can be grouped into two pairs $\pm\lambda_1$ and $\pm\lambda_2$. In the trivial case where none of them is purely imaginary, these pairs are complex conjugates of one another. If one of the pairs is purely imaginary, $\lambda_1 = \mathrm{i}\alpha$ with $\alpha > 0$ say, we are in one of the following two cases.

1. $\lambda_2 = \mu$ is real and positive: since $-\mu < 0 < \mu$, both the stable manifold and the unstable manifold have dimension one and the subcentre manifold C_α is in fact a centre manifold for m. Correspondingly, in the 2-dimensional subspace spanned

by the eigenspaces of the real characteristic exponents $\pm\mu$, the linearized flow has the form of a saddle (type 2(c) in Example 3.6.13). By the Lyapunov Centre Theorem, in the limit $E \to H(m)$, the Floquet multipliers of the periodic integral curves γ_E of C_α approach $\mathrm{e}^{\pm\frac{2\pi}{\alpha}\mu}$, which are positive real numbers distinct from 1. Since, for dimensional reasons, the Floquet multipliers can only arrive at the real axis if they pass through 1 or -1, there exists $\varepsilon > 0$ such that for all E with $|E - H(m)| < \varepsilon$, the Floquet multipliers are real and distinct from 1 and the periodic integral curves γ_E are hyperbolic on their energy surfaces.

2. $\lambda_2 = \mathrm{i}\beta$ with $\beta > 0$ and rationally independent from α: here, we have two subcentre manifolds, C_α and C_β, consisting of periodic integral curves γ_E^α and γ_E^β, respectively. Since α and β are rationally independent, m is r-elementary for any r, and Remark 9.6.9 yields that we can find Darboux coordinates q_1, q_2, p_1, p_2 at m such that H is a polynomial in the action variables $I_i = \frac{1}{2}(q_i^2 + p_i^2)$ whose lowest order term is given by
$$H(q, p) = \frac{1}{2}(\pm\alpha I_1 \pm \beta I_2) + \cdots .$$
Since H can be expressed entirely in terms of I_1 and I_2, it follows that the subcentre manifolds C_α and C_β are given by $q_2 = p_2 = 0$ and $q_1 = p_1 = 0$, respectively. If the contributions of α and β have the same sign, m is a local minimum or maximum of H and the energy surfaces Σ_E near m are three-spheres containing two periodic integral curves γ_E^α and γ_E^β each, which collapse to m as E approaches $H(m)$. If the contributions of α and β have opposite sign, m is a saddle point of H and Σ_E are hyperboloids, each of which contain one periodic integral curve. To be definite, let us assume that the contribution of α has positive and that of β has negative sign. Then, Σ_E contains γ_E^α for $E > H(m)$ and γ_E^β for $E < H(m)$. As one increases the energy starting from a value below $H(m)$, the periodic integral curves γ_E^β shrink, degenerate to m at $E = H(m)$ and reappear as the periodic integral curves γ_E^α, though in another dimension. This could be imagined as an "evolution" of orbit cylinders, where C_β "dies" at $E = H(m)$ and is "reborn" as the orbit cylinder C_α. In either case, for $E \to H(M)$, the Floquet multipliers of γ_E^α approach $\mathrm{e}^{\pm\mathrm{i}\frac{2\pi}{\alpha}\beta}$ and those of γ_E^β approach $\mathrm{e}^{\pm\mathrm{i}\frac{2\pi}{\beta}\alpha}$.

The authors of [1] call case 1 the phantom burst and case 2 the stable burst catastrophe. For a further discussion of bifurcations in the case of two degrees of freedom we refer to Sect. 8.6 of [1] (where this example was taken from).

Finally, we discuss the stability of a periodic integral curve γ of a Hamiltonian system (M, ω, H) with two degrees of freedom, that is, with $\dim M = 4$. Assume that γ is elliptic, 4-elementary and non-degenerate. From Example 9.6.13 we know that every neighbourhood of γ in its (local) energy surface Σ_E contains uncountably many 2-tori which are invariant under X_H and in which every integral curve is dense. The latter property implies that the tori cannot intersect. Therefore, they divide the 3-dimensional energy surface Σ_E into disjoint subsets which are also invariant under X_H. As a consequence, γ is stable within its energy surface Σ_E. If one combines this observation with the existence of an orbit cylinder, one finds that γ is stable:

Proposition 9.7.5 *Every elliptic*, 4*-elementary and non-degenerate periodic integral curve of a* 4*-dimensional Hamiltonian system is stable.*

Proof Denote the periodic integral curve under consideration by γ_{E_0}, where $E_0 = H(\gamma_{E_0})$. Since H does not have critical points in some open neighbourhood of γ_{E_0} in M, and since stability is a local property, for convenience we may assume that H is a submersion and hence M is foliated by regular energy surfaces Σ_E. Since, by assumption, 1 is not a Floquet multiplier of γ_{E_0}, Theorem 9.5.4 yields that γ_{E_0} is contained in a regular orbit cylinder $\zeta : \mathrm{S}^1 \times I \to M$. Consider energy values $E \in I$. Since the conditions on γ_E to be elliptic, to be 4-elementary and to be non-degenerate are all open conditions on the Birkhoff invariants[26] of γ_E and since the latter depend smoothly on E, by possibly shrinking I we may assume that γ_E is elliptic, 4-elementary and non-degenerate for all $E \in I$. Now, let U be a neighbourhood of γ_{E_0} in M. By possibly further shrinking I, we may assume that $\zeta(\mathrm{S}^1 \times I) \subset U$. Then, for all $E \in I$, $U \cap \Sigma_E$ is a neighbourhood of γ_E in the energy surface Σ_E. Since γ_E is stable in Σ_E, we find a neighbourhood V_E of γ_E in Σ_E where the flow of X_H is defined for all times and whose points never leave U under this flow. Then, $V = \bigcup_{E \in I} V_E$ is a neighbourhood of γ in M with the same property. □

For an application of the above stability analysis to celestial mechanics we refer to [273], §34. For example, the motion of an asteroid in the asteroid belt between the planets Mars and Jupiter can be viewed as a restricted 3-body problem,[27] with the three bodies being the asteroid, Jupiter and the sun and with the mass of the asteroid being completely neglected. If one assumes that the asteroid moves in the same plane as the sun and Jupiter, this system has two degrees of freedom. The ratios between the frequencies ω_A of the asteroid and ω_J of Jupiter for which the Floquet multipliers of the corresponding periodic integral curve do not fulfil the non-resonance condition are $\frac{1}{4}, \frac{2}{5}, \frac{1}{3}$ and $\frac{1}{2}$. For these ratios one finds well-defined gaps in the distribution of asteroids, known as the Kirkwood gaps.[28] Hence, these orbits are not stable, indeed. After the long history of the solar system, the asteroids have disappeared from them. For an extensive discussion of these and further gaps, see [216].

Remark 9.7.6 (Arnold diffusion) For systems with more than two degrees of freedom, the invariant tori have codimension greater than 1 in their energy surface. Consequently, these tori do not divide the energy surface into disjoint subsets, so that an

[26] Defined by an isoenergetic Poincaré mapping for γ_E, cf. Remark 9.6.5/1.

[27] See also Sect. 4.4 of the book of Thirring [286].

[28] Named after the astronomer who first observed them in 1866.

integral curve can escape from a neighbourhood of γ without running through such a torus. Thus, the above stability argument does no longer apply. This phenomenon is known as Arnold diffusion.

9.8 Time-Dependent Systems. Parametric Resonance

For the remainder of this chapter we depart from autonomous systems and consider time-dependent Hamiltonian systems (M, ω, H). Here, the Hamiltonian H is a smooth function on $\mathbb{R} \times M$, cf. Sect. 9.3. We restrict attention to the case where the time dependence is periodic, that is, where there exists a minimal positive real number T, called the period, such that $H_{t+T} = H_t$ as functions on M for all t. In the present section, we study the problem of stability of equilibria and in the next section we address the problem of the existence of periodic integral curves.

Let us start with some introductory remarks. The Hamiltonian vector field X_H generated by a T-periodic Hamiltonian function H is also T-periodic, because

$$(X_H)_{t+T} \equiv X_{H_{t+T}} = X_{H_t} \equiv (X_H)_t$$

for all t. Let Φ denote the flow of X_H. From Sect. 3.4 we know that Φ satisfies

$$\Phi_{t+T,0} = \Phi_{t,0} \circ \Phi_{T,0}, \qquad \Phi_{kT,0} = \Phi_{T,0}^k \tag{9.8.1}$$

for all $t \in \mathbb{R}$ and $k \in \mathbb{Z}$. The local symplectomorphism $\Phi_{T,0}$ is called the period mapping of X_H. Note that the period mapping is a Hamiltonian symplectomorphism in the sense of Definition 8.8.7 and that every integral curve is invariant under $\Phi_{T,0}$. In particular, equilibria of X_H are fixed points of $\Phi_{T,0}$.

Proposition 9.8.1 *A critical integral curve of a periodically time-dependent Hamiltonian system is stable iff it is stable under the period mapping.*

Proof In the proof of the equivalence of points 1 and 2 of Proposition 3.8.5, allow γ to be an equilibrium as well and replace the period mapping Φ_T of the periodic integral curve given there by the period mapping $\Phi_{T,0}$ of X_H. □

Now, we turn to the study of the stability properties of an equilibrium of X_H. Since stability is a local concept, we may assume that $M = \mathbb{R}^{2n}$ with the standard symplectic structure and that the equilibrium under consideration is given by the origin. We restrict attention to the case where H_t is a quadratic form on $\mathbb{R}^{2n}$ for all t. This is a model for a system of coupled harmonic oscillators with T-periodic characteristic frequencies and coupling constants. Since H is quadratic in the variable $\mathbf{x} \in \mathbb{R}^{2n}$, the Hamiltonian vector field X_H is linear, and so are the flow Φ and the period mapping $\Phi_{T,0}$. Thus, $\Phi_{T,0} \in \mathrm{Sp}(n, \mathbb{R})$. For the elements of $\mathrm{Sp}(n, \mathbb{R})$, Proposition 3.8.9/2 yields the following simple stability criterion.

Proposition 9.8.2 *Let $a \in \mathrm{Sp}(n, \mathbb{R})$. If all eigenvalues of a have absolute value* 1 *and multiplicity* 1, *there exists a neighbourhood of a in* $\mathrm{Sp}(n, \mathbb{R})$ *such that the origin is stable under all symplectomorphisms from this neighbourhood.*

Since the multiplicity of an eigenvalue ± 1 must be even, the assumptions imply that the eigenvalues of a are distinct from ± 1. Note that the proposition provides both a criterion for the stability of the origin under a symplectomorphism and a criterion for the structural stability[29] of a family of linear symplectomorphisms with respect to the property to have the origin as a stable fixed point.

Proof It is enough to show that a possesses a neighbourhood in $\mathrm{Sp}(n,\mathbb{R})$ whose elements have non-degenerate eigenvalues lying on the unit circle, because any automorphism b of $\mathbb{R}^{2n}$ with this property satisfies $\mathrm{spec}(b) = \mathrm{spec}_1^d(b)$, so that stability follows by Proposition 3.8.9/2. To construct the desired neighbourhood of a, we choose pairwise disjoint neighbourhoods U_λ of the eigenvalues λ of a in $\mathbb{C}$, satisfying

$$U_{\lambda^{-1}} = U_\lambda^{-1} = \overline{U_\lambda},$$

where $\overline{U_\lambda}$ means the complex conjugate subset. Since the eigenvalues depend continuously on the elements of $\mathrm{Sp}(n,\mathbb{R})$, there exists a neighbourhood V of a in $\mathrm{Sp}(n,\mathbb{R})$ whose elements have exactly one eigenvalue in each neighbourhood U_λ. For every element b of this neighbourhood, the eigenvalues necessarily have multiplicity 1. We show that they have absolute value 1. Indeed, if μ is an eigenvalue of b, then so is $\overline{\mu}^{-1}$. If λ is the eigenvalue of a such that $\mu \in U_\lambda$, then

$$\overline{\mu}^{-1} \in \overline{U_\lambda}^{-1} = U_\lambda .$$

Since U_λ contains only one eigenvalue of b, we conclude $\overline{\mu}^{-1} = \mu$ and hence $|\mu| = 1$. □

Remark 9.8.3 In the case of degenerate eigenvalues on the unit circle, Proposition 9.8.2 does not provide information about stability. In this situation, the origin may be stable as in the case $a = \mathbb{1}$ or unstable as in the case

$$a = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

(Exercise 9.8.1). The proposition implies, however, that the transition between stable and unstable linear symplectomorphisms is only possible via degenerate eigenvalues. This transition can be imagined as follows. Two pairs of eigenvalues on the unit circle run (in opposite directions) into one another, meet at a certain eigenvalue pair and then escape along the two radial lines defined by this pair into opposite directions.[30] Thus, for a linear symplectomorphism with degenerate eigenvalues, even a small perturbation can cause a transition between stability and instability. This phenomenon is referred to as parametric resonance. Note, however, that the fact that a has degenerate eigenvalues does not necessarily imply that in every neighbourhood of a there exist stable and unstable linear symplectomorphisms. For a criterion in that case, established by M.G. Krein, we refer to [18, §42].

---

[29] Cf. Remark 3.8.19.

[30] Compare with Fig. 7.1.

In what follows we study the case of one degree of freedom, that is, $M = \mathbb{R}^2$. Under the assumption that H_t be a quadratic form on $\mathbb{R}^2$ for all t, H is of the form

$$H^{\pm}(q, p) = \frac{1}{2m} p^2 \pm \frac{m\omega(t)^2}{2} q^2, \tag{9.8.2}$$

with T-periodic frequency ω. The positive sign refers to an attractive and the negative sign to a repelling force. The Hamiltonian vector field is

$$X_{H^{\pm}} = \mp m\omega(t)^2 q \partial_p + \frac{p}{m}\partial_q. \tag{9.8.3}$$

Denote the flow of $X_{H^{\pm}}$ by $\Phi^{H^{\pm}}$. The stability criterion of Proposition 9.8.2 takes the following form.

Proposition 9.8.4 *Let* $a \in \mathrm{Sp}(1, \mathbb{R})$.

1. *If* $|\operatorname{tr}(a)| < 2$, *there exists a neighbourhood of* a *in* $\mathrm{Sp}(1, \mathbb{R})$ *such that the origin is stable under all symplectomorphisms from that neighbourhood.*
2. *If* $|\operatorname{tr}(a)| = 2$, *then* $a = \pm\mathbb{1}$ *and the origin is stable under* a.
3. *If* $|\operatorname{tr}(a)| > 2$, *there exists a neighbourhood of* a *in* $\mathrm{Sp}(1, \mathbb{R})$ *such that the origin is unstable under all symplectomorphisms from that neighbourhood.*

As a consequence, the equation $|\operatorname{tr}(a)| = 2$ defines the boundary of the subset of $\mathrm{Sp}(1, \mathbb{R})$ of symplectomorphisms under which the origin is stable.

Proof Let λ_1 and λ_2 be the eigenvalues of a. By the Symplectic Eigenvalue Theorem 7.4.3, λ_2 must coincide with $\overline{\lambda_1}$ or λ_1^{-1}. Therefore, the following cases can occur:

(a) $\lambda_1 = \lambda_2$: here, $\lambda_1 = \lambda_2 = \pm 1$ and hence $|\operatorname{tr}(a)| = 2$.
(b) $|\lambda_1| = |\lambda_2| = 1$ but $\lambda_1 \neq \lambda_2$: here, $\lambda_2 = \overline{\lambda_1} \neq \pm 1$ and hence

$$\big|\operatorname{tr}(a)\big| = |\lambda_1 + \overline{\lambda_1}| = 2\big|\operatorname{Re}(\lambda_1)\big| < 2.$$

(c) $|\lambda_1| \neq |\lambda_2|$: here, $\lambda_2 = \lambda_1^{-1}$, where λ_1 is real and distinct from ± 1. Hence,

$$0 < |\lambda_1|^{-1}\big(|\lambda_1| - 1\big)^2 = |\lambda_1| + |\lambda_1|^{-1} - 2 = \big|\operatorname{tr}(a)\big| - 2.$$

It follows that the cases (a), (b) and (c) are equivalently characterized by the conditions, respectively, $|\operatorname{tr}(a)| = 2$, $|\operatorname{tr}(a)| < 2$ and $|\operatorname{tr}(a)| > 2$. Thus, for the cases $|\operatorname{tr}(a)| = 2$ and $|\operatorname{tr}(a)| > 2$, the assertion is obvious and for the case $|\operatorname{tr}(a)| < 2$, it follows from Proposition 9.8.2. □

Combining Proposition 9.8.4 with Proposition 9.8.1, we obtain

Corollary 9.8.5 *For a Hamiltonian system on* $\mathbb{R}^2$ *whose Hamiltonian function is of the form* (9.8.2), *the origin is stable iff* $|\operatorname{tr}(\Phi^{H^{\pm}}_{T,0})| \leq 2$.

This way, the stability analysis for systems with Hamiltonian function (9.8.2) reduces to the problem of calculating the quantity $\operatorname{tr}(\Phi^{H^{\pm}}_{T,0})$.

Remark 9.8.6 In the case where $\omega(t)$ performs small oscillations of period T about a constant value ω_0, Proposition 9.8.4 yields further information about the qualitative behaviour of the system. Write $\omega(t) = \omega_\varepsilon(t)$ with $\omega_0(t) = \omega_0$ and, correspondingly, $H^\pm = H^\pm_\varepsilon$. Assume that the mapping $\varepsilon \mapsto \omega_\varepsilon$ is continuous with respect to the supremum norm for functions on $[0, T]$, that is, $\varepsilon_n \to \varepsilon$ implies[31]

$$\sup\{|\omega_{\varepsilon_n}(t) - \omega_\varepsilon(t)| : t \in [0, T]\} \to 0.$$

Then, $|\operatorname{tr}(\Phi^{H^\pm_\varepsilon}_{T,0})|$ depends continuously on ε, too. Let us compute $|\operatorname{tr}(\Phi^{H^\pm_0}_{T,0})|$. In the case of an attractive force, the flow of the harmonic oscillator with constant frequency ω_0 is given by

$$\Phi^{H^+_0}_{t,0} = \begin{bmatrix} \cos(\omega_0 t) & \frac{1}{m\omega_0}\sin(\omega_0 t) \\ -m\omega_0 \sin(\omega_0 t) & \cos(\omega_0 t) \end{bmatrix},$$

hence

$$\left|\operatorname{tr}\left(\Phi^{H^+_0}_{T,0}\right)\right| = 2\left|\cos(\omega_0 T)\right|.$$

Thus, if $T \notin \frac{\pi}{\omega_0}\mathbb{Z}$, Proposition 9.8.4 implies that the origin remains stable for sufficiently small ε. If $T \in \frac{\pi}{\omega_0}\mathbb{Z}$, the origin is stable for $\varepsilon = 0$, but a transition to instability is possible for arbitrarily small values of ε. For such values of the period, parametric resonance can occur. In the case of a repulsive force,

$$\Phi^{H^-_0}_{t,0} = \begin{bmatrix} \cosh(\omega_0 t) & \frac{1}{m\omega_0}\sinh(\omega_0 t) \\ -m\omega_0 \sinh(\omega_0 t) & \cosh(\omega_0 t) \end{bmatrix},$$

and hence

$$\left|\operatorname{tr}\left(\Phi^{H^-_0}_{T,0}\right)\right| = 2\left|\cosh(\omega_0 T)\right|.$$

Since $\omega_0 > 0$ and $T > 0$, we have $|\operatorname{tr}(\Phi^{H^-_0}_{T,0})| > 2$, so that the origin remains unstable for sufficiently small values of ε.

Now, we consider an example where $\Phi^{H^\pm_\varepsilon}_{T,0}$ can be calculated explicitly. This allows to go beyond the qualitative statements of Remark 9.8.6 and to analyze the stability of the origin for arbitrary values of ε.

Example 9.8.7 (Swing) Let

$$\omega_\varepsilon(t) = \omega_0 + \begin{cases} \varepsilon & |\ t \in [kT, \frac{2k+1}{2}T[\\ -\varepsilon & |\ t \in [\frac{2k-1}{2}T, kT[\end{cases}, \quad k \in \mathbb{Z},\ 0 \le \varepsilon < \omega_0. \tag{9.8.4}$$

With this time-dependence of the frequency, the Hamiltonian (9.8.2) yields a simple model for a swing. To see this, consider a planar pendulum with time-dependent

[31] Note that we do not assume $\omega_\varepsilon(t)$ to be continuous in t.

length $l(t)$. Let g be the gravitational acceleration and let ϕ be the angle of deviation from the lower equilibrium point. Then, the equation of motion reads

$$\ddot{\phi} + 2\frac{\dot{l}\dot{\phi}}{l} + \frac{g}{l}\sin\phi = 0. \tag{9.8.5}$$

The mixed term $2\frac{\dot{l}\dot{\phi}}{l}$ will be ignored. The motion of the person sitting on the swing is modelled by the following law for $l(t)$:

$$l = \begin{cases} l_0 \frac{\omega_0^2}{(\omega_0+\varepsilon)^2} & | \; t \in [kT, \frac{2k+1}{2}T[\\ l_0 \frac{\omega_0^2}{(\omega_0-\varepsilon)^2} & | \; t \in [\frac{2k-1}{2}T, kT[\end{cases}, \quad k \in \mathbb{Z}, \; \omega_0 = \sqrt{\frac{g}{l_0}}.$$

Here, $l_0 \frac{\omega_0^2}{(\omega_0+\varepsilon)^2} < l_0$ is the pendulum length which effectively occurs if the person on the swing leans forward and $l_0 \frac{\omega_0^2}{(\omega_0-\varepsilon)^2} > l_0$ is the effective length if the person leans backward. For small deviations about the lower or the upper equilibrium point, $\sin\phi$ can be approximated by $\pm\phi$, respectively. Under this approximation, Eq. (9.8.5) is equivalent to the Hamilton equations for the Hamiltonian (9.8.2) with $p := m\dot{\phi}$, $q := l\phi$ and with frequency given by (9.8.4).

First, let us consider H_ε^+, which corresponds to small deviations from the lower equilibrium point. Integration of the Hamilton equations yields (Exercise 9.8.2)

$$\left|\mathrm{tr}\left(\Phi_{T,0}^{H_\varepsilon^+}\right)\right| = 2\left|\frac{\omega_0^2}{\omega_0^2 - \varepsilon^2}\cos(\omega_0 T) - \frac{\varepsilon^2}{\omega_0^2 - \varepsilon^2}\cos(\varepsilon T)\right|. \tag{9.8.6}$$

We keep the period T fixed and study the stability of the origin, that is, of the lower equilibrium point of the swing, depending on the values of the parameters ω_0 and ε. Regions in the ω_0–ε-plane, for which the origin is stable or unstable, respectively, will be called stable or unstable regions, respectively. The boundaries between such regions are called stability boundaries. According to Proposition 9.8.4, these boundaries are defined by $|\,\mathrm{tr}(\Phi_{T,0}^{H_\varepsilon^+})| = 2$. In our example, this corresponds to the equation

$$\omega_0^2\big(1 \pm \cos(\omega_0 T)\big) = \varepsilon^2\big(1 \pm \cos(\varepsilon T)\big), \tag{9.8.7}$$

see Fig. 9.4. We see that, apart from the points $\omega_0 = \frac{k\pi}{T}$, the ω_0-axis is enclosed by stable regions, that is, there the origin remains stable for sufficiently small values of ε. At the exceptional points $\omega_0 = \frac{k\pi}{T}$, the unstable regions touch the ω_0-axis. Thus, at these points the origin is unstable for arbitrarily small values of ε. For the swing, this means that when swinging with a multiple of half the characteristic frequency of the swing, an arbitrarily small perturbation suffices to set the swing in motion. Indeed, for $\omega_0 T = (2k+1)\pi$, Eq. (9.8.6) yields

$$\left|\mathrm{tr}\left(\Phi_{T,0}^{H_\varepsilon^+}\right)\right| = 2\left|\frac{\omega_0^2 + \varepsilon^2\cos(\varepsilon\frac{(2k+1)\pi}{\omega_0})}{\omega_0^2 - \varepsilon^2}\right| > 2, \quad 0 < \varepsilon < \frac{\omega_0}{2k+1},$$

and for $\omega_0 T = 2k\pi$ we obtain

$$\left|\mathrm{tr}\left(\Phi_{T,0}^{H_\varepsilon^+}\right)\right| = 2\left|\frac{\omega_0^2 - \varepsilon^2\cos(\varepsilon\frac{2k\pi}{\omega_0})}{\omega_0^2 - \varepsilon^2}\right| > 2, \quad 0 < \varepsilon < \frac{\omega_0}{k}.$$

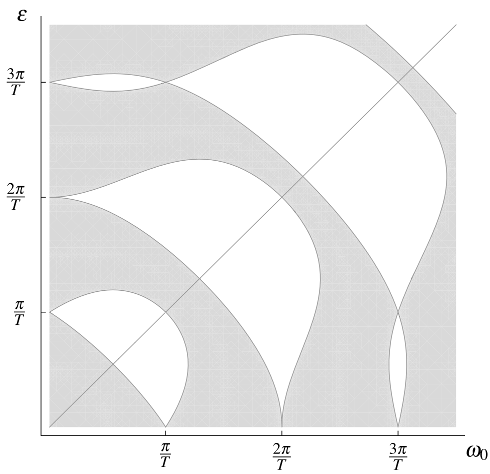


Fig. 9.4 Stable (*grey*) and unstable (*white*) regions in the ω_0–ε-plane for the harmonic oscillator with attractive force and time-dependent frequency give by (9.8.4)

However, Fig. 9.4 shows that the opening angles under which the unstable regions touch the ω_0-axis depend drastically on whether an integer multiple or a half integer multiple of the characteristic frequency is approached. In the first case, one must meet the characteristic frequency rather exactly or change the effective length heavily in order to reach the unstable region.[32] On the other hand, for a half integer multiple of the characteristic frequency it is enough to meet the characteristic frequency approximately and to produce a small change of the effective length by swinging. To calculate the opening angles, we write $\omega_0 = \frac{k\pi}{T} + x$ and expand the left and the right hand side of Eq. (9.8.7) to the lowest non-vanishing order about $x = 0$ and $\varepsilon = 0$, respectively. Using that the sign in (9.8.7) is positive for odd k, we obtain $\frac{k^2\pi^2}{2}x^2 + O(x^3)$ for the left hand side and $2\varepsilon^2 + O(\varepsilon^4)$ for the right hand side. Thus, $\varepsilon \sim \pm\frac{k\pi}{2}x$ and the opening angle is

$$\alpha = \arccos\left(\frac{k^2\pi^2 - 4}{k^2\pi^2 + 4}\right).$$

A similar calculation for even k yields $\varepsilon \sim \pm\sqrt{\frac{k\pi}{T}x}$ and hence $\alpha = 0$.

Next, let us consider the case H_ε^-, which corresponds to small deviations from the upper equilibrium position. Here, one finds (Exercise 9.8.2)

$$\left|\mathrm{tr}\left(\Phi_{T,0}^{H_\varepsilon^-}\right)\right| = 2\left|\frac{\omega_0^2}{\omega_0^2 - \varepsilon^2}\cosh(\omega_0 T) - \frac{\varepsilon^2}{\omega_0^2 - \varepsilon^2}\cosh(\varepsilon T)\right|. \tag{9.8.8}$$

Since the expression under the absolute value signs is larger than 1 for all $0 < \varepsilon < \omega_0$, the upper equilibrium point remains unstable for any choice of the time dependence of the pendulum length.

[32] The reader should try to set a swing in motion by leaning forward and backward over a whole period each time.

In a similar manner, the reader may discuss the following example (Exercise 9.8.3).

Example 9.8.8 Let

$$\omega_\varepsilon(t)^2 = \begin{cases} \omega_0^2 + \varepsilon^2 & \mid t \in [kT, \frac{2k+1}{2}T[\\ \omega_0^2 - \varepsilon^2 & \mid t \in [\frac{2k-1}{2}T, kT[\end{cases}, \quad k \in \mathbb{Z}. \tag{9.8.9}$$

The corresponding Hamiltonians (9.8.2) model small deviations from the equilibria of a planar pendulum whose suspension point moves vertically according to

$$a(t) = \begin{cases} \varepsilon^2 l(\frac{(t-kT)^2}{2} - \frac{T^2}{16}) & \mid t \in [kT, \frac{2k+1}{2}T[\\ -\varepsilon^2 l(\frac{(t-kT)^2}{2} - \frac{T^2}{16}) & \mid t \in [\frac{2k-1}{2}T, kT[\end{cases}, \quad k \in \mathbb{Z}.$$

Finally, we discuss an example which can be worked out explicitly by using the theory of Mathieu functions. The latter is treated in great detail in [209].

Example 9.8.9 Let

$$\omega_\varepsilon(t)^2 = \omega_0^2 - \varepsilon^2 \cos(\Omega t), \quad \Omega = \frac{2\pi}{T}. \tag{9.8.10}$$

The corresponding Hamiltonians (9.8.2) model small deviations from the equilibria of a planar pendulum whose suspension point moves vertically according to

$$a(t) = \frac{\varepsilon^2}{\Omega^2} \cos(\Omega t).$$

The equation of motion is

$$\ddot{x} \pm \big(\omega_0^2 - \varepsilon^2 \cos(\Omega t)\big)x = 0,$$

where the positive (negative) sign stands for the lower (upper) equilibrium. Setting $s := \frac{\Omega}{2}t$ and denoting the derivative with respect to s by x', for both cases we obtain the Mathieu equation

$$x'' + \big(a - 2q\cos(2s)\big)x = 0, \tag{9.8.11}$$

with parameters

$$a = \pm 4\frac{\omega_0^2}{\Omega^2}, \qquad q = \pm 2\frac{\varepsilon^2}{\Omega^2}.$$

In the theory of the Mathieu equation one shows that there exist families a_r, $r = 0, 1, 2, \dots$, and b_r, $r = 1, 2, \dots$, of real-valued functions such that Eq. (9.8.11) has an even periodic solution for every $q \in \mathbb{R}$ and $a = a_r(q)$ and an odd periodic solution for every $q \in \mathbb{R}$ and $a = b_r(q)$. After an appropriate normalization, these solutions are called Mathieu functions and are denoted by ce_r and se_r, respectively.

Let us calculate $\Phi_{T,0}^{H_\varepsilon^\pm}$ under the simplifying assumption that $b_r(q) \neq a \neq a_r(q)$. In this case, the general solution of the Mathieu equation can be shown to have the form

$$x(s) = A\mathrm{e}^{\mathrm{i}\nu s}P(s) + B\mathrm{e}^{-\mathrm{i}\nu s}P(-s), \tag{9.8.12}$$

with P being a π-periodic function and $\nu = \nu(a,q)$ being a complex-valued analytic function.[33] Under the assumptions made, either $\nu(q,a) \in \mathbb{R}$ or $\mathrm{Re}(\nu) \in \mathbb{Z}$ and $\mathrm{Im}(\nu) \neq 0$ [3, §20.3]. The first case occurs if $q \geq 0$ and $a_r(q) < a < b_{r+1}(q)$ or $q \leq 0$ and $a_r(q) < a < b_r(q)$, whereas the second case occurs if $q \geq 0$ and $b_r(q) < a < a_r(q)$ or $q \leq 0$ and $b_{2r-1}(q) < a < b_{2r}(q)$ or $a_{2r}(q) < a < a_{2r+1}(q)$. In particular, $\nu(a,q)$ is not an integer. With $p(t) = ml^2\dot{x}(t)$, from (9.8.12) we read off

$$\begin{aligned} x(t) &= A\mathrm{e}^{\mathrm{i}\tilde{\nu}t} P\left(\frac{\Omega}{2}t\right) + B\mathrm{e}^{-\mathrm{i}\tilde{\nu}t} P\left(-\frac{\Omega}{2}t\right), \\ p(t) &= ml^2\left\{A\mathrm{e}^{\mathrm{i}\tilde{\nu}t}\left(\mathrm{i}\tilde{\nu}P\left(\frac{\Omega}{2}t\right) + \frac{\Omega}{2}P'\left(\frac{\Omega}{2}t\right)\right)\right. \\ &\qquad \left. - B\mathrm{e}^{-\mathrm{i}\tilde{\nu}t}\left(\mathrm{i}\tilde{\nu}P\left(-\frac{\Omega}{2}t\right) + \frac{\Omega}{2}P'\left(-\frac{\Omega}{2}t\right)\right)\right\}, \end{aligned}$$

with $\tilde{\nu} = \nu\frac{\Omega}{2}$. The constants A and B are determined by the initial conditions

$$x(0) = (A+B)P(0), \qquad p(0) = ml^2(A-B)\left(\mathrm{i}\tilde{\nu}P(0) + \frac{\Omega}{2}P'(0)\right).$$

Since $P(0) \neq 0$, we can choose $P(0) = 1$. Then,

$$\Phi^{H_\varepsilon^\pm}_{t,0} = \begin{bmatrix} \frac{\mathrm{e}^{\mathrm{i}\tilde{\nu}t}P(\frac{\Omega}{2}t) + \mathrm{e}^{-\mathrm{i}\tilde{\nu}t}P(-\frac{\Omega}{2}t)}{2} & \frac{\mathrm{e}^{\mathrm{i}\tilde{\nu}t}P(\frac{\Omega}{2}t) - \mathrm{e}^{-\mathrm{i}\tilde{\nu}t}P(-\frac{\Omega}{2}t)}{2ml^2F(0)} \\ \frac{ml^2(\mathrm{e}^{\mathrm{i}\tilde{\nu}t}F(t) - \mathrm{e}^{-\mathrm{i}\tilde{\nu}t}F(-t))}{2} & \frac{\mathrm{e}^{\mathrm{i}\tilde{\nu}t}F(t) + \mathrm{e}^{-\mathrm{i}\tilde{\nu}t}F(-t)}{2F(0)} \end{bmatrix},$$

with

$$F(t) = \mathrm{i}\tilde{\nu}P\left(\frac{\Omega}{2}t\right) + \frac{\Omega}{2}P'\left(\frac{\Omega}{2}t\right).$$

For $t = T$ we obtain $P(\frac{\Omega}{2}T) = P(\pi) = P(0) = 1$ and $F(T) = F(0)$. Thus,

$$\Phi^{H_\varepsilon^\pm}_{T,0} = \begin{bmatrix} \cos(\pi\nu) & -\frac{1}{\mathrm{i}ml^2F(0)}\sin(\pi\nu) \\ \mathrm{i}ml^2F(0)\sin(\pi\nu) & \cos(\pi\nu) \end{bmatrix},$$

and, therefore,

$$\left|\mathrm{tr}\left(\Phi^{H_\varepsilon^\pm}_{T,0}\right)\right| = 2\left|\cos(\pi\nu)\right| = 2\sqrt{\cos^2\left(\mathrm{Re}(\nu)\right) + \sinh^2\left(\mathrm{Im}(\nu)\right)}.$$

In the case where $\nu(a,q) \in \mathbb{R}$, we have $|\mathrm{tr}(\Phi^{H_\varepsilon^\pm}_{T,0})| \leq 2$, whereas in the case where $\mathrm{Re}(\nu) \in \mathbb{Z}$ and $\mathrm{Im}(\nu) \neq 0$, we find $|\mathrm{tr}(\Phi^{H_\varepsilon^\pm}_{T,0})| > 2$. Thus, the boundaries of stability are given by the curves $a_r(q)$ and $b_r(q)$. The corresponding curves in the ω_0–ε-plane are given by

$$\omega_0(\varepsilon) = \frac{\Omega}{2}\sqrt{\pm f\left(\pm 2\frac{\varepsilon^2}{\Omega^2}\right)}, \quad f = a_r, b_r,$$

[33] Obviously, ν is determined up to addition of $2k\pi$, $k \in \mathbb{Z}$.

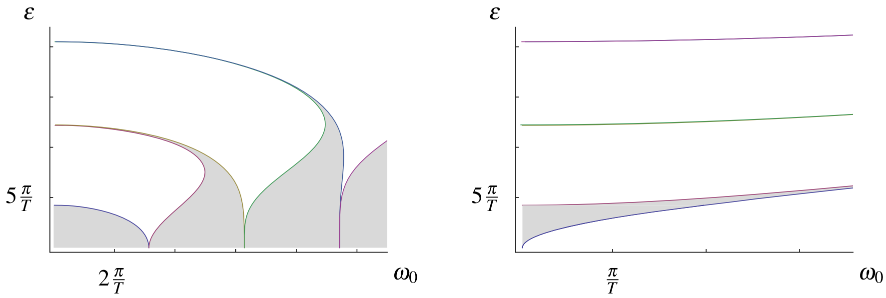


Fig. 9.5 Stable (*grey*) and unstable (*white*) regions in the ω_0–ε-plane for the harmonic oscillator with time-dependent frequency (9.8.10) and attractive (*left*) or repulsive force (*right*)

with the positive sign corresponding to small deviations from the lower equilibrium point and the negative sign corresponding to small deviations from the upper equilibrium point, see Fig. 9.5. The reader can analyze this figure in the same manner as in Example 9.8.7.

Remark 9.8.10 Let us consider the influence of friction, thus leaving the realm of Hamiltonian systems. We restrict attention to the case of the lower equilibrium point. Assuming the friction force to be proportional to the velocity, the equation of motion reads

$$\ddot{x} + 2\beta\dot{x} + \omega_\varepsilon(t)^2 x = 0, \qquad \beta > 0.$$

The corresponding vector field is linear,

$$X^\varepsilon = \frac{p}{m}\partial_x - \big(m\omega_\varepsilon(t)^2 x + \beta p\big)\partial_p,$$

and so is its flow Φ^ε. Therefore, the origin is the only equilibrium. Of course, this vector field is not Hamiltonian and thus the stability criterion of Proposition 9.8.4 cannot be applied. Nonetheless, Proposition 9.8.1 carries over to the present case, and Proposition 3.8.9/2 yields that the origin is asymptotically stable under $\Phi^\varepsilon_{T,0}$, and hence under Φ^ε, if the absolute values of all the eigenvalues of $\Phi^\varepsilon_{T,0}$ are smaller than 1. Let us calculate the eigenvalues of $\Phi^\varepsilon_{T,0}$ for $\varepsilon = 0$, that is, for $\omega_\varepsilon(t) = \omega_0$. According to Example 3.6.13,

$$\Phi^{\varepsilon=0}_{T,0} = \mathrm{e}^{-\beta T}\begin{bmatrix} \cos(\Omega t) + \frac{\beta}{\Omega}\sin(\Omega t) & \frac{1}{m\Omega}\sin(\Omega t) \\ -\frac{m\omega_0^2}{\Omega}\sin(\Omega t) & \cos(\Omega t) - \frac{\beta}{\Omega}\sin(\Omega t) \end{bmatrix},$$

where $\Omega = \sqrt{\omega_0^2 - \beta^2}$. Thus, the eigenvalues are $\lambda_\pm = \mathrm{e}^{-\beta T}(\cos(\Omega T) \pm \mathrm{i}\sin(\Omega T))$ and their absolute value is

$$|\lambda_\pm| = \mathrm{e}^{-\beta T}.$$

Since $\Phi^\varepsilon_{T,0}$ and, therefore, also its eigenvalues are continuous in ε, we conclude that for every ω_0, there exists $\varepsilon_0 > 0$ such that the origin is stable, and hence asymptotically stable, under Φ^ε for all $\varepsilon < \varepsilon_0$. This means that under the influence of friction the regions of instability cease to touch the ω_0-axis, that is, for every period T we need a finite ε for reaching instability.

Remark 9.8.11 If we give up the assumption on H to be a quadratic form, the flow Φ and consequently also the period mapping $\Phi_{T,0}$ become non-linear. There exists a certain iteration procedure which can be used for calculating the period mapping approximately. Using this idea, a lot of non-linear Hamiltonian systems have been studied in the literature, see e.g. [153, §§ 7.4.2, 7.5.1, 7.5.2].

Exercises

9.8.1 Show that the matrix

$$a = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

defines a linear symplectomorphism on $\mathbb{R}^2$ under which the origin is unstable.

9.8.2 Verify Eqs. (9.8.6) and (9.8.8).

9.8.3 Work out Example 9.8.8.

9.8.4 Consider a mathematical pendulum whose suspension point oscillates vertically according to $a(t) = \frac{1}{\Omega^2}\varepsilon^2 \cos(\Omega t)$, cf. Example 9.8.9. Show that the approximations of the equation of motion for small deviations from the equilibria are equivalent to the Hamilton equations of the Hamiltonian systems (9.8.2) with time-dependent frequency (9.8.10).

9.9 On the Arnold Conjecture

In this section, we turn to the problem of the existence of critical integral curves for a periodically time-dependent Hamiltonian system (M, ω, H). More precisely, we will discuss a special type of critical integral curves which is related to the fixed points of the period mapping $\Phi_{T,0}$. For every fixed point m of $\Phi_{T,0}$, we have $\Phi_{t+T,0}(m) = \Phi_t(m)$ for all t, so that the maximal integral curve through m is either an equilibrium or periodic of period $\frac{T}{k}$ for some positive integer k. A typical example is the motion of a harmonic oscillator acted upon by a periodic external force. Therefore, periodic integral curves whose period is an integer part of that of H are referred to as forced oscillations.

Remark 9.9.1 For convenience, one may restrict attention to time-dependent Hamiltonian systems of period 1. Indeed, if H has period T, the function

$$\tilde{H} : \mathbb{R} \times M \to \mathbb{R}, \qquad \tilde{H}(t, m) := T \cdot H(Tt, m),$$

has period 1 and generates the flow

$$\Phi^{\tilde{H}}_{t_2,t_1} = \Phi^{H}_{Tt_2,Tt_1}.$$

Thus, up to a simultaneous rescaling of the time parameter and the energy scale, one may assume that H has period $T = 1$. In this situation, the period mapping $\Phi_{1,0}$ is usually referred to as the time-1 mapping of the flow.

For what follows, it is important to notice that the period mapping of the flow of a periodic Hamiltonian is a Hamiltonian symplectomorphism, cf. Definition 8.8.7. In this sense, any statement about Hamiltonian symplectomorphisms applies in particular to the period mappings of periodic Hamiltonians.[34] The following conjecture, formulated by Arnold in the 1960s, predicts a universal lower bound on the number of fixed points of a Hamiltonian symplectomorphism and hence on the number of critical integral curves of a periodically time-dependent Hamiltonian system:

Arnold conjecture *Let (M, ω) be a compact symplectic manifold and let $\Psi : M \to M$ be a Hamiltonian symplectomorphism. Then, Ψ must have at least as many fixed points as a function on M must have critical points. If the fixed points are all non-degenerate, then their number is at least equal to the minimal number of critical points of a Morse function on M.*

Remark 9.9.2

1. The Arnold conjecture is certainly true in the time-independent case, that is, for symplectomorphisms Ψ which are given by the flow Φ of an autonomous Hamiltonian function H on M, taken at an arbitrary instant of time, say $t = 1$. Indeed, every critical point of H is an equilibrium of the Hamiltonian vector field X_H and hence a fixed point of Φ_1. Hence, Ψ has at least as many fixed points as H has critical points. If in addition the fixed points of Ψ are non-degenerate, so are the critical points of H. Therefore, H is a Morse function, and Ψ has at least as many fixed points as a Morse function on M must have.
2. The Morse inequalities of Remark 8.9.10/3 imply the following weaker version of the Arnold conjecture. If the fixed points of Ψ are all non-degenerate, their number is greater than or equal to the sum of the Betti numbers of M.

The Arnold conjecture may be viewed as a higher-dimensional generalization of the following theorem, which was formulated by Poincaré in 1912, shortly before he died, and proved by Birkhoff [45][35] in 1913.

Theorem 9.9.3 (Poincaré-Birkhoff Theorem) *Every area preserving homeomorphism of an annulus*

$$A = \left\{(q, p) \in \mathbb{R}^2 : a \leq q^2 + p^2 \leq b\right\}$$

which leaves the boundary circles invariant but twists them in opposite directions possesses at least two fixed points.

---

[34] One can also show the converse, that is, every Hamiltonian symplectomorphisms can be represented as the period mapping of some periodic Hamiltonian, see Exercise 11.8 in [206].

[35] See also [46] for a generalization to ring shaped regions with arbitrary boundary curves.

This theorem is also known as Poincaré's Last Geometric Theorem. Poincaré was led to formulate it when he studied the problem of periodic solutions of the restricted three body problem in celestial mechanics. As mentioned in Sect. 9.6, the Birkhoff-Lewis Theorem 9.6.10 generalizes this theorem in a different direction. We note that both the requirement of being area preserving and the twist condition are important, see Exercise 9.9.1. Following McDuff and Salamon [206], we give a beautiful and simple proof of this theorem under the stronger condition of monotone twist.

Proof Denote the homeomorphism under consideration by φ. In polar coordinates ϕ and r on A, it is given by $\varphi(\phi, r) = (f(\phi, r), h(\phi, r))$, where f and h are continuous functions on A satisfying

$$f(\phi + 2\pi, r) = f(\phi, r) + 2\pi, \qquad h(\phi + 2\pi, r) = h(\phi, r).$$

Under the assumption of monotone twist, that is,

$$r < r' \quad \Rightarrow \quad f(\phi, r) < f\big(\phi, r'\big) \tag{9.9.1}$$

for every angle ϕ there exists a unique radius $r = F(\phi) \in (a, b)$ such that

$$f\big(\phi, F(\phi)\big) = \phi.$$

The mapping F is continuous and 2π-periodic. All points on the curve $\phi \mapsto (\phi, F(\phi))$ in A are mapped radially under φ. Since φ preserves the area, this curve must intersect its image under φ at least twice. Since all points of the intersection are fixed points of φ, this proves the theorem. □

Remark 9.9.4 We draw the attention of the reader to the fact that the above intersection argument is a simple version of the argument used in the last step of the proof of the Birkhoff-Lewis Theorem. In more detail, the counterpart of φ is the iterate $\Psi^{(N)}$ and the counterpart of the annulus A is $\mathrm{T}^n \times B_r$. Thus, to f and to h there correspond $\vartheta_k^{(N)}$ and $I_k^{(N)}$, respectively.

Now, let us prove the Arnold conjecture in the simplest nontrivial case.

Proposition 9.9.5 *The Arnold conjecture holds for Hamiltonian symplectomorphisms of a compact symplectic manifold which are sufficiently close to the identity in the C^1-topology.*

Proof Let (M, ω) be a compact symplectic manifold. By Propositions 8.8.4 and 8.8.10, in the C^1-topology, there exists a neighbourhood of id_M in $\mathrm{Symp}(M, \omega)$ and a local homeomorphism from this neighbourhood onto a neighbourhood of the zero section of T^*M in $Z^1(M)$, mapping Hamiltonian symplectomorphisms φ to exact 1-forms $\mathrm{d}f$ on M. Fixed points of φ thereby correspond to zeros of $\mathrm{d}f$. Thus, fixed points of φ are in one-to-one correspondence with critical points of f. Obviously, they are non-degenerate iff f is a Morse function. □

The Arnold conjecture has attracted the attention of leading mathematicians over the last decades. It was first proved by Eliashberg [85] for Riemann surfaces and next by Conley and Zehnder [66] for tori of arbitrary dimension. Conley and Zehnder showed that the standard torus T^{2n} has at least 2^{2n} distinct fixed points, provided they are all non-degenerate. In the degenerate case there are at least $2n + 1$ fixed points. A breakthrough was made by Floer [93, 94], who developed a new approach to infinite-dimensional Morse theory[36] and applied it to prove the weak Arnold conjecture for so-called monotone symplectic manifolds. Thereafter, Hofer and Salamon [137] generalized this result to the so-called weakly monotone case. Finally, using Floer theory, the weak version of the Arnold conjecture has been proven for arbitrary symplectic manifolds independently by Fukaya and Ono [100] and Liu and Tian [186]. To our knowledge, the strong version is still open.

Exercises

9.9.1 Take the annulus $A = \{(q, p) \in \mathbb{R}^2 : a \le q^2 + p^2 \le b\}$ with polar coordinates ϕ and r and consider the mappings

$$\psi_1(\phi, r) = \left(\phi + \frac{1}{2}, r\right), \qquad \psi_2(\phi, r) = \left(\phi + r - \frac{1}{2}, r^2\right).$$

Find out whether these mappings fulfil the requirements of the Poincaré-Birkhoff Theorem. Do they possess critical points?

[36] Called Floer homology, see [260] or [269] for a review.

Chapter 10
Symmetries

Symmetries play a fundamental role in the study of the dynamics of physical systems, because they give rise to conserved quantities. These can be used to eliminate a number of variables. In the Hamiltonian context, this procedure is called symplectic reduction of systems with symmetries.[1] It is the aim of this chapter to present this procedure in a systematic way. In the Hamiltonian context, the conserved quantities corresponding to the symmetry of the system are encoded in a mapping from the phase space of the system to the dual space of the Lie algebra of the symmetry group. It is called momentum mapping, because it generalizes well-known constants of motion, like momentum or angular momentum etc. In Sect. 10.1 we discuss this notion in detail, including a number of examples. Next, in Sect. 10.2, we present some algebraic basics needed for the symmetry reduction procedure. Then, the classical result of Marsden, Weinstein and Meyer on symmetry reduction is discussed. It states that the reduced phase space, obtained by factorizing a level set of the momentum mapping with respect to the freely acting residual symmetry group, carries a natural symplectic structure and that the dynamics of the systems reduces to this space. In particular, also the relation to orbit reduction is studied. In Sect. 10.4 we present the Symplectic Tubular Neighbourhood Theorem.[2] This is an important technical tool for generalizing the above classical result to the so-called singular case, where the assumption about the free action of the residual symmetry group is removed. The theory of singular reduction is presented in detail in Sect. 10.5. In Sects. 10.6 and 10.7 the reader will find a large number of applications. First, we discuss the following examples from mechanics: the geodesic flow on the three-sphere, the Kepler problem (including the Moser regularization), the Euler top and the spherical pendulum. Section 10.7 contains a model of gauge theory, which can be viewed as obtained from approximation of gauge theory on a finite lattice. Finally, we give an introduction to the study of qualitative dynamics of systems with symmetries in terms of the energy-momentum mapping.

---

[1] In the sequel, we shall cite important contributions to the subject. However, for a quite exhaustive list of references and also for a lot of historical remarks, we refer the reader to [196] and [232].

[2] Also called the Symplectic Slice Theorem.

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*,
Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7_10,

10.1 Momentum Mappings

Recall that an action Ψ of a Lie group G on a symplectic manifold (M, ω) is called symplectic if

$$\Psi_g^* \omega = \omega \tag{10.1.1}$$

for all $g \in G$ and that, in this case, the tuple (M, ω, Ψ) is called a symplectic G-manifold (Definition 8.6.2). Correspondingly, an action[3] ψ of a Lie algebra $\mathfrak{g}$ on (M, ω) is called symplectic if the vector field $\psi(A)$ is symplectic for all $A \in \mathfrak{g}$. Note that the action Ψ of a Lie group G induces an action ψ of its Lie algebra $\mathfrak{g}$ by $\psi(A) = A_*$, with A_* denoting the Killing vector field generated by $A \in \mathfrak{g}$. If Ψ is symplectic, differentiation of (10.1.1) yields

$$\mathscr{L}_{A_*} \omega = 0$$

for all $A \in \mathfrak{g}$, hence ψ is symplectic, too.

Definition 10.1.1 A Hamiltonian system (M, ω, H) is called symmetric

1. under a symplectic action Ψ of a Lie group G if $\Psi_g^* H = H$ for all $g \in G$,
2. under a symplectic action ψ of a Lie algebra $\mathfrak{g}$ if $\psi(A)H = 0$ for all $A \in \mathfrak{g}$.

If a Hamiltonian system is symmetric with respect to a Lie group action, then it is also symmetric with respect to the induced Lie algebra action. Thus, every Lie group symmetry of a given Hamiltonian system induces a Lie algebra symmetry. The converse need not be true, because not every Lie algebra action integrates into a Lie group action.

As we have seen, Killing vector fields of a symplectic G-action are automatically symplectic, but they need not be Hamiltonian. This special case is, however, of particular importance. In this case, for every $A \in \mathfrak{g}$ there exists a function $J_A \in C^\infty(M)$ such that $X_{J_A} = A_*$.

Definition 10.1.2 (Momentum mapping) Let (M, ω) be a symplectic manifold and let Ψ be an action of a Lie group G on M. A mapping $J : M \to \mathfrak{g}^*$ is called a momentum mapping[4] for Ψ if

$$X_{J_A} = A_* \tag{10.1.2}$$

for all $A \in \mathfrak{g}$, where the functions $J_A : M \to \mathbb{R}$ are given by

$$J_A(m) := \big\langle J(m), A \big\rangle. \tag{10.1.3}$$

The action Ψ is called Hamiltonian if it is symplectic and if there exists a momentum mapping. In this case, the tuple (M, ω, Ψ) is called a Hamiltonian G-manifold.[5]

[3] See Definition 6.2.6.

[4] The notion of momentum mapping has a long history, see [309] and [194].

[5] If a momentum mapping J is fixed, we may include it in the tuple, thus writing (M, ω, Ψ, J).

Since ω is non-degenerate, Formula (10.1.2) is equivalent to

$$A_* \lrcorner \omega = -\mathrm{d}J_A. \tag{10.1.4}$$

Proposition 10.1.3 *Let (M,ω) be a symplectic manifold and let Ψ be an action of a Lie group G on M. A momentum mapping for Ψ exists iff all Killing vector fields are Hamiltonian.*

Proof If every Killing vector field is Hamiltonian, then, for a chosen basis $\{e_i\}$ in $\mathfrak{g}$, there exist smooth functions $f_i \in C^\infty(M)$ such that $(e_i)_* = X_{f_i}$. Using the dual basis $\{e^{*i}\}$, we define

$$J : M \to \mathfrak{g}^*, \quad J(m) := f_i(m)e^{*i}.$$

Then, for every $A \in \mathfrak{g}$, we obtain $J_A = \langle e^{*i}, A\rangle\, f_i$ and hence

$$X_{J_A} = X_{\langle e^{*i},A\rangle f_i} = \big\langle e^{*i}, A\big\rangle X_{f_i} = \big\langle e^{*i}, A\big\rangle (e_i)_* = \big(\big\langle e^{*i}, A\big\rangle e_i\big)_* = A_*.$$

Thus, J is a momentum mapping for Ψ. The converse direction is obvious. □

Remark 10.1.4

1. We see that a momentum mapping exists if only condition (10.1.2) is fulfilled for each $A \in \mathfrak{g}$ separately, with an arbitrary smooth function J_A. As shown in the proof of Proposition 10.1.3, the functions J_A can always be chosen so that the mapping $A \mapsto J_A$ is linear.
2. Let (M,ω,Ψ) be a Hamiltonian G-manifold and let J and $\tilde{J}$ be momentum mappings. By (10.1.4), we have
$$\mathrm{d}(J_A - \tilde{J}_A) = 0$$
for all $A \in \mathfrak{g}$, that is, $J_A - \tilde{J}_A$ is constant on every connected component of M. Thus, if M is connected, the momentum mapping is fixed up to an additive constant $\mu_0 \in \mathfrak{g}^*$,
$$J = \tilde{J} + \mu_0.$$
3. Definition 10.1.2 uses only the Lie algebra $\mathfrak{g}$ and its dual vector space. Therefore, it extends automatically to the case of the action of a Lie algebra. If there exists a momentum mapping, the action is symplectic. In this case, the tuple (M,ω,ψ) is called a Hamiltonian $\mathfrak{g}$-manifold. In this book we have in mind Lie algebras of finite-dimensional Lie groups only. However, these notions can be extended to infinite-dimensional Lie algebras as well.

Proposition 10.1.5 *Let (M,ω) be a symplectic manifold and let Ψ be an action of a Lie group G on M. If G is connected and if there exists a momentum mapping for Ψ, then Ψ is symplectic (and hence Hamiltonian).*

Proof If there exists a momentum mapping J, for every $A \in \mathfrak{g}$ we have

$$\mathscr{L}_{A_*}\omega = A_* \lrcorner \mathrm{d}\omega + \mathrm{d}(A_* \lrcorner \omega) = -\mathrm{d}(\mathrm{d}J_A) = 0$$

and hence, by (4.1.28),

$$\frac{\mathrm{d}}{\mathrm{d}t}\Psi^*_{\exp(tA)}\omega = \Psi^*_{\exp(tA)}\mathscr{L}_{A_*}\omega = 0.$$

It follows that $\Psi^*_{\exp(tA)}\omega = \omega$ for all t. Since G is connected, by Proposition 5.1.7, it is generated by a neighbourhood of the identity. Thus, $\Psi^*_g\omega = \omega$ for all $g \in G$. □

Remark 10.1.6 As a consequence of Propositions 10.1.3 and 10.1.5, for an action of a connected Lie group on a symplectic manifold to be Hamiltonian it is sufficient that all Killing vector fields be Hamiltonian. In particular, the condition that the action be symplectic is automatically fulfilled then.

We note that for arbitrary symplectic G-manifolds a momentum mapping need not exist, because the Killing vector fields of a symplectic group action need not be Hamiltonian, see Remark 8.2.7. The following proposition provides conditions on M and G under which a momentum mapping exists.

Proposition 10.1.7 *Let (M, ω, Ψ) be a symplectic G-manifold. A momentum mapping for Ψ exists iff the linear mapping*

$$F: \mathfrak{g}/[\mathfrak{g},\mathfrak{g}] \to H^1(M,\mathbb{R}), \qquad F\big([A]\big) := [A_* \lrcorner\, \omega] \tag{10.1.5}$$

vanishes identically.

Proof We give the proof for a left action. The mapping F is well defined, that is, if $A = [B, C]$ for some $B, C \in \mathfrak{g}$, then $A_* \lrcorner\, \omega$ is exact: since Ψ is symplectic, we have

$$\mathscr{L}_{B_*}\omega = 0 = \mathscr{L}_{C_*}\omega,$$

that is, B_* and C_* are locally Hamiltonian. According to Proposition 8.2.6/2, then $[C_*, B_*] = X_{\omega(C_*,B_*)}$. Thus, using Proposition 6.2.2/2, we obtain

$$A_* \lrcorner\, \omega = [B, C]_* \lrcorner\, \omega = [C_*, B_*] \lrcorner\, \omega = -\mathrm{d}\big(\omega(C_*, B_*)\big).$$

The linearity of F is obvious. Now, according to Proposition 10.1.3 and Formula (10.1.4), a momentum mapping J exists iff $A_* \lrcorner \omega$ is exact for all $A \in \mathfrak{g}$, that is, iff $[A_* \lrcorner\, \omega] = F([A]) = 0$. □

Corollary 10.1.8 *Let (M, ω, Ψ) be a symplectic G-manifold. In each of the following two cases, a momentum mapping for Ψ exists.*

1. *The first de Rham-cohomology group $H^1(M,\mathbb{R})$ is trivial.*
2. *There holds $\mathfrak{g} = [\mathfrak{g},\mathfrak{g}]$. This means that the first cohomology group of the Lie algebra $\mathfrak{g}$ is trivial.*

The most important property of momentum mappings is that they provide constants of motion. Recall that a function $f \in C^\infty(M)$ is called G-invariant if $f \circ \Psi_g = f$ for all $g \in G$ and that the subspace of G-invariant functions is denoted by $C^\infty(M)^G$.

Theorem 10.1.9 (Noether) *Let (M, ω, Ψ, J) be a Hamiltonian G-manifold and let $H \in C^\infty(M)^G$. For $A \in \mathfrak{g}$, the function J_A is a constant of motion for the Hamiltonian system (M, ω, H), that is,*

$$\frac{\mathrm{d}}{\mathrm{d}t} J_A\big(\gamma(t)\big) = 0$$

for all integral curves γ of the Hamiltonian vector field X_H.

Proof For all $m \in M$, we have

$$\{J_A, H\}(m) = X_{J_A}(H)(m) = A_*(H)(m) = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} H \circ \Psi_{\exp tA}(m) = 0.$$

Hence, the assertion follows from Proposition 9.1.10/1. □

Corollary 10.1.10 *If the action is free, for every basis $\{e_i\}$ in $\mathfrak{g}$, the functions J_{e_i} constitute a linearly independent system of constants of motion.*

Next, we study the transformation properties of momentum mappings with respect to the action of G on M and the coadjoint action of G on $\mathfrak{g}^*$. For that purpose, let (M, ω, Ψ, J) be a left Hamiltonian G-manifold and let $A \in \mathfrak{g}$ and $g \in G$. Using Proposition 8.2.9 and point 1 of Proposition 6.2.2, we obtain

$$X_{J_A \circ \Psi_g} = (\Psi_{g^{-1}})_* X_{J_A} = (\Psi_{g^{-1}})_* A_* = \big(\mathrm{Ad}\big(g^{-1}\big)A\big)_* = X_{J_{\mathrm{Ad}(g^{-1})A}}$$

and hence

$$0 = (X_{J_A \circ \Psi_g} - X_{J_{\mathrm{Ad}(g^{-1})A}}) \lrcorner\, \omega = -\mathrm{d}(J_A \circ \Psi_g - J_{\mathrm{Ad}(g^{-1})A}).$$

Thus, for all $A \in \mathfrak{g}$ the function $J_A \circ \Psi_g - J_{\mathrm{Ad}(g^{-1})A}$ is constant on every connected component of M. Using

$$J_{\mathrm{Ad}(g^{-1})A}(m) = \big\langle J(m), \mathrm{Ad}\big(g^{-1}\big)A\big\rangle = \big\langle \mathrm{Ad}^*(g) \circ J(m), A\big\rangle,$$

we can rewrite this function in the form

$$m \mapsto \big\langle J \circ \Psi_g(m) - \mathrm{Ad}^*(g) \circ J(m), A\big\rangle.$$

Thus, if M is connected, the linear functional $J \circ \Psi_g(m) - \mathrm{Ad}^*(g) \circ J(m)$ on $\mathfrak{g}$ does not depend on m and we obtain a mapping

$$\sigma : G \to \mathfrak{g}^*, \qquad \sigma(g) := J \circ \Psi_g(m) - \mathrm{Ad}^*(g) \circ J(m), \tag{10.1.6}$$

where m is an arbitrarily chosen point of M. Obviously, $\sigma = 0$ iff J is equivariant with respect to the coadjoint action of G, that is, iff

$$J \circ \Psi_g = \mathrm{Ad}^*(g) \circ J. \tag{10.1.7}$$

In this case, $G_m \subset G_{J(m)}$ and Proposition 6.2.4/2 implies

$$J' \circ A_* = A_*^{\mathrm{Ad}^*} \circ J, \quad A \in \mathfrak{g}, \tag{10.1.8}$$

where $A_*^{\mathrm{Ad}^*}$ is the Killing vector field generated by A under the coadjoint action.

Remark 10.1.11 The mapping σ defined by (10.1.6) is a measure for the momentum mapping to fail the equivariance property. It satisfies

$$\sigma(gh) = \sigma(g) + \mathrm{Ad}^*(g)\sigma(h). \tag{10.1.9}$$

A mapping $\sigma : G \to \mathfrak{g}^*$ with this property is called a coadjoint 1-cocycle with values in $\mathfrak{g}^*$. A cocycle τ is called a coboundary if there exists an element $\mu \in \mathfrak{g}^*$ such that

$$\tau(g) = \mu - \mathrm{Ad}^*(g)\mu. \tag{10.1.10}$$

The set of $\mathfrak{g}^*$-valued 1-cocycles carries a vector space structure and the coboundaries form a vector subspace. The quotient vector space $H^1(G, \mathfrak{g}^*)$ is called the first cohomology group of G with values in $\mathfrak{g}^*$.

Proposition 10.1.12 *For a connected Hamiltonian G-manifold (M, ω, Ψ), the class $[\sigma] \in H^1(G, \mathfrak{g}^*)$, defined by (10.1.6), does not depend on the momentum mapping J. An equivariant momentum mapping exists iff $[\sigma] = 0$.*

Proof Let J, $\tilde{J}$ be momentum mappings for Ψ and let σ and $\tilde{\sigma}$ denote the corresponding coadjoint 1-cocycles. By Remark 10.1.4/2, we have $J = \tilde{J} + \mu_0$ for some $\mu_0 \in \mathfrak{g}^*$. Thus,

$$\sigma(g) = (\tilde{J} + \mu_0) \circ \Psi_g(m) - \mathrm{Ad}^*(g) \circ (\tilde{J} + \mu_0)(m) = \tilde{\sigma}(g) + \mu_0 - \mathrm{Ad}^*(g)\mu_0$$

and hence $[\sigma] = [\tilde{\sigma}]$. Moreover, if an equivariant momentum mapping exists, then obviously $[\sigma] = 0$. Conversely, if $[\sigma] = 0$, there exists an element $\mu \in \mathfrak{g}^*$ such that $\sigma(g) = \mu - \mathrm{Ad}^*(g)\mu$. Then, $\tilde{J} = J - \mu$ is an equivariant momentum mapping. □

Remark 10.1.13 If (M, ω, Ψ) is a connected Hamiltonian G-manifold with $[\sigma] \neq 0$, a given momentum mapping J can be made equivariant by the following modification of the coadjoint action on $\mathfrak{g}^*$:

$$\Phi : G \times \mathfrak{g}^* \to \mathfrak{g}^*, \qquad \Phi(g, \mu) := \mathrm{Ad}^*(g)\mu + \sigma(g). \tag{10.1.11}$$

Using (10.1.9), one can check that Φ is a left action. Moreover, by construction, we have $J \circ \Psi_g = \Phi_g \circ J$, that is, J is equivariant with respect to Φ.

Proposition 10.1.14 *Let (M, ω, Ψ, J) be a left Hamiltonian G-manifold. If J is equivariant, for all $A, B \in \mathfrak{g}$ one has*

$$\{J_A, J_B\} = J_{[B,A]}, \tag{10.1.12}$$

that is, J is an anti-homomorphism[6] *of the Lie algebras $\mathfrak{g}$ and $(C^\infty(M), \{\,,\,\})$.*

Proof By equivariance of J, we have

$$\langle J \circ \Psi_{\exp tA}(m), B\rangle = \langle \mathrm{Ad}^*(\exp tA) \circ J(m), B\rangle.$$

[6] If Ψ is a right action, J is a homomorphism.

Differentiating this equation with respect to t at $t=0$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \big\langle J\circ\Psi_{\exp tA}(m), B\big\rangle = A_*(J_B)(m) = X_{J_A}(J_B)(m) = \{J_A, J_B\}(m)$$

for the left hand side and

$$\frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \big\langle \mathrm{Ad}^*(\exp tA)\circ J(m), B\big\rangle = \big\langle J(m), [B,A]\big\rangle = J_{[B,A]}(m)$$

for the right-hand side. □

Corollary 10.1.15 *Let (M,ω,Ψ,J) be a left Hamiltonian G-manifold and let J be equivariant. Let $m\in M$ and $J(m)=\mu$, let $\iota: G\cdot m\to M$ be the natural inclusion mapping and denote the coadjoint orbit through μ by $\mathscr{O}_\mu\subset\mathfrak{g}^*$. Then,*[7]

$$\iota^*\omega = J^*\omega^{\mathscr{O}_\mu^-}, \tag{10.1.13}$$

where $\omega^{\mathscr{O}_\mu^-}$ denotes the negative Kirillov form on $\mathscr{O}_\mu$, see Theorem 8.4.1.

For a right action, one obtains the positive Kirillov form on $\mathscr{O}_\mu$.

Proof Since ω and $\omega^{\mathscr{O}_\mu^-}$ are G-invariant, it is enough to prove (10.1.13) at m. On the one hand, using (10.1.12) and the fact that $\mathrm{T}_m(G\cdot m)$ is spanned by Killing vector fields, we get

$$\omega_m(A_*,B_*) = \omega_m(X_{J_A}, X_{J_B}) = \{J_A,J_B\}(m) = J_{[B,A]}(m) = -\big\langle\mu,[A,B]\big\rangle$$

for all $A,B\in\mathfrak{g}$. On the other hand, the equivariance property (10.1.7) yields

$$J\circ\Psi_m(g) = J\circ\Psi_g(m) = \mathrm{Ad}^*(g)\circ J(m) = \mathrm{Ad}^*(g)\mu,$$

that is,

$$J'_m\circ\Psi'_m(A) = \mathrm{ad}^*(A)\mu,$$

where $\mathrm{ad}^*(A)\mu$ is the value of the Killing vector field $A_*^{\mathrm{Ad}^*}$ at μ. Now, we have

$$\begin{aligned}\big(J^*\omega^{\mathscr{O}_\mu^-}\big)_m(A_*,B_*) &= \big(\omega^{\mathscr{O}_\mu^-}\big)_\mu\big(J'_mA_*, J'_mB_*\big)\\ &= \big(\omega^{\mathscr{O}_\mu^-}\big)_\mu\big(A_*^{\mathrm{Ad}^*}, B_*^{\mathrm{Ad}^*}\big)\\ &= -\big\langle\mu,[A,B]\big\rangle \end{aligned}\tag{10.1.14}$$

for all $A,B\in\mathfrak{g}$. □

Definition 10.1.16 A Hamiltonian G-manifold ($\mathfrak{g}$-manifold) is said to be strongly Hamiltonian if there exists a momentum mapping satisfying (10.1.12).

By Proposition 10.1.14, every Hamiltonian G-manifold which admits an equivariant momentum mapping is strongly Hamiltonian.

[7] Since $\mathscr{O}_\mu$ is an initial submanifold of $\mathfrak{g}^*$, J restricts to a smooth mapping $J: G\cdot m\to\mathscr{O}_\mu$, denoted by the same symbol.

For a connected Hamiltonian G-manifold (M, ω, Ψ, J), we can derive a criterion for J to satisfy (10.1.12) in terms of the mapping $\sigma : G \to \mathfrak{g}^*$ defined by (10.1.6). For that purpose, we determine the tangent mapping $\sigma_e' \colon \mathrm{T}_e G \cong \mathfrak{g} \to \mathrm{T}_0\mathfrak{g}^* \cong \mathfrak{g}^*$. By the calculation in the proof of Proposition 10.1.14, we obtain

$$\big\langle \sigma_e'(A), B \big\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \big\langle \sigma(\exp tA), B \big\rangle = \{J_A, J_B\}(m) - J_{[B,A]}(m). \tag{10.1.15}$$

This means that J satisfies (10.1.12) iff σ_e' vanishes identically, that is, iff J is infinitesimally equivariant.

Remark 10.1.17 Choose a point $m \in M$ and define the following antisymmetric bilinear form Σ on $\mathfrak{g}$:

$$\Sigma(A, B) := \{J_A, J_B\}(m). \tag{10.1.16}$$

Using $\Sigma(A, B) = \omega_m(A_*, B_*)$, Proposition 4.1.6 and point 2 of Proposition 8.2.6, one can check that

$$\Sigma\big([A, B], C\big) + \Sigma\big([B, C], A\big) + \Sigma\big([C, A], B\big) = 0 \tag{10.1.17}$$

for arbitrary $A, B, C \in \mathfrak{g}$ (Exercise 10.1.1). An antisymmetric bilinear form on $\mathfrak{g}$ with this property is called a coadjoint 2-cocycle on $\mathfrak{g}$. By the Jacobi identity, every $\mu \in \mathfrak{g}^*$ defines a 2-cocycle $\delta\mu$ by

$$\delta\mu(A, B) := \mu\big([A, B]\big). \tag{10.1.18}$$

A 2-cocycle of this form is called a 2-coboundary.[8] The 2-cocycles form a vector space and the 2-coboundaries form a vector subspace. The quotient vector space is called the second cohomology group of the Lie algebra $\mathfrak{g}$ and is denoted by $\mathscr{H}^2(\mathfrak{g})$. One can show that another choice of the point m in (10.1.16) yields an equivalent cocycle (Exercise 10.1.1). From (10.1.15) we read off

$$\big\langle \sigma_e'(A), B \big\rangle = \Sigma(A, B) + \delta\big(J(m)\big)(A, B), \tag{10.1.19}$$

that is, $[\Sigma] = [\sigma_e']$ as elements of $\mathscr{H}^2(\mathfrak{g})$.

The above discussion yields (Exercise 10.1.1)

Proposition 10.1.18 *For a connected Hamiltonian G-manifold (M, ω, Ψ), the class $[\Sigma] \in \mathscr{H}^2(g)$, defined by (10.1.16), does not depend on the momentum mapping J. The action Ψ is strongly Hamiltonian iff $[\Sigma] = 0$.*

Remark 10.1.19

1. Let (M, ω, Ψ) be a symplectic G-manifold. If $\mathfrak{g}$ is semisimple, then

$$[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g} \quad \text{and} \quad \mathscr{H}^2(\mathfrak{g}) = 0,$$

[8] And δ is called the coboundary operator.

see [314] and [315]. According to Corollary 10.1.8, the property $[\mathfrak{g},\mathfrak{g}]=\mathfrak{g}$ implies that a momentum mapping exists. The property $\mathscr{H}^2(\mathfrak{g})=0$ ensures that (M,ω,Ψ) is strongly Hamiltonian.
2. Let (M,ω,Ψ) be a Hamiltonian G-manifold. If $[\sigma]$ vanishes, there exists an equivariant momentum mapping. By Proposition 10.1.14, then (M,ω,Ψ) is strongly Hamiltonian, so that $[\Sigma]$ must vanish, too. Conversely, if $[\Sigma]=0$, there exists a momentum mapping J satisfying (10.1.12). Then, (10.1.15) yields $\sigma'_e=0$. Hence, by plugging $h=\exp(tA)$ into (10.1.9) and differentiating at $t=0$, we obtain
$$\sigma'_g\circ L'_g=\mathrm{Ad}^*(g)\circ\sigma'_e=0.$$
We conclude that σ is constant on each connected component of G. Thus, if G is connected, then $[\sigma]=0$. To summarize, if G is connected, there exists an equivariant momentum mapping iff Ψ is strongly Hamiltonian.

In what follows we discuss Hamiltonian G-manifolds admitting an invariant symplectic potential.

Proposition 10.1.20 *Let (M,ω,Ψ) be a left symplectic G-manifold. If $\omega=\mathrm{d}\theta$ for some Ψ-invariant 1-form θ, then*
$$J\colon M\to\mathfrak{g}^*,\qquad \big\langle J(m),A\big\rangle:=\theta_m(A_*),\tag{10.1.20}$$
is an equivariant momentum mapping.

Proof Obviously, for every $m\in M$, $J(m)$ is a linear functional on $\mathfrak{g}$. G-invariance of θ implies
$$0=\mathscr{L}_{A_*}\theta=A_*\lrcorner\,\mathrm{d}\theta+\mathrm{d}(A_*\lrcorner\,\theta)$$
and hence $A_*\lrcorner\,\omega=-\mathrm{d}J_A$ for all $A\in\mathfrak{g}$. Equivariance follows from
$$\theta_m(A_*)=\big(\Psi_g^*\theta\big)_m(A_*)=\theta_{\Psi_g(m)}(\Psi_{g*}A_*)=\theta_{\Psi_g(m)}\Big(\big(\mathrm{Ad}\big(g^{-1}\big)A\big)_*\Big),$$
where we have used Proposition 6.2.2/1. □

An important class of Hamiltonian G-manifolds which admit an invariant symplectic potential is constituted by the cotangent bundles of G-manifolds. Let (Q,ψ) be a left G-manifold and let $\pi:\mathrm{T}^*Q\to Q$ denote the canonical projection. According to Example 6.1.2/5, for every $g\in G$, the diffeomorphism $\psi_g:Q\to Q$ induces a point transformation $\Psi_g:\mathrm{T}^*Q\to\mathrm{T}^*Q$ by
$$\big\langle\Psi_g(\xi),X\big\rangle=\big\langle\xi,\psi'_{g^{-1}}X\big\rangle,\quad \xi\in\mathrm{T}^*Q,\ X\in\mathrm{T}_{\psi_g\circ\pi(\xi)}Q,\tag{10.1.21}$$
see also (8.3.9) and (8.3.10). The assignment $g\mapsto\Psi_g$ defines a mapping
$$\Psi:G\times\mathrm{T}^*Q\to\mathrm{T}^*Q,$$
which is a left G-action. By construction, the mappings Ψ_g are vector bundle automorphisms covering the diffeomorphisms ψ_g,
$$\pi\circ\Psi_g=\psi_g\circ\pi.\tag{10.1.22}$$

In particular, the canonical projection π is equivariant. An action Ψ on T^*Q with the property (10.1.22) is called a lift of the action ψ to T^*Q.

By Proposition 8.3.6, the canonical 1-form θ on T^*Q, and hence the natural symplectic form $\omega = \mathrm{d}\theta$, is invariant under Ψ. Thus, $(\mathrm{T}^*Q, \omega, \Psi)$ is a symplectic G-manifold and the assumptions of Proposition 10.1.20 are fulfilled. This yields

Corollary 10.1.21 *In the case of the symplectic G-manifold $(\mathrm{T}^*Q, \omega, \Psi)$ associated with the G-manifold (Q, ψ), the equivariant momentum mapping defined by (10.1.20) is given by*

$$J : \mathrm{T}^*Q \to \mathfrak{g}^*, \qquad \big\langle J(\xi), A\big\rangle = \big\langle \xi, A_*^{\psi}\big(\pi(\xi)\big)\big\rangle, \tag{10.1.23}$$

where A_*^{ψ} is the Killing vector field generated by A under the action ψ on Q.

Proof For $A \in \mathfrak{g}$, let A_*^{Ψ} denote the Killing vector field generated by A under Ψ. Since π is equivariant, Proposition 6.2.4/2 yields

$$\big\langle J(\xi), A\big\rangle = \theta_\xi\big(A_*^{\Psi}\big) = \big\langle \xi, \pi'\big(A_*^{\Psi}(\xi)\big)\big\rangle = \big\langle \xi, A_*^{\psi}\big(\pi(\xi)\big)\big\rangle. \qquad \square$$

Example 10.1.22 (Momentum) Let $Q = \mathbb{R}^3$ and let $G = \mathbb{R}^3$ act by translations:

$$\psi_{\mathbf{a}}(\mathbf{x}) := \mathbf{x} + \mathbf{a}.$$

Under the identification $\mathrm{T}Q \cong \mathbb{R}^3 \times \mathbb{R}^3$, the tangent mapping of $\psi_{\mathbf{a}}$ is given by $\psi'_{\mathbf{a}}(\mathbf{x}, \mathbf{y}) = (\mathbf{x} + \mathbf{a}, \mathbf{y})$. Hence, the lift Ψ to $\mathrm{T}^*Q \cong \mathbb{R}^3 \times \mathbb{R}^3$ is given by

$$\Psi_{\mathbf{a}}(\mathbf{x}, \mathbf{p}) = (\mathbf{x} + \mathbf{a}, \mathbf{p}).$$

Using Example 5.3.15, we compute the Killing vector field generated by $\mathbf{b} \in \mathfrak{g} \cong \mathbb{R}^3$:

$$\mathbf{b}_*^{\psi}(\mathbf{x}) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 (\mathbf{x} + t\mathbf{b}) = \mathbf{b}.$$

Consequently, for the momentum mapping (10.1.23) we obtain

$$\big\langle J\big((\mathbf{x}, \mathbf{p})\big), \mathbf{b}\big\rangle = \mathbf{p} \cdot \mathbf{b},$$

that is,

$$J\big((\mathbf{x}, \mathbf{p})\big) = \mathbf{p}.$$

Thus, J coincides with momentum. This explains the origin of the name momentum mapping.

Example 10.1.23 (Angular momentum) Let $Q = \mathbb{R}^3$ and let $G = \mathrm{SO}(3)$ act by rotations:

$$\psi_g \mathbf{x} = g\mathbf{x}.$$

Since Ψ leaves invariant the Euclidean scalar product, the lift Ψ is given by

$$\Psi_g(\mathbf{x}, \mathbf{p}) = (g\mathbf{x}, g\mathbf{p}).$$

By Example 6.2.5/1, the Killing vector field generated by $A \in \mathfrak{g} = \mathfrak{so}(3)$ is

$$A_*^{\psi}(\mathbf{x}) = A\mathbf{x}.$$

Hence, for the momentum mapping (10.1.23) we obtain

$$\langle J(\mathbf{x}, \mathbf{p}), A \rangle = \mathbf{p} \cdot A\mathbf{x}.$$

Using

$$\mathbf{p} \cdot A\mathbf{x} = (\mathbf{x} \times \mathbf{p}) \cdot \mathbf{A}, \tag{10.1.24}$$

where $\times$ denotes the vector product and $\mathbf{A} \in \mathbb{R}^3$ denotes the vector corresponding to A via the isomorphism (5.2.6), we obtain

$$J(\mathbf{x}, \mathbf{p}) = \mathbf{x} \times \mathbf{p},$$

that is, J coincides with angular momentum.

Example 10.1.24 (Translations on a Lie group) Consider $Q = G$ with G acting by left translation:

$$\psi : G \times G \to G, \qquad \psi(g, a) := ga,$$

that is, $\psi_g = \mathrm{L}_g$. The action induced on T^*G is given by

$$\Psi : G \times \mathrm{T}^*G \to \mathrm{T}^*G, \qquad \Psi_g(\xi) = (\mathrm{L}_{g^{-1}})'^{\mathrm{T}}(\xi),$$

and the associated equivariant momentum mapping (10.1.23) is given by

$$J : \mathrm{T}^*G \to \mathfrak{g}^*, \qquad \langle J(\xi), A \rangle = \langle \xi, \big(A_*^{\mathrm{L}}\big)_{\pi(\xi)} \rangle.$$

Here, A_*^{L} denotes the Killing vector field of A with respect to the action by left translation. According to Example 6.2.5/2, A_*^{L} coincides with the right-invariant vector field generated by A. Hence, $A_*^{\mathrm{L}}(a) = (\mathrm{R}_a)'_e(A)$ and we obtain

$$J(\xi) = (\mathrm{R}_{\pi(\xi)})'^{\mathrm{T}}(\xi). \tag{10.1.25}$$

In the trivialization $\chi : G \times \mathfrak{g}^* \to \mathrm{T}^*G$ induced by left translation, cf. (8.3.6), Ψ is given by the action

$$\mathscr{L} := \chi^{-1} \circ \Psi \circ \chi : G \times \big(G \times \mathfrak{g}^*\big) \to \big(G \times \mathfrak{g}^*\big), \qquad \mathscr{L}_g(a, \nu) = (ga, \nu), \tag{10.1.26}$$

and the momentum mapping $J^{\mathscr{L}} := J \circ \chi$ has the form

$$J^{\mathscr{L}}(a, \nu) = \mathrm{Ad}^*(a)\nu. \tag{10.1.27}$$

We encourage the reader to check that for $G = \mathbb{R}^3$, this example boils down to Example 10.1.22.

Analogously, one deals with the right action induced by right translation on G. Here, the momentum mapping is given by

$$J(\xi) = (\mathrm{L}_{\pi(\xi)})'^{\mathrm{T}}(\xi). \tag{10.1.28}$$

The corresponding left action is induced by right translation with the inverse group element,

$$\psi: G \times G \to G, \qquad \psi(g, a) := ag^{-1} \equiv \mathrm{R}_{g^{-1}}(a).$$

It has the momentum mapping

$$J(\xi) = -(\mathrm{L}_{\pi(\xi)})'^{\mathrm{T}}(\xi). \tag{10.1.29}$$

In the trivialization χ, this action is given by

$$\mathscr{R}: G \times \big(G \times \mathfrak{g}^*\big) \to \big(G \times \mathfrak{g}^*\big), \qquad \mathscr{R}_g(a, \nu) = \big(ag^{-1}, \mathrm{Ad}^*(g)\nu\big), \tag{10.1.30}$$

and the momentum mapping (10.1.29) takes the form

$$J^{\mathscr{R}}(a, \nu) = -\nu. \tag{10.1.31}$$

Example 10.1.25 (Inner automorphisms of a Lie group) Consider $Q = G$ with G acting by inner automorphisms:

$$\psi: G \times G \to G, \qquad \psi(g, a) := gag^{-1},$$

that is, $\psi_g = \mathrm{C}_g$. The Killing vector field generated by $A \in \mathfrak{g}$ is

$$A_*^{\mathrm{C}}(g) = \mathrm{R}'_g(A) - \mathrm{L}'_g(A)$$

and the induced action on T^*G reads

$$\Psi : G \times \mathrm{T}^*G \to \mathrm{T}^*G, \qquad \Psi_g(\xi) = (\mathrm{C}_{g^{-1}})'^{\mathrm{T}}(\xi).$$

For the equivariant momentum mapping (10.1.23), we obtain

$$\big\langle J(\xi), A\big\rangle = \big\langle (\mathrm{R}_{\pi(\xi)})'^{\mathrm{T}}(\xi) - (\mathrm{L}_{\pi(\xi)})'^{\mathrm{T}}(\xi), A\big\rangle, \quad \xi \in \mathrm{T}_g^*G,$$

that is,

$$J(\xi) = \mu - \mathrm{Ad}^*\big(\pi(\xi)\big)\mu, \tag{10.1.32}$$

with $\mu = (\mathrm{R}_{\pi(\xi)})'^{\mathrm{T}}(\xi)$.

Example 10.1.26 (Coadjoint orbits) Let $\mathfrak{g}$ be a Lie algebra and let $\mathscr{O} \subset \mathfrak{g}^*$ be a coadjoint orbit. According to Theorem 8.4.1, $\mathscr{O}$ endowed with the Kirillov form

$$\omega^{\mathscr{O}}(A_*, B_*)(\mu) := \big\langle \mu, [A, B]\big\rangle, \quad \mu \in \mathscr{O},$$

is a symplectic G-manifold. Every $A \in \mathfrak{g}$ defines a linear function

$$J_A: \mathfrak{g}^* \to \mathbb{R}, \qquad J_A(\mu) := -\langle \mu, A\rangle.$$

For $B \in \mathfrak{g}$, we find

$$(B_*)_\mu(J_A) = -\frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \big\langle \mathrm{Ad}^*\big(\exp(tB)\big)\mu, A\big\rangle = -\big\langle \mu, [A, B]\big\rangle = -\omega^{\mathscr{O}}_\mu(A_*, B_*).$$

Hence, $A_* \lrcorner\, \omega^{\mathscr{O}} = -\mathrm{d}J_A$. This means that the mapping

$$J: \mathscr{O} \to \mathfrak{g}^*, \qquad J(\mu) := -\mu, \tag{10.1.33}$$

is a momentum mapping. It is obviously equivariant.

Exercises

10.1.1 Show that the cocycles on $\mathfrak{g}$ defined in (10.1.16) by the help of different points differ by a coboundary. Prove Formula (10.1.17) and Proposition 10.1.18.

10.1.2 Let (V, ω) be a symplectic vector space and let $G \subset \mathrm{Sp}(V, \omega)$ be a closed subgroup. Show that

$$J: V \to \mathfrak{g}^*, \qquad \langle J(v), A \rangle := -\frac{1}{2}\omega(Av, v), \tag{10.1.34}$$

is an equivariant momentum mapping for the action of G on (V, ω). (This applies in particular to the isotropy representation at an arbitrary point of a symplectic G-manifold.)

10.1.3 Let (M, ω, Ψ) be a Hamiltonian G-manifold with equivariant momentum mapping J. Let $\mathscr{O} \subset \mathfrak{g}^*$ be a coadjoint orbit, endowed with the Kirillov form $\omega^{\mathscr{O}}$. Consider the direct product of G-manifolds $M \times \mathscr{O}$ and let pr_M and $\mathrm{pr}_{\mathscr{O}}$ denote the natural projections to the factors. Show that

$$\tilde{\omega} = \mathrm{pr}_1^* \omega + \mathrm{pr}_2^* \omega^{\mathscr{O}} \tag{10.1.35}$$

is a G-invariant symplectic form on $M \times \mathscr{O}$ and that

$$K: M \times \mathscr{O} \to \mathfrak{g}^*, \qquad K(m, \mu) := J(m) - \mu, \tag{10.1.36}$$

is an equivariant momentum mapping. (This momentum mapping is used for the so-called shifting trick, which will be explained in Remark 10.3.9.)

10.1.4 Let G be a Lie group and let $H \subset G$ be a closed subgroup. Consider the action of H on G by left translation. Determine the induced action on T^*G and the corresponding equivariant momentum mapping (10.1.23).

10.1.5 Prove Formula (10.1.24).

10.1.6 Let $M = \mathbb{R}^2$ and consider the action of $G = \mathbb{R}^2$ on M by translations,

$$\Psi : G \times M \to M, \qquad \Psi(\mathbf{a}, \mathbf{x}) := \mathbf{x} + \mathbf{a}.$$

Show that

$$J: M \to \mathfrak{g}^* \cong \mathbb{R}^2, \qquad J(\mathbf{x}) \cdot \mathbf{A} := A_1 x_2 - A_2 x_1,$$

is a momentum mapping. Calculate $[\sigma]$ and the modified G-action (10.1.11) on $\mathfrak{g}^*$ for which J is equivariant.

10.1.7 Show that the phase space of the n-dimensional isotropic harmonic oscillator admits a symplectic action of the unitary group $\mathrm{U}(n)$ which leaves the Hamiltonian

$$H(q, p) = \frac{1}{2}p^2 + \frac{\omega^2}{2}q^2$$

invariant. Find a momentum mapping.
Hint. Identify $\mathrm{T}^*\mathbb{R}^n$ with $\mathbb{C}^n$ via $(q, p) \mapsto aq + \mathrm{i}bp$ with appropriately chosen constants $a, b \in \mathbb{R}$.

10.2 The Witt-Artin Decomposition

Let (M, ω, Ψ, J) be a Hamiltonian G-manifold. From the Noether Theorem 10.1.9 we know that J is constant along the integral curves of the Hamiltonian vector field generated by a G-invariant function H. Thus, the level sets

$$M_\mu := J^{-1}(\mu) \subset M, \quad \mu \in \mathfrak{g}^*,$$

are invariant under the flow of X_H. This means that the dynamics can be reduced to such level sets, which leads to an elimination of some of the variables. It can then be further reduced by factorizing with respect to the residual symmetry. In this section, we provide the algebraic basics needed for this reduction procedure. For that purpose, we investigate the algebraic structure of the tangent spaces of M induced from the tangent mapping J' and from the orbit structure of the action Ψ. The key for this analysis is the Witt-Artin decomposition induced by the kernel of J'. In what follows, we assume J to be equivariant.

Let $m \in M$ and let $\mu = J(m)$. Let G_μ be the stabilizer of μ under the coadjoint action, let G_m be the stabilizer of m under Ψ and let $\mathfrak{g}_\mu$ and $\mathfrak{g}_m$ denote the corresponding Lie algebras. Since J is equivariant, we have

$$G_m \subset G_\mu, \qquad \mathfrak{g}_m \subset \mathfrak{g}_\mu.$$

For the convenience of the reader, we recall the following.

(a) The orbits of G and G_μ through m are denoted by $G \cdot m$ and $G_\mu \cdot m$, respectively.
(b) G_m acts on $\mathrm{T}_m M$ via the isotropy representation, cf. Proposition 6.1.5/4.
(c) The ω_m-orthogonal complement of a subspace $V \subset \mathrm{T}_m M$ is given by

$$V^{\omega_m} = \{X_m \in \mathrm{T}_m M : \omega_m(X_m, Y_m) = 0 \text{ for all } Y_m \in V\}.$$

(d) The annihilator of a vector subspace $\mathfrak{h} \subset \mathfrak{g}$ in $\mathfrak{g}^*$ is given by

$$\mathfrak{h}^0 = \{\mu \in \mathfrak{g}^* : \langle \mu, A \rangle = 0 \text{ for all } A \in \mathfrak{h}\}.$$

The following lemma characterizes the kernel and the image of J'_m.

Lemma 10.2.1 *Let (M, ω, Ψ, J) be a Hamiltonian G-manifold with equivariant momentum mapping and let $m \in M$ and $\mu = J(m)$. Then, we have*

$$\ker J'_m = \mathrm{T}_m(G \cdot m)^{\omega_m}, \tag{10.2.1}$$

$$\operatorname{im} J'_m = \mathfrak{g}_m^0, \tag{10.2.2}$$

$$\left(\ker J'_m\right)^{\omega_m} \cap \ker J'_m = \mathrm{T}_m(G_\mu \cdot m). \tag{10.2.3}$$

Proof For $Y_m \in \mathrm{T}_m M$ and $A \in \mathfrak{g}$, we calculate

$$\omega\big(A_*(m), Y_m\big) = \omega\big(X_{J_A}(m), Y_m\big) = -\mathrm{d}J_A(Y_m) = -Y_m(J_A) = -\big\langle J'_m(Y_m), A\big\rangle.$$

Since the Killing vector fields of Ψ span $\mathrm{T}_m(G \cdot m)$, this proves (10.2.1). Next, for the mapping $(J'_m)^{\mathrm{T}} : \mathfrak{g} \to \mathrm{T}^*_m M$ dual to J'_m, we find

$$\big\langle (J'_m)^{\mathrm{T}}(A), Y_m\big\rangle = \big\langle J'_m(Y_m), A\big\rangle = \mathrm{d}J_A(Y_m) = -\langle A_* \lrcorner\, \omega_m, Y_m\rangle$$

and thus $(J'_m)^{\mathrm{T}}(A) = -A_* \lrcorner \omega_m$. Due to Proposition 6.2.2/3, this implies

$$\ker\big(J'_m\big)^{\mathrm{T}} = \{A \in \mathfrak{g} \colon A_* \lrcorner \omega_m = 0\} = \big\{A \in \mathfrak{g} \colon A_*(m) = 0\big\} = \mathfrak{g}_m .$$

Since the image of a linear mapping coincides with the annihilator of the kernel of the dual mapping, we obtain $\operatorname{im} J'_m = (\ker(J'_m)^{\mathrm{T}})^0 = \mathfrak{g}_m^0$, which proves (10.2.2). Finally, according to (10.2.1), the elements of $(\ker J'_m)^{\omega_m}$ can be written in the form $\Psi'_m(A)$ with $A \in \mathfrak{g}$. By Proposition 6.2.4/2 and (6.2.3), the equivariance property (10.1.8) implies

$$J'_m \circ \Psi'_m(A) = \operatorname{ad}^*(A)\mu .$$

Thus, $\Psi'_m(A) \in \ker J'_m$ iff $\operatorname{ad}^*(A)\mu = 0$. By Proposition 6.2.2/3, this holds iff $A \in \mathfrak{g}_\mu$. Since the Killing vector fields of $\mathfrak{g}_\mu$ span the tangent spaces of the orbit $G_\mu \cdot m$, we obtain (10.2.3). □

Corollary 10.2.2 *Let (M, ω, Ψ, J) be a Hamiltonian G-manifold.*

1. *The rank of J'_m is equal to the dimension of the orbit through m.*
2. *The momentum mapping J is a submersion at m iff G_m is a discrete group. If Ψ is proper, this is equivalent to G_m being a finite group.*

Proof 1. Using (10.2.1) and point 2 of Proposition 7.2.1, we obtain

$$\begin{aligned}
\dim(\mathrm{T}_m M) &= \dim\big(\ker J'_m\big) + \dim\big(\operatorname{im} J'_m\big) \\
&= \dim\big(\mathrm{T}_m(G\cdot m)^{\omega_m}\big) + \operatorname{rank} J'_m \\
&= \dim(\mathrm{T}_m M) - \dim\big(\mathrm{T}_m(G\cdot m)\big) + \operatorname{rank} J'_m
\end{aligned} \tag{10.2.4}$$

and hence $\dim(\mathrm{T}_m(G\cdot m)) = \operatorname{rank} J'_m$.

2. The momentum mapping J is a submersion at m iff its rank is maximal, that is, iff $\operatorname{rank} J'_m = \dim \mathfrak{g}^*$. By point 1 and $\dim \mathrm{T}_m(G\cdot m) = \dim \mathfrak{g} - \dim \mathfrak{g}_m$, this is true iff $\dim \mathfrak{g}_m = 0$, that is, iff G_m is discrete. If Ψ is proper, then G_m is compact and, therefore, finite. □

Remark 10.2.3

1. Since G_m is closed, one has $\dim \mathfrak{g}_m = 0$ for all $m \in M$ iff Ψ is locally free, which means that for every $m \in M$ there exists an open neighbourhood U of the identity $e \in G$, such that
$$U \cap G_m = \{e\}.$$
Thus, if Ψ is locally free, point 2 of Corollary 10.2.2 yields that the momentum mapping is a submersion and hence every value of J is regular. This is of course in particular true for a free G-action.
2. In complete analogy, the action of G_μ on M_μ, viewed as an action of a topological group on a topological space, is locally free iff $\dim \mathfrak{g}_m = 0$ for all $m \in M_\mu$, that is, iff $\mu \in \mathfrak{g}^*$ is regular.

Now, we will discuss the Witt-Artin decomposition of the symplectic vector space $\mathrm{T}_m M$, $m \in M$, with respect to the subspace $W = \ker J'_m$. Choose subspaces E and F such that

$$W = E \oplus W \cap W^{\omega_m}, \qquad W^{\omega_m} = F \oplus W \cap W^{\omega_m}.$$

By Theorem 7.3.3, E and F are symplectic and we have the following direct sum decomposition of $\mathrm{T}_m M$ into symplectic vector subspaces:

$$\mathrm{T}_m M = E \oplus F \oplus (E \oplus F)^{\omega}. \tag{10.2.5}$$

Note that $(E \oplus F)^{\omega}$ contains $W \cap W^{\omega_m}$ as a Lagrangian subspace. According to Lemma 10.2.1,

$$W = \mathrm{T}_m(G \cdot m)^{\omega_m}, \qquad W^{\omega_m} = \mathrm{T}_m(G \cdot m), \qquad W \cap W^{\omega_m} = \mathrm{T}_m(G_\mu \cdot m).$$

For what follows, we assume that G acts properly. Then, the stabilizer G_m is compact and there exists a G_m-invariant scalar product $\langle \cdot, \cdot \rangle$ in $\mathrm{T}_m M$, cf. Proposition 5.5.6. Let us choose E and F as orthogonal complements of the subspace $W \cap W^{\omega_m}$ with respect to this scalar product.

Lemma 10.2.4 *The symplectic subspaces E, F and $(E \oplus F)^{\omega}$ are G_m-invariant.*

Proof By Remark 6.2.10/1, $W^{\omega_m} = \mathrm{T}_m(G \cdot m)$ is G_m-invariant. Since ω_m is G_m-invariant, the ω_m-orthogonal complement $W = \mathrm{T}_m(G \cdot m)^{\omega_m}$ is invariant, too. This proves the invariance of $W \cap W^{\omega_m}$. Now, the assertion follows from the G_m-invariance of the scalar product and of ω_m. □

Since G_m is compact, there exists an $\mathrm{Ad}(G_m)$-invariant scalar product in $\mathfrak{g}$ and we obtain $\mathrm{Ad}(G_m)$-invariant orthogonal vector space decompositions

$$\mathfrak{g}_\mu = \mathfrak{g}_m \oplus \mathfrak{m}, \qquad \mathfrak{g} = \mathfrak{g}_m \oplus \mathfrak{m} \oplus \mathfrak{q}. \tag{10.2.6}$$

There correspond $\mathrm{Ad}^*(G_m)$-invariant decompositions of the dual vector spaces,

$$\mathfrak{g}^*_\mu = \mathfrak{g}^*_m \oplus \mathfrak{m}^*, \qquad \mathfrak{g}^* = \mathfrak{g}^*_m \oplus \mathfrak{m}^* \oplus \mathfrak{q}^*. \tag{10.2.7}$$

Here, $\mathfrak{g}^*_m$, $\mathfrak{m}^*$ and $\mathfrak{q}^*$ are identified with the annihilators in $\mathfrak{g}^*$ of, respectively, $\mathfrak{m} \oplus \mathfrak{q}$, $\mathfrak{g}_m \oplus \mathfrak{q}$ and $\mathfrak{g}_m \oplus \mathfrak{m}$. Let us rewrite the Witt-Artin decomposition (10.2.5) using the subspaces $\mathfrak{m}$ and $\mathfrak{q}$. First, it is clear that the mapping $\Psi'_m : \mathfrak{g} \to \mathrm{T}_m M$ induces vector space isomorphisms

$$W \cap W^{\omega_m} = \mathrm{T}_m(G_\mu \cdot m) \cong \mathfrak{m}, \qquad W^{\omega_m} = \mathrm{T}_m(G \cdot m) \cong \mathfrak{m} \oplus \mathfrak{q}. \tag{10.2.8}$$

Second, if we choose the G_m-invariant scalar product on $\mathrm{T}_m M$ so that

$$\left\langle \Psi'_m(A), \Psi'_m(B) \right\rangle = \langle A, B \rangle$$

for all $A, B \in \mathfrak{m} \oplus \mathfrak{q}$, then Ψ'_m induces an isomorphism

$$F \cong \mathfrak{q}.$$

Third, by choosing a Lagrangian complement L of $W \cap W^{\omega_m}$ in $(E \oplus F)^{\omega_m}$, from Proposition 7.2.9 we obtain a symplectomorphism

$$(E \oplus F)^{\omega} \cong W \cap W^{\omega_m} \oplus \bigl(W \cap W^{\omega_m}\bigr)^*$$

and hence, by (10.2.8), a symplectomorphism

$$j_m : (E \oplus F)^{\omega} \to \mathfrak{m} \oplus \mathfrak{m}^*. \tag{10.2.9}$$

The latter induces a linear embedding $\lambda_L : \mathfrak{m}^* \to \mathrm{T}_m M$ with image L. Then, the inverse of j_m is given by the mapping

$$\bigl(\Psi_m'\bigr)_{\restriction \mathfrak{m}} \oplus \lambda_L : \mathfrak{m} \oplus \mathfrak{m}^* \to \mathrm{T}_m M. \tag{10.2.10}$$

Finally, we introduce the standard notation $V_m \equiv E$. To summarize, from the Witt-Artin decomposition (10.2.5) we obtain the ω_m-orthogonal decomposition

$$\mathrm{T}_m M \cong \mathfrak{q} \oplus \bigl(\mathfrak{m} \oplus \mathfrak{m}^*\bigr) \oplus V_m, \tag{10.2.11}$$

given by the vector space isomorphism $(\Psi_m')_{\restriction \mathfrak{q}} \oplus (\Psi_m')_{\restriction \mathfrak{m}} \oplus \lambda_L \oplus \iota_m$, where $\iota_m : V_m \to \mathrm{T}_m M$ is the natural inclusion mapping. This decomposition is usually referred to as the G_m-invariant Witt-Artin decomposition of the tangent space.[9]

Proposition 10.2.5 *The Witt-Artin decomposition* (10.2.11) *induces the following decomposition of the symplectic form ω at $m \in M_\mu$:*

$$\omega_m = \bigl(J^* \omega^{\mathscr{O}_\mu^-}\bigr)_m + j_m^* \omega^{\mathfrak{m} \oplus \mathfrak{m}^*} + \omega^{V_m}. \tag{10.2.12}$$

Here, $\omega^{\mathscr{O}_\mu^-}$ is the (*negative*) *Kirillov form on $\mathscr{O}_\mu$, $\omega^{\mathfrak{m} \oplus \mathfrak{m}^*}$ denotes the canonical symplectic form on $\mathfrak{m} \oplus \mathfrak{m}^*$ given by* (7.1.5) *and ω^{V_m} is the restriction of ω_m to the symplectic subspace V_m.*

Proof Since the decomposition (10.2.11) is ω_m-orthogonal, for $A_i \in \mathfrak{q}$, $B_i \in \mathfrak{m}$, $\sigma_i \in \mathfrak{m}^*$ and $v_i \in V$, $i = 1, 2$, we find

$$\begin{aligned} &\omega_m\bigl(A_{1*} + B_{1*} + \lambda_L(\sigma_1) + v_1, A_{2*} + B_{2*} + \lambda_L(\sigma_2) + v_2\bigr) \\ &\quad = \omega_m(A_{1*}, A_{2*}) + \omega_m\bigl(B_{1*} + \lambda_L(\sigma_1), B_{2*} + \lambda_L(\sigma_2)\bigr) + \omega_m(v_1, v_2). \end{aligned}$$

By Corollary 10.1.15,

$$\omega_m(A_{1*}, A_{2*}) = \bigl(J^* \omega^{\mathscr{O}_\mu^-}\bigr)_m (A_{1*}, A_{2*}).$$

Since j_m is a symplectomorphism and since its inverse is given by (10.2.10),

$$\begin{aligned} \omega_m\bigl(B_{1*} + \lambda_L(\sigma_1), B_{2*} + \lambda_L(\sigma_2)\bigr) &= \langle \sigma_2, B_1 \rangle - \langle \sigma_1, B_2 \rangle \\ &= \omega^{\mathfrak{m} \oplus \mathfrak{m}^*}\bigl((B_1, \sigma_1), (B_2, \sigma_2)\bigr) \\ &= j_m^* \omega^{\mathfrak{m} \oplus \mathfrak{m}^*}\bigl(B_{1*} + \lambda_L(\sigma_1), B_{2*} + \lambda_L(\sigma_2)\bigr). \end{aligned}$$

□

[9] Note that this decomposition need not be orthogonal with respect to the G_m-invariant scalar product chosen above. If one wants to have an orthogonal Witt-Artin decomposition, one has to redefine the original scalar product by choosing a scalar product on each component and taking the orthogonal direct sum.

Remark 10.2.6

1. Let $m_0 \in M$. The symplectic vector space V_{m_0} is called a linear symplectic slice for Ψ at m_0. Since ω is invariant, along the orbit $G \cdot m_0$, the symplectic slices can be chosen as $V_{\Psi_g(m_0)} = \Psi'_g V_{m_0}$.[10] Then, the union over all $m \in G \cdot m_0$ of the subspaces V_m forms a vertical vector subbundle of $(\mathrm{T}M)_{\restriction G\cdot m_0}$, called the symplectic normal bundle over $G \cdot m_0$. It is isomorphic to the vector bundle $G \times_{G_m} V_{m_0}$ associated with the principal G_{m_0}-bundle $G \to G/G_{m_0}$.
2. The Witt-Artin decomposition will be used for the proof of the Symplectic Tubular Neighbourhood Theorem in Sect. 10.4. In particular, we will use that, according to (10.2.11), the injective linear mapping

$$\lambda := \lambda_L \oplus \iota_m : \mathfrak{m}^* \oplus V_m \to \mathrm{T}_m M$$

sends $\mathfrak{m}^* \oplus V_m$ onto a vector space complement of $\mathrm{T}_m(G \cdot m)$ in $\mathrm{T}_m M$. This implies

$$\mathrm{T}_m M \cong \mathrm{T}_m(G \cdot m) \oplus \mathfrak{m}^* \oplus V_m, \tag{10.2.13}$$

which may be interpreted as the infinitesimal version of the Symplectic Tubular Neighbourhood Theorem. According to Proposition 10.2.5, with respect to this decomposition, the symplectic form ω_m is given by

$$\begin{aligned}&\omega_m\big(A_{1*} + \lambda(\sigma_1, v_1), A_{2*} + \lambda(\sigma_2, v_2)\big)\\ &\quad = \langle \sigma_2, A_1\rangle - \langle \sigma_1, A_2\rangle - \big\langle \mu, [A_1, A_2]\big\rangle + w_m^V(v_1, v_2),\end{aligned} \tag{10.2.14}$$

where $A_i \in \mathfrak{g}$, $\sigma_i \in \mathfrak{m}^*$ and $v_i \in V_m$.
3. If μ is regular, Corollary 10.2.2 implies that $\mathfrak{g}_m = 0$ and hence $\mathfrak{g}_\mu = \mathfrak{m}$. Moreover, by the Level Set Theorem 1.8.3, M_μ is an embedded submanifold of M and $\ker J'_m = \mathrm{T}_m M_\mu$. Hence, Lemma 10.2.1 yields

$$\mathrm{T}_m(G \cdot m)^{\omega_m} = \mathrm{T}_m M_\mu, \tag{10.2.15}$$

$$\mathrm{T}_m(G \cdot m) \cap \mathrm{T}_m M_\mu = \mathrm{T}_m(G_\mu \cdot m), \tag{10.2.16}$$

and thus

$$\mathrm{T}_m(G \cdot m) = (\mathrm{T}_m M_\mu)^{\omega_m}. \tag{10.2.17}$$

These facts will be used in Sect. 10.3.
4. If $\mu = 0$, then $G_\mu = G$ and hence $\mathfrak{q} = 0$. If, in addition, 0 is a regular value, then $\mathfrak{g}_m = 0$ and thus $\mathfrak{m} = \mathfrak{g}_\mu = \mathfrak{g}$. Then, (10.2.17) yields

$$(\mathrm{T}_m M_0)^{\omega_m} = \mathrm{T}_m(G \cdot m) \subset \mathrm{T}_m M_0,$$

that is, $\mathrm{T}_m M_0$ is a coisotropic subspace and the symplectic slice V_m is isomorphic to the symplectic vector space obtained by linear symplectic reduction, cf. (7.3.3).

---

[10] When working with scalar products, this would require the scalar product on $\mathrm{T}_{\Psi_g(m_0)}M$ to be defined by $\langle \Psi'_{g^{-1}}\cdot, \Psi'_{g^{-1}}\cdot\rangle$, which makes sense, because $\langle\cdot,\cdot\rangle$ is G_{m_0}-invariant.

10.3 Regular Symplectic Reduction

Now we are prepared to turn to symplectic reduction in the context of Hamiltonian G-manifolds. In this section, we present the reduction theorem for the regular case. This classical result is due to Marsden and Weinstein [195] and to Meyer [210].

Let (M, ω, Ψ, J) be a left Hamiltonian G-manifold with proper action and equivariant momentum mapping and let μ be a regular value of J. Then, by the Level Set Theorem 1.8.3, $M_\mu = J^{-1}(\mu)$ is an embedded submanifold of M. Since it is G_μ-invariant, according to Proposition 6.3.4/1, Ψ restricts to a proper action

$$\Psi^\mu : G_\mu \times M_\mu \to M_\mu .$$

According to Remark 10.2.3/1, this action is locally free. One can show that in this case the orbit space M_μ / G_μ is a symplectic orbifold, see [69]. In the sequel, we restrict ourselves to the special case where the G_μ-action is not only locally free, but free.

Theorem 10.3.1 (Regular Reduction) *Let (M, ω, Ψ, J) be a Hamiltonian G-manifold with proper action and equivariant momentum mapping. Let $\mu \in \mathfrak{g}^*$ be a regular value of J and assume that the induced G_μ-action is free.*

1. *The topological space M_μ / G_μ carries a unique manifold structure such that the natural projection $\pi_\mu : M_\mu \to M_\mu / G_\mu$ is a submersion.*
2. *There exists a unique symplectic form ω^μ on M_μ / G_μ such that*
$$\pi_\mu^* \omega^\mu = j_\mu^* \omega, \tag{10.3.1}$$
where $j_\mu : M_\mu \to M$ denotes the natural inclusion mapping.

In the context of Hamiltonian systems, the symplectic manifold $(M_\mu / G_\mu, \omega^\mu)$ is referred to as the reduced phase space.

Proof Point 1 is due to Corollary 6.5.1. To prove point 2, we observe that since π_μ is a surjective submersion, every tangent vector of M_μ / G_μ can be written in the form $\pi'_\mu X$ for some $X \in \mathrm{T}M_\mu$. Thus, we may define ω^μ by

$$\omega^\mu_{\pi_\mu(m)}\big(\pi'_\mu X, \pi'_\mu Y\big) := \omega_m(X, Y), \tag{10.3.2}$$

where $m \in M_\mu$ and $X, Y \in \mathrm{T}_m M_\mu$. To prove that ω^μ is well-defined, we must show that the right hand side does not depend on the choice of m, X and Y. For that purpose, let $\tilde{m} \in M_\mu$ and $\tilde{X}, \tilde{Y} \in \mathrm{T}_{\tilde{m}} M_\mu$ such that $\pi_\mu(\tilde{m}) = \pi_\mu(m)$, $\pi'_\mu \tilde{X} = \pi'_\mu X$ and $\pi'_\mu \tilde{Y} = \pi'_\mu Y$. Then, there exists $g \in G$ such that

$$m = \Psi^\mu_g(\tilde{m}), \qquad \big(\Psi^\mu_g\big)' \tilde{X} - X \in \mathrm{T}_m(G_\mu \cdot m), \qquad \big(\Psi^\mu_g\big)' \tilde{Y} - Y \in \mathrm{T}_m(G_\mu \cdot m).$$

By G-invariance of ω, we have

$$\begin{aligned}
\omega_{\tilde{m}}(\tilde{X}, \tilde{Y}) &= \big(\big(\Psi^\mu_{g^{-1}}\big)^* \omega\big)_m \big(\big(\Psi^\mu_g\big)' \tilde{X}, \big(\Psi^\mu_g\big)' \tilde{Y}\big) \\
&= \omega_m\big(\big(\Psi^\mu_g\big)' \tilde{X}, \big(\Psi^\mu_g\big)' \tilde{Y}\big) \\
&= \omega_m\big(X + \big(\big(\Psi^\mu_g\big)' \tilde{X} - X\big), Y + \big(\big(\Psi^\mu_g\big)' \tilde{Y} - Y\big)\big).
\end{aligned}$$

By (10.2.17), the right hand side equals $\omega_m(X, Y)$. Hence, the definition (10.3.2) makes sense, indeed.

By construction, the pointwise defined 2-form ω^μ fulfils (10.3.1). Since π_μ is a surjective submersion, this implies that ω^μ is uniquely determined and smooth. For the same reason, $\mathrm{d}\omega = 0$ implies $\mathrm{d}\omega^\mu = 0$. It remains to show that ω^μ is non-degenerate. Thus, let $m \in M_\mu$ and $X \in \mathrm{T}_m M_\mu$ such that

$$\omega^\mu_{\pi_\mu(m)}\big(\pi'_\mu X, \pi'_\mu Y\big) = 0 \quad \text{for all } Y \in \mathrm{T}_m M_\mu. \tag{10.3.3}$$

We have to show that this implies $\pi'_\mu X = 0$. By definition of ω^μ, (10.3.3) implies $\omega_m(X, Y) = 0$ for all $Y \in \mathrm{T}_m M_\mu$, that is, $X \in (\mathrm{T}_m M_\mu)^{\omega_m}$. Now, by (10.2.16) and (10.2.17),

$$(\mathrm{T}_m M_\mu)^{\omega_m} \cap \mathrm{T}_m M_\mu = \mathrm{T}_m(G_\mu \cdot m).$$

Thus, $\pi'_\mu X = 0$ holds, indeed. □

Remark 10.3.2 The assumption that $\mu \in \mathfrak{g}^*$ be regular can be weakened. According to Proposition 1.7.6, if J is a subimmersion, M_μ is still a closed submanifold. In this case, μ is sometimes called weakly regular. For a discussion of reduction under this weaker assumption, we refer to [181].

Now we can discuss the reduction of a G-invariant Hamiltonian system (M, ω, H) with equivariant momentum mapping:

Proposition 10.3.3 *Let (M, ω, Ψ, J) be a Hamiltonian G-manifold with proper action and equivariant momentum mapping and let $H \in C^\infty(M)^G$. Let $\mu \in \mathfrak{g}^*$ be a regular value of J and assume that the induced G_μ-action is free.*

1. *M_μ is invariant under the flow of X_H and X_H restricts to a vector field X^μ_H on M_μ which is j_μ-related to X_H,*

$$j'_\mu \circ X^\mu_H = X_H \circ j_\mu. \tag{10.3.4}$$

2. *H defines a unique smooth function H_μ on M_μ/G_μ by*

$$H_\mu \circ \pi_\mu = H \circ j_\mu. \tag{10.3.5}$$

   The corresponding Hamiltonian vector field X_{H_μ} is π_μ-related to X^μ_H,

$$X_{H_\mu} \circ \pi_\mu = \pi'_\mu \circ X^\mu_H, \tag{10.3.6}$$

The Hamiltonian system $(M_\mu/G_\mu, \omega^\mu, H_\mu)$ is referred to as the reduction of the Hamiltonian system (M, ω, H) at μ.

Proof 1. Invariance of M_μ follows from the Noether Theorem 10.1.9. Then, X_H is tangent to M_μ and the rest of the assertion follows from Proposition 2.7.16.

2. Since H is G-invariant, a function H_μ on M_μ/G_μ satisfying (10.3.5) exists. Since π_μ is surjective, H_μ is uniquely determined and since π_μ is a submersion,

H_μ is smooth. It remains to prove (10.3.6). According to (6.7.5), since X_H^μ is G_μ-invariant, it projects to a vector field $\hat{X}_H^\mu$ on M_μ/G_μ, uniquely determined by

$$\hat{X}_H^\mu \circ \pi_\mu = \pi'_\mu \circ X_H^\mu. \tag{10.3.7}$$

It suffices to show that $\hat{X}_H^\mu = X_{H_\mu}$. By (10.3.5), we have $\mathrm{d}H_\mu \circ \pi'_\mu = \mathrm{d}H \circ j'_\mu$. Using this, as well as (10.3.4) and (10.3.7), for $m \in M_\mu$ and $Y \in \mathrm{T}_m M_\mu$ we obtain

$$\begin{aligned}(\mathrm{d}H_\mu)_{\pi_\mu(m)}\big(\pi'_\mu Y\big) &= (\mathrm{d}H)_{j_\mu(m)}\big(j'_\mu Y\big)\\ &= -\omega_{j_\mu(m)}\big((X_H)_{j_\mu(m)}, j'_\mu Y\big)\\ &= -\big(j_\mu^*\omega\big)_m\big(\big(X_H^\mu\big)_m, Y\big)\\ &= -\big(\pi_\mu^*\omega^\mu\big)_m\big(\big(X_H^\mu\big)_m, Y\big)\\ &= -\omega^\mu_{\pi_\mu(m)}\big(\big(\hat{X}_H^\mu\big)_{\pi_\mu(m)}, \pi'_\mu Y\big).\end{aligned}$$

It follows that $\mathrm{d}H_\mu = -\hat{X}_H^\mu \lrcorner\, \omega^\mu$ and hence $\hat{X}_H^\mu = X_{H_\mu}$, indeed. □

Corollary 10.3.4 *The Poisson structures defined by ω and ω^μ are compatible, that is, one has*

$$\{f_\mu, h_\mu\} \circ \pi_\mu = \{f, h\} \circ j_\mu, \tag{10.3.8}$$

for all functions $f, h \in C^\infty(M_\mu)^G$ and $f_\mu, h_\mu \in C^\infty(M_\mu/G_\mu)$ related by (10.3.5).

Proof Using (10.3.4) and (10.3.6), we calculate

$$\begin{aligned}\{f_\mu, h_\mu\}\big(\pi_\mu(m)\big) &= \omega^\mu_{\pi_\mu(m)}(X_{f_\mu}, X_{h_\mu})\\ &= \big(\pi_\mu^*\omega^\mu\big)_m\big(X_f^\mu, X_h^\mu\big)\\ &= \big(j_\mu^*\omega\big)_m\big(X_f^\mu, X_h^\mu\big)\\ &= \omega_{j_\mu(m)}(X_f, X_h)\\ &= \{f, h\}\big(j_\mu(m)\big).\end{aligned}$$

□

Now, let us apply the Regular Reduction Theorem 10.3.1 to the action of G on T^*G induced by left translation, cf. Example 10.1.24. Thus, we consider $Q = G$ with G acting by

$$\psi: G \times Q \to Q, \qquad \psi(g, a) := \mathrm{L}_g(a) = ga,$$

and with the induced action Ψ on T^*Q given by

$$\Psi : G \times \mathrm{T}^*Q \to \mathrm{T}^*Q, \qquad \Psi_g(\xi) = (\mathrm{L}_{g^{-1}})'^{\mathrm{T}}(\xi).$$

Recall from Example 10.1.24 that the corresponding equivariant momentum mapping is given by

$$J(\xi) = (\mathrm{R}_{\pi(\xi)})'^{\mathrm{T}}(\xi),$$

cf. (10.1.25). The action Ψ is obviously free. Since ψ is proper, Remark 6.3.9 implies that Ψ is proper, too. Let $\mu \in \mathfrak{g}^* \cong \mathrm{T}_e^*Q$ and denote by $\alpha_\mu \in \Omega^1(Q)$ the right invariant differential form on Q defined by $\alpha_\mu(e) = \mu$. Thus, $\alpha_\mu(a) = (\mathrm{R}_{a^{-1}})'^{\mathrm{T}}(\mu)$. Since

$$\big\langle J\big(\alpha_\mu(a)\big), A\big\rangle = \big\langle \alpha_\mu(a), (\mathrm{R}_a)'_e(A)\big\rangle = \langle \mu, A\rangle,$$

we obtain $J(\alpha_\mu(a)) = \mu$. This implies

$$(\mathrm{T}^*Q)_\mu = J^{-1}(\mu) = \alpha_\mu(Q),$$

that is, the level set of μ is given by the image of the 1-form α_μ. Thus, $(\mathrm{T}^*Q)_\mu$ is an embedded submanifold, diffeomorphic to the group manifold G. Due to

$$\Psi_g\big(\alpha_\mu(a)\big) = (\mathrm{L}_{g^{-1}})'^{\mathrm{T}} \circ (\mathrm{R}_{a^{-1}})'^{\mathrm{T}}(\mu) = (\mathrm{R}_{g^{-1}} \circ \mathrm{R}_{a^{-1}})'^{\mathrm{T}} \circ \mathrm{Ad}^*\big(g^{-1}\big)\mu,$$

the G_μ-action on $(\mathrm{T}^*Q)_\mu$ is given by

$$\Psi_g\big(\alpha_\mu(a)\big) = \alpha_\mu(ga). \tag{10.3.9}$$

Since $(\mathrm{R}_{\pi(\xi)})'^{\mathrm{T}}$ is fibrewise bijective, J is a submersion. Hence, we can apply the Regular Reduction Theorem 10.3.1, which yields a symplectic manifold structure on the quotient $(\mathrm{T}^*Q)_\mu/G_\mu$. The following theorem yields an explicit description of this quotient.

Theorem 10.3.5 *Let G be a Lie group and let $(\mathrm{T}^*G, \omega, \Psi, J)$ be the Hamiltonian G-manifold of Example 10.1.24. For every $\mu \in \mathfrak{g}^*$, the reduced phase space $((\mathrm{T}^*G)_\mu/G_\mu, \omega^\mu)$ is isomorphic to the coadjoint orbit of μ endowed with the positive Kirillov structure.*

For $\mu = 0$, both symplectic spaces obviously degenerate to the one-point-space.

Proof Let θ be the canonical 1-form and $\omega = \mathrm{d}\theta$ the canonical symplectic form on T^*G. Denote by $\omega^{\mathscr{O}_\mu^+}$ the positive Kirillov form on the coadjoint orbit $\mathscr{O}_\mu$ of μ, see Sect. 8.4. Consider the mapping

$$\big(\mathrm{T}^*G\big)_\mu \to \mathscr{O}_\mu, \qquad \alpha_\mu(a) \mapsto \mathrm{Ad}^*\big(a^{-1}\big)\mu.$$

Since for any $g \in G_\mu$, this mapping sends $\Psi_g\alpha_\mu(a) = \alpha_\mu(ga)$ to

$$\mathrm{Ad}^*\big((ga)^{-1}\big)\mu = \mathrm{Ad}^*\big(a^{-1}\big) \circ \mathrm{Ad}^*\big(g^{-1}\big)\mu = \mathrm{Ad}^*\big(a^{-1}\big)\mu,$$

it induces a bijection

$$\varphi\colon \big(\mathrm{T}^*G\big)_\mu/G_\mu \to \mathscr{O}_\mu. \tag{10.3.10}$$

We decompose φ as follows:

$$\big(\mathrm{T}^*G\big)_\mu/G_\mu \to G/G_\mu \to \mathscr{O}_\mu.$$

Here, the first mapping is obviously a diffeomorphism, induced by the natural projection of T^*G. The second mapping is the diffeomorphism (6.1.8) provided by the Orbit Theorem 6.2.8. Thus, φ is a diffeomorphism. We prove

$$\varphi^*\omega^{\mathscr{O}_\mu^+} = \omega^\mu.$$

Due to (10.3.1), for that purpose it is enough to show that

$$\pi_\mu^* \circ \varphi^* \omega^{\mathscr{O}_\mu^+} = \omega_{\restriction \mathrm{T}(\mathrm{T}^*G)_\mu}. \tag{10.3.11}$$

Now, the tangent vectors to $(\mathrm{T}^*G)_\mu$ at $\alpha_\mu(a)$ can be written as $\alpha'_\mu \tilde{A}(a)$, with $\tilde{A}$ denoting the right invariant vector field on G, generated by $A \in \mathfrak{g}$. Thus, for the left hand side we get:

$$\pi_\mu^* \circ \varphi^* \omega^{\mathscr{O}_\mu^+} \big(\alpha'_\mu \tilde{A}(a), \alpha'_\mu \tilde{B}(a)\big) = \omega^{\mathscr{O}_\mu^+} \big((\varphi \circ \pi_\mu \circ \alpha_\mu)' \tilde{A}(a), (\varphi \circ \pi_\mu \circ \alpha_\mu)' \tilde{B}(a)\big).$$

For any $f \in C^\infty(\mathscr{O}_\mu)$,

$$\begin{aligned}
\big((\varphi \circ \pi_\mu \circ \alpha_\mu)' \tilde{A}(a)\big)(f) &= \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} f \circ \varphi \circ \pi_\mu \circ \alpha_\mu \big(\exp(tA)a\big) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} f\big(\mathrm{Ad}^*\big(a^{-1}\exp(-tA)\big)\mu\big) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} f\big(\mathrm{Ad}^*\big(a^{-1}\exp(-tA)a\big) \circ \mathrm{Ad}^*\big(a^{-1}\big)\mu\big) \\
&= -\big(\mathrm{Ad}\big(a^{-1}\big)A\big)_*^{\mathrm{Ad}^*}(f)\big(\mathrm{Ad}^*\big(a^{-1}\big)\mu\big),
\end{aligned}$$

where $A_*^{\mathrm{Ad}^*}$ denotes the Killing vector field of the coadjoint action generated by A. It follows that

$$(\varphi \circ \pi_\mu \circ \alpha_\mu)' \tilde{A}(a) = -\big(\mathrm{Ad}\big(a^{-1}\big)A\big)_*^{\mathrm{Ad}^*}\big(\mathrm{Ad}^*\big(a^{-1}\big)\mu\big)$$

and the left hand side of (10.3.11) takes the form

$$\omega^{\mathscr{O}_\mu^+}\big(\big(\mathrm{Ad}\big(a^{-1}\big)A\big)_*^{\mathrm{Ad}^*}, \big(\mathrm{Ad}\big(a^{-1}\big)B\big)_*^{\mathrm{Ad}^*}\big)\big(\mathrm{Ad}^*\big(a^{-1}\big)\mu\big) = -\big\langle \mu, [A, B]\big\rangle.$$

Here we have used (8.4.2). For the right hand side of (10.3.11) we have

$$\begin{aligned}
\omega\big(\alpha'_\mu \tilde{A}(a), \alpha'_\mu \tilde{B}(a)\big) &= \big(\alpha_\mu^* \omega\big)_a (\tilde{A}, \tilde{B}) \\
&= (\mathrm{d}\alpha_\mu)_a (\tilde{A}, \tilde{B}) \\
&= \tilde{A}_a\big(\langle \alpha_\mu, \tilde{B}\rangle\big) - \tilde{B}_a\big(\langle \alpha_\mu, \tilde{A}\rangle\big) - \big\langle \alpha_\mu, [\tilde{A}, \tilde{B}]\big\rangle(a),
\end{aligned}$$

where we have used (8.3.3). By right-invariance, $\langle \alpha_\mu, \tilde{B}\rangle = \langle \mu, B\rangle$. Hence, the first two terms vanish and we obtain

$$-\big\langle \alpha_\mu, [\tilde{A}, \tilde{B}]\big\rangle(a) = \big\langle \mu, [A, B]\big\rangle.$$

Thus, we have proved (10.3.11), which implies that the mapping (10.3.10) is a symplectomorphism, indeed. □

To conclude this section, we discuss an alternative reduction prescription leading to the same reduced phase spaces. It is usually referred to as orbit reduction. We will see that the choice of a concrete value $\mu \in \mathfrak{g}^*$ is not important and that it is in fact the orbit $\mathscr{O}_\mu$ through μ which is relevant. Consider the preimage

$$\mathscr{M}_\mu := J^{-1}(\mathscr{O}_\mu) \tag{10.3.12}$$

of the orbit $\mathscr{O}_\mu$ under J. Since J is equivariant, we have $\mathscr{M}_\mu = G \cdot M_\mu$. As in the Regular Reduction Theorem, we assume that μ is a regular value and that G_μ acts freely on $J^{-1}(\mu)$, which is equivalent to the assumption that G acts freely on $\mathscr{M}_\mu$. By the Orbit Theorem 6.2.8, $\mathscr{O}_\mu$ is an initial submanifold of $\mathfrak{g}^*$. Since μ is regular, J is a submersion at every point of $\mathscr{M}_\mu$ and hence the Transversal Mapping Theorem 1.8.2 implies that $\mathscr{M}_\mu$ is an initial submanifold of M. Since $\mathscr{M}_\mu$ is G-invariant, the action Ψ restricts to an action of G on $\mathscr{M}_\mu$. By Proposition 6.3.4/1, this action is proper. Hence, Corollary 6.5.1 endows $\mathscr{M}_\mu/G$ with a unique smooth structure such that the natural projection $\tilde{\pi}_\mu : \mathscr{M}_\mu \to \mathscr{M}_\mu/G$ is a submersion. Since $\mathscr{M}_\mu$ is initial, the natural inclusion mapping $M_\mu \to \mathscr{M}_\mu$ is smooth. It is easy to see that it descends to a bijection

$$\varphi : M_\mu/G_\mu \to \mathscr{M}_\mu/G. \tag{10.3.13}$$

Since π_μ and $\tilde{\pi}_\mu$ are submersions, φ is a diffeomorphism.

Remark 10.3.6

1. The G-manifold $\mathscr{M}_\mu$ has the structure of a bundle with fibre M_μ associated with the principal G_μ-bundle $\pi_\mu : M_\mu \to M_\mu/G_\mu$. Indeed, the mapping

$$F : G \times_{G_\mu} M_\mu \to \mathscr{M}_\mu, \qquad F\big([(g, m)]\big) := \Psi_g(m), \tag{10.3.14}$$

is a G-equivariant diffeomorphism (Exercise 10.3.1).
2. Alternatively, one can use the Tubular Neighbourhood Theorem for showing that $\mathscr{M}_\mu$ is a submanifold of M on which G acts smoothly, freely and properly (Exercise 10.3.2).

Proposition 10.3.7 *Let (M, ω, Ψ, J) be a left Hamiltonian G-manifold with proper action and equivariant momentum mapping. Let $\mu \in \mathfrak{g}^*$ be a regular value and let $\mathscr{O}_\mu$ be the coadjoint orbit of μ. Assume that the restriction of Ψ to the submanifold $\mathscr{M}_\mu$ is free. Let $\iota_\mu : \mathscr{M}_\mu \to M$ denote the natural inclusion mapping.*

1. *There exists a unique symplectic form $\tilde{\omega}^\mu$ on $\mathscr{M}_\mu/G$ such that*

$$\tilde{\pi}_\mu^* \tilde{\omega}^\mu = \iota_\mu^* \omega + J^* \omega^{\mathscr{O}_\mu^+}. \tag{10.3.15}$$

2. *The symplectic manifolds $(M_\mu/G_\mu, \omega^\mu)$ and $(\mathscr{M}_\mu/G, \tilde{\omega}^\mu)$ are symplectomorphic.*

Proof 1. The right hand side of (10.3.15) defines a 2-form pointwise on $\mathscr{M}_\mu/G$: for $A, B \in \mathfrak{g}$ and $m \in \mathscr{M}_\mu$, we calculate, denoting $J(m) \equiv \nu$,[11]

$$\begin{aligned}
&\big(\iota_\mu^* \omega\big)_m (A_*, B_*) + \big(J^* \omega^{\mathscr{O}_\mu^+}\big)_m (A_*, B_*) \\
&= \omega_m(X_{J_A}, X_{J_B}) + \omega_\nu^{\mathscr{O}_\mu^+} \big(J'_m(A_*)_m, J'_m(B_*)_m\big)
\end{aligned}$$

[11] Alternatively, one may observe that, on Killing vector fields, $(\iota_\mu^* \omega)_m$ coincides with the pull-back of ω to the orbit $G \cdot m$ and apply Corollary 10.1.5.

$$= J_{[B,A]}(m) + \langle \nu, [A, B] \rangle$$
$$= \langle \nu, [B, A] \rangle + \langle \nu, [A, B] \rangle = 0.$$

Thus, $\tilde{\omega}^\mu$ is well-defined. Since $\tilde{\pi}_\mu$ is a surjective submersion, $\tilde{\omega}^\mu$ is smooth and $\mathrm{d}\omega = 0 = \mathrm{d}\omega^{\mathscr{O}_\mu^+}$ implies $\mathrm{d}\tilde{\omega}^\mu = 0$. Non-degeneracy will follow from point 2.

2. We show that the diffeomorphism φ given by (10.3.13) fulfils $\varphi^*\tilde{\omega}^\mu = \omega^\mu$. This implies in particular that $\tilde{\omega}^\mu$ is non-degenerate and hence symplectic. Since π_μ is a submersion, it is enough to show that $\pi_\mu^* \circ \varphi^*\tilde{\omega}^\mu = \pi_\mu^*\omega^\mu$. With $i_\mu : M_\mu \to \mathscr{M}_\mu$ denoting the natural inclusion mapping, we have $\varphi \circ \pi_\mu = \tilde{\pi}_\mu \circ i_\mu$ and hence

$$\pi_\mu^* \circ \varphi^*\tilde{\omega}^\mu = (\tilde{\pi}_\mu \circ i_\mu)^*\tilde{\omega}^\mu = (\iota_\mu \circ i_\mu)^*\omega + (J \circ i_\mu)^*\omega^{\mathscr{O}_\mu^+}.$$

Using $\iota_\mu \circ i_\mu = j_\mu$ and (10.3.1), for the first term of this sum we get $j_\mu^*\omega = \pi_\mu^*\omega^\mu$. The second term vanishes, because $J \circ i_\mu$ is constant. This proves $\varphi^*\tilde{\omega}^\mu = \omega^\mu$. □

Corollary 10.3.8 *The Poisson structures of* $(M_\mu/G_\mu, \omega^\mu)$ *and* $(\mathscr{M}_\mu/G, \tilde{\omega}^\mu)$ *are isomorphic.*

Remark 10.3.9 Orbit reduction can be also performed using the shifting trick, see [232], Theorem 6.5.2: for a given Hamiltonian G-manifold (M, ω, Ψ, J) with equivariant momentum mapping J and a coadjoint orbit $\mathscr{O}$, one considers the G-manifold $M \times \mathscr{O}$ with symplectic form

$$\tilde{\omega} = \pi_1^*\omega + \pi_2^*\omega^{\mathscr{O}}$$

and with equivariant momentum mapping

$$K : M \times \mathscr{O} \to \mathfrak{g}^*, \qquad K(m, \mu) := J(m) - \mu,$$

cf. Exercise 10.1.3. Thus, instead of considering the preimage $J^{-1}(\mathscr{O})$, one can consider $K^{-1}(0)$. This way, the symplectic reduction problem is reduced to the case $\mu = 0$.

Exercises

10.3.1 Show that the mapping (10.3.14) is a G-equivariant diffeomorphism.

10.3.2 Using the Tubular Neighbourhood Theorem, show that the subset $\mathscr{M}_\mu$ of M, defined by (10.3.12), is a submanifold on which G acts smoothly, freely and properly.

10.4 The Symplectic Tubular Neighbourhood Theorem

The results of this section will be used for the subsequent discussion of singular symplectic reduction. Let (M, ω, Ψ) be a Hamiltonian G-manifold with proper action and equivariant momentum mapping J. Let O be an orbit of Ψ and let

$\mathscr{O} = J(O)$ be the corresponding coadjoint orbit. Choose a point $m \in O$ and denote $\mu = J(m)$. The starting point of our analysis is the Witt-Artin decomposition[12] (10.2.13) of $\mathrm{T}_m M$ defined by the choice of an appropriate linear embedding $\lambda : \mathfrak{m}^* \oplus V \to \mathrm{T}_m M$ onto a vector space complement of $\mathrm{T}_m O$. This way, $\mathfrak{m}^* \oplus V$ may be viewed as a model for the normal space to the orbit at m. According to the Tubular Neighbourhood Theorem 6.4.3 and Remark 6.5.8 there exists a tubular neighbourhood U of O in M and a G-equivariant diffeomorphism

$$\chi: U \to E = G \times_{G_m} \left(\mathfrak{m}^* \oplus V\right)$$

onto an open neighbourhood $\tilde{E}$ of the zero section of E, viewed as a vector bundle over G/G_m. Our aim is to construct a Hamiltonian G-manifold structure on $\tilde{E}$ and to deform χ in a G-equivariant way so that it becomes symplectic and thus an isomorphism of symplectic G-manifolds from U onto $\tilde{E}$. In the course of this, we will use the following two commuting left G-actions on $G \times \mathfrak{g}^*$ introduced in Example 10.1.24:

(a) the action induced by left translation on G,

$$\mathscr{L}: G \times \left(G \times \mathfrak{g}^*\right) \to \left(G \times \mathfrak{g}^*\right), \qquad \mathscr{L}_a(g, \nu) = (ag, \nu), \tag{10.4.1}$$

with equivariant momentum mapping $J^{\mathscr{L}}(g, \nu) = \mathrm{Ad}^*(g)\nu$,

(b) the action induced by right translation with the inverse group element,

$$\mathscr{R}: G \times \left(G \times \mathfrak{g}^*\right) \to \left(G \times \mathfrak{g}^*\right), \qquad \mathscr{R}_a(g, \nu) = \left(ga^{-1}, \mathrm{Ad}^*(a)\nu\right) \tag{10.4.2}$$

with equivariant momentum mapping $J^{\mathscr{R}}(g, \nu) = -\nu$.

Since E is the quotient of a free G_m-action on $G \times (\mathfrak{m}^* \oplus V)$, it is natural to construct $\tilde{E}$ by regular symplectic reduction of a free Hamiltonian G_m-manifold. For that purpose, we add the factor $\mathfrak{g}_m^*$ as a symplectic partner for the necessarily isotropic G_m-orbits and consider the following auxiliary trivial vector bundle over G:

$$\mathfrak{E} := G \times \left(\mathfrak{g}_m^* \oplus \mathfrak{m}^* \oplus V\right). \tag{10.4.3}$$

This bundle is endowed with the G_m-action

$$\left(a, (g, \eta, \rho, v)\right) \mapsto \left(ga^{-1}, \mathrm{Ad}^*(a)\eta, \mathrm{Ad}^*(a)\rho, (\Psi_a)'_m v\right) \tag{10.4.4}$$

and with the G-action induced by left translation on the first factor,

$$\left(h, (g, \eta, \rho, v)\right) \mapsto (hg, \eta, \rho, v). \tag{10.4.5}$$

By means of a chosen G_m-invariant scalar product on $\mathfrak{g}$, we can identify $\mathfrak{g}_m^* \oplus \mathfrak{m}^*$ with $\mathfrak{g}_\mu^*$ and the latter with a linear subspace of $\mathfrak{g}^*$. Accordingly, for $(g, \eta, \rho, v) \in \mathfrak{E}$, we can decompose

$$\mathrm{T}_{(g,\eta,\rho,v)}\mathfrak{E} = \mathfrak{q} \oplus \mathfrak{g}_\mu \oplus \mathfrak{g}_\mu^* \oplus V, \tag{10.4.6}$$

[12] In the sequel, for convenience, we write $V \equiv V_m$.

cf. Sect. 10.2 for the notations. Here, $\mathfrak{q} \cong \mathfrak{g}/\mathfrak{g}_\mu$ can be viewed as the tangent space to the coadjoint orbit $\mathscr{O}_\mu$ at μ and $\mathfrak{g}_\mu \oplus \mathfrak{g}_\mu^*$ models the tangent spaces of T^*G_μ. This suggests to take the following closed 2-form as a candidate for a symplectic structure on $\mathfrak{E}$:

$$\omega^\times := p_\mu^* \omega^{\mathscr{O}_\mu^-} + \iota_\mu^* \omega^{\mathrm{T}^*G} + \omega^V. \tag{10.4.7}$$

Here, $\omega^{\mathscr{O}_\mu^-}$, ω^{T^*G} and ω^V denote, respectively, the (negative) Kirillov form on $\mathscr{O} = \mathscr{O}_\mu$, the natural symplectic form on T^*G and the symplectic form on $V \subset \mathrm{T}_m M$ induced from ω_m. The mapping p_μ is given by

$$p_\mu : G \to \mathscr{O}_\mu, \qquad p_\mu(g) := \mathrm{Ad}^*(g)\mu$$

and $\iota_\mu : G \times \mathfrak{g}_\mu^* \to \mathrm{T}^*G$ is the embedding defined by the left trivialization of T^*G, cf. (8.3.6). According to (10.4.3), we write tangent vectors of $\mathfrak{E}$ at (g, η, ρ, v) in the form

$$\big(\mathrm{L}'_g A, \xi, \sigma, u\big), \quad A \in \mathfrak{g},\ \xi \in \mathfrak{g}_m^*,\ \sigma \in \mathfrak{m}^*,\ u \in V.$$

A straightforward calculation using (8.3.8) and (8.4.2) yields

$$\begin{aligned} &\omega^\times_{(g,\eta,\rho,v)}\Big(\big(\mathrm{L}'_g A_1, \xi_1, \sigma_1, u_1\big), \big(\mathrm{L}'_g A_2, \xi_2, \sigma_2, u_2\big)\Big) \\ &\quad = \langle \xi_1 + \sigma_1, A_2\rangle - \langle \xi_2 + \sigma_2, A_1\rangle - \big\langle \mu + \eta + \rho, [A_1, A_2]\big\rangle + \omega_m(u_1, u_2). \end{aligned} \tag{10.4.8}$$

Recall from Exercise 10.1.2 that the linear G_m-action on V induced from the isotropy representation is Hamiltonian with momentum mapping

$$J^V : V \to \mathfrak{g}_m^*, \qquad \big\langle J^V(v), A\big\rangle := \frac{1}{2}\omega_m\big(v, \mathrm{Hess}_m(A_*)v\big). \tag{10.4.9}$$

Lemma 10.4.1 *There exists an open neighbourhood $\mathfrak{E}^\times$ of the zero section in $\mathfrak{E}$, invariant under both G and G_m, such that $\omega^\times$ is symplectic on $\mathfrak{E}^\times$. Moreover,*

1. $(\mathfrak{E}^\times, \omega^\times)$ *endowed with the action* (10.4.4) *is a Hamiltonian G_m-manifold with G_m-equivariant and G-invariant momentum mapping*

$$K : \mathfrak{E}^\times \to \mathfrak{g}_m^*, \qquad K(g, \eta, \rho, v) := J^V(v) - \eta,$$

2. $(\mathfrak{E}^\times, \omega^\times)$ *endowed with the action* (10.4.5) *is a Hamiltonian G-manifold with G-equivariant and G_m-invariant momentum mapping*

$$J^\times : \mathfrak{E}^\times \to \mathfrak{g}_m^*, \qquad J^\times(g, \eta, \rho, v) := \mathrm{Ad}^*(g)(\mu + \eta + \rho).$$

Denote the restriction of the G-action (10.4.5) to $\mathfrak{E}^\times$ by $\Psi^\times$.

Proof First, we check that $\omega^\times$ is invariant under both the action (10.4.4) and the action (10.4.5). This is a consequence of the following facts:

(a) $p_\mu(hg) = \mathrm{Ad}^*(h)p_\mu(g)$ for all $h \in G$ and $p_\mu(ga^{-1}) = p_\mu(g)$ for all $a \in G_\mu$,
(b) the left trivialization of T^*G intertwines the action of G on T^*G induced from left translation (right translation by the inverse group element) on G with the actions $\mathscr{L}$ and $\mathscr{R}$, respectively,
(c) $\omega^{\mathscr{O}_\mu^-}$ is invariant under the action of G on $\mathscr{O}_\mu$, ω^{T^*G} is invariant under point transformations and ω^V is invariant under the isotropy representation.

Second, from (10.4.8) we read off that under the identification (10.4.6), $\omega^\times_{(\mathbb{1},0,0,0)}$ coincides with the direct sum of the Kirillov form on $\mathrm{T}_\mu\mathscr{O} \cong \mathfrak{q}$, the canonical symplectic form on $\mathfrak{g}_\mu \oplus \mathfrak{g}_\mu^*$ and ω^V. Hence, it is non-degenerate. By G-invariance, then $\omega^\times$ is non-degenerate on the zero section of $\mathfrak{E}$ and hence on some G-invariant open neighbourhood $\mathfrak{E}^\times$ of the zero section. Since G_m is compact, $\mathfrak{E}^\times$ can be chosen to be G_m-invariant as well.

1. To see that K is a momentum mapping for the G_m-action (10.4.4), we interpret $\mathfrak{E}$ as the direct product of the symplectic G_m-manifolds

$$\left(G \times \mathfrak{g}_m^* \oplus \mathfrak{m}^*, \, p_\mu^*\omega^{\mathscr{O}_\mu^-} + \iota_\mu^*\omega^{\mathrm{T}^*G}\right)$$

and (V, ω^V). One can check that the first one is Hamiltonian, where the momentum mapping is obtained from that of the action $\mathscr{R}$ on $(G \times \mathfrak{g}^*, \omega^{\mathrm{T}^*G})$ by restriction to $G \times \mathfrak{g}_m^* \oplus \mathfrak{m}^*$ and by composition with the projection[13] $\mathfrak{g}^* \to \mathfrak{g}_m^*$. That is, the momentum mapping is given by $(g, \eta, \rho) \mapsto -\eta$ (Exercise 10.4.2). Since K is the sum of the latter and the momentum mapping J^V for the G_m-action on V, it is a momentum mapping for the direct product. Finally, G_m-equivariance follows from that of J^V and G-invariance is obvious.

2. To prove that $J^\times$ is a momentum mapping for the G-action $\Psi^\times$, we show that

$$\omega^\times\left(A_*^\times, Z\right) = -Z\left(J_A^\times\right) \tag{10.4.10}$$

for all $A \in \mathfrak{g}$ and all vector fields Z on $\mathfrak{E}^\times$. Here, $A_*^\times$ denotes the Killing vector field generated by A under $\Psi^\times$. For $(g, \eta, \rho, v) \in \mathfrak{E}^\times$, we find

$$\left(A_*^\times\right)_{(g,\eta,\rho,v)} = \left(\mathrm{L}'_g\left(\mathrm{Ad}\left(g^{-1}\right)A\right), 0, 0, 0\right).$$

Hence, writing $Z_{(g,\eta,\rho,v)} = (\mathrm{L}'_g B, \xi, \sigma, u)$ and using (10.4.8), we compute

$$\omega^\times_{(g,\eta,\rho,v)}\left(A_*^\times, Z\right) = -\left\langle \xi + \sigma, \mathrm{Ad}\left(g^{-1}\right)A\right\rangle - \left\langle \mu + \eta + \rho, \left[\mathrm{Ad}\left(g^{-1}\right)A, B\right]\right\rangle.$$

On the other hand,

$$\begin{aligned} Z_{(g,\eta,\rho,v)}\left(J_A^\times\right) &= \frac{\mathrm{d}}{\mathrm{d}t}\Big\restriction_0 \left\langle \mathrm{Ad}^*\left(g\exp(tB)\right)\left(\mu + \eta + \rho + t(\xi + \sigma)\right), A\right\rangle \\ &= \left\langle \mu + \eta + \rho, \left[\mathrm{Ad}\left(g^{-1}\right)A, B\right]\right\rangle + \left\langle \xi + \sigma, \mathrm{Ad}\left(g^{-1}\right)A\right\rangle. \end{aligned}$$

This proves (10.4.10) and, hence, that $J^\times$ is a momentum mapping for $\Psi^\times$. Equivariance is obvious and G_m-invariance follows at once from $G_m \subset G_\mu$. □

[13] Induced by the injection $\mathfrak{g}_m \to \mathfrak{g}$.

Obviously, K is a submersion and the action of G_m on $\mathfrak{E}^\times$ is free and proper. Hence, we can apply regular reduction to construct the symplectic manifold $K^{-1}(0)/G_m$.

Lemma 10.4.2 *The neighbourhood $\mathfrak{E}^\times$ can be chosen so that there exist G_m-invariant open neighbourhoods $\tilde{\mathfrak{m}}^*$ of the origin in $\mathfrak{m}^*$ and $\tilde{V}$ of the origin in V such that the mapping*

$$\tilde{E} := G \times_{G_m} \left(\tilde{\mathfrak{m}}^* \times \tilde{V}\right) \to K^{-1}(0)/G_m, \qquad \left[(g, \rho, v)\right] \mapsto \left[\left(g, J^V(v), \rho, v\right)\right],$$

is a diffeomorphism.

Proof For given $\mathfrak{E}^\times$, there exist $\tilde{\mathfrak{m}}^*$ and $\tilde{V}$ such that the mapping

$$G \times \tilde{\mathfrak{m}}^* \times \tilde{V} \to \mathfrak{E}^\times, \qquad (g, \rho, v) \mapsto \left(g, J^V(v), \rho, v\right), \tag{10.4.11}$$

is defined. By construction, this mapping is bijective onto an open subset of $K^{-1}(0)$. Differentiability (in both directions) follows from the fact that $K^{-1}(0)$ is an embedded submanifold. Since $K^{-1}(0)$ carries the relative topology induced from $\mathfrak{E}^\times$, the latter can be shrunk so that the mapping (10.4.11) becomes a diffeomorphism onto $K^{-1}(0)$. Since J^V is G_m-equivariant, so is this diffeomorphism. Hence, it descends to a bijection of the quotients. Since the natural projections involved are submersions, this bijection is in fact a diffeomorphism. □

Via the diffeomorphism of Lemma 10.4.2, the symplectic form of $K^{-1}(0)/G_m$ inherited from $\omega^\times$ induces a symplectic form $\tilde{\omega}$ on $\tilde{E}$. This form is uniquely determined by the relation

$$\pi_0^* \tilde{\omega} = \iota_0^* \omega^\times,$$

where $\iota_0 : K^{-1}(0) \to \mathfrak{E}^\times$ is the natural inclusion mapping and $\pi_0 : K^{-1}(0) \to \tilde{E}$ is the submersion obtained by composing the natural projection $K^{-1}(0) \to K^{-1}(0)/G_m$ with the inverse of the diffeomorphism of Lemma 10.4.2. To derive an explicit formula for $\tilde{\omega}$, we denote the natural projection $G \times \tilde{\mathfrak{m}}^* \times \tilde{V} \to \tilde{E}$ by π and write tangent vectors of $\tilde{E}$ at $[(g, \rho, v)]$ in the form $\pi'(\mathrm{L}'_g A, \sigma, u)$, where $A \in \mathfrak{g}$, $\sigma \in \mathfrak{m}^*$ and $u \in V$. Then, from (10.4.8) we read off

$$\begin{aligned} &\tilde{\omega}_{[(g,\rho,v)]}\left(\pi'\left(\mathrm{L}'_g A_1, \sigma_1, u_1\right), \pi'\left(\mathrm{L}'_g A_2, \sigma_2, u_2\right)\right) \\ &\quad = \left\langle \sigma_1 + \left(J^V\right)'_v(u_1), A_2 \right\rangle - \left\langle \sigma_2 + \left(J^V\right)'_v(u_2), A_1 \right\rangle \\ &\qquad - \left\langle \mu + \rho + J^V(v), [A_1, A_2] \right\rangle + \omega_m(u_1, u_2). \end{aligned} \tag{10.4.12}$$

Moreover, since K is G-invariant, so is $K^{-1}(0)$. Hence, $\Psi^\times$ restricts to an action Ψ^0 of G on $K^{-1}(0)$. Since Ψ^0 commutes with the G_m-action and since π_0 is a submersion, Ψ^0 descends to an action $\tilde{\Psi}$ of G on $\tilde{E}$, explicitly given by

$$\tilde{\Psi}_h\left(\left[(g, \rho, v)\right]\right) = \left[(hg, \rho, v)\right].$$

By construction, for every $g \in G$, we have

$$\tilde{\Psi}_g \circ \pi_0 = \pi_0 \circ \Psi_g^0, \qquad \iota_0 \circ \Psi_g^0 = \Psi_g^\times \circ \iota_0. \tag{10.4.13}$$

Finally, since $J^\times$ is G_m-invariant, and again since π_0 is a submersion, $J^\times$ induces a smooth mapping $\tilde{J} : \tilde{E} \to \mathfrak{g}^*$ by

$$\tilde{J} \circ \pi_0 = J^\times \circ \iota_0. \tag{10.4.14}$$

Explicitly, one finds

$$\tilde{J}\big(\big[(g, \rho, v)\big]\big) = \mathrm{Ad}^*(g)\big(\mu + \rho + J^V(v)\big). \tag{10.4.15}$$

Lemma 10.4.3 *$(\tilde{E}, \tilde{\omega}, \tilde{\Psi})$ is a Hamiltonian G-manifold with equivariant momentum mapping $\tilde{J}$.*

Proof In the proof, we use the relations (10.4.12), (10.4.13) and (10.4.14) without further notice. First, we show that the action $\tilde{\Psi}$ is symplectic: for $g \in G$, we find

$$\pi_0^*\big(\tilde{\Psi}_g^*\tilde{\omega}\big) = \Psi_g^{0*}\big(\pi_0^*\tilde{\omega}\big) = \Psi_g^{0*}\big(\iota_0^*\omega\big) = \iota_0^*\big(\Psi_g^{\times*}\omega\big) = \iota_0^*\omega = \pi_0^*\tilde{\omega}.$$

Since π_0 is a submersion, it follows that $\tilde{\Psi}_g^*\tilde{\omega} = \tilde{\omega}$. Next, we show that $\tilde{J}$ is a momentum mapping for $\tilde{\Psi}$: for $A \in \mathfrak{g}$, denote the Killing vector fields generated by A under the actions $\Psi^\times$, Ψ^0 and $\tilde{\Psi}$ by, respectively, $A_*^\times$, A_*^0 and $\tilde{A}_*$. According to Proposition 6.2.4/2, we have

$$\tilde{A}_* \circ \pi_0 = \pi_0' \circ A_*^0, \qquad A_*^\times \circ \iota_0 = \iota_0' \circ A_*^0.$$

Using this, we calculate

$$\pi_0^*(\tilde{A}_* \lrcorner\, \tilde{\omega}) = A_*^0 \lrcorner\, \pi_0^*\tilde{\omega} = A_*^0 \lrcorner\, \iota_0^*\omega^\times = \iota_0^*\big(A_*^\times \lrcorner\, \omega^\times\big).$$

Since $J^\times$ is a momentum mapping for $\Psi^\times$, the right hand side equals $-\iota_0^*\mathrm{d}J_A^\times$ and hence $-\pi_0^*\mathrm{d}\tilde{J}_A$. Since π_0 is a surjective submersion, we conclude $\tilde{A}_* \lrcorner\, \tilde{\omega} = -\mathrm{d}\tilde{J}_A$. Finally, we show that $\tilde{J}$ is equivariant: for $g \in G$ we find

$$\tilde{J} \circ \tilde{\Psi}_g \circ \pi_0 = \tilde{J} \circ \pi_0 \circ \Psi_g^0 = J^\times \circ \iota_0 \circ \Psi_g^0 = J^\times \circ \Psi_g^\times \circ \iota_0.$$

Since $J^\times$ is equivariant, the right hand side equals

$$\mathrm{Ad}^*(g) \circ J^\times \circ \iota_0 = \mathrm{Ad}^*(g) \circ \tilde{J} \circ \pi_0.$$

Since π_0 is surjective, this yields equivariance. □

Theorem 10.4.4 (Symplectic Tubular Neighbourhood Theorem) *Let (M, ω, Ψ) be a symplectic G-manifold with proper action and equivariant momentum mapping J.*

1. *For every $m \in M$, there exists a G-invariant open neighbourhood U of $G \cdot m$ in M and a G-equivariant symplectomorphism $\chi : U \to \tilde{E}$ such that $\chi(m) = [(\mathbb{1}, 0, 0)]$.*
2. *$\tilde{J} \circ \chi$ is a momentum mapping for Ψ. If G is connected, then $\tilde{J} \circ \chi = J_{\restriction U}$.*

As a result, $(\tilde{E}, \tilde{\omega}, \tilde{\Psi}, \tilde{J})$ yields a local normal form for the Hamiltonian G-manifold (M, ω, Ψ, J) near the orbit $G \cdot m$. It will be referred to as a symplectic tubular neighbourhood of $G \cdot m$. Accordingly, χ will be referred to as a tube symplectomorphism. The representation (10.4.15) of the momentum mapping J so obtained is called the Marle-Guillemin-Sternberg normal form of J, see [190–192] and [117].

Proof 1. By the Witt-Artin decomposition (10.2.13), there exists a G_m-equivariant injective linear mapping $\lambda : \mathfrak{m}^* \oplus V \to \mathrm{T}_m M$ onto a vector space complement of $\mathrm{T}_m(G \cdot m)$ such that ω_m is given by (10.2.14). Given λ, Theorem 6.4.3 and Remark 6.5.8 yield a G-equivariant diffeomorphism

$$\chi : U \to \tilde{E} \subset G \times_{G_m} \big(\mathfrak{m}^* \oplus V\big) \tag{10.4.16}$$

satisfying $\chi(m) = [(\mathbb{1}, 0, 0)]$ and (6.5.3), which in the present situation reads

$$\big(\chi^{-1}\big)'_{[(e,0,0)]}\big(\pi'(A, \sigma, u)\big) = (A_*)_m + \lambda(\sigma, u), \tag{10.4.17}$$

where $\pi : G \times (\tilde{\mathfrak{m}}^* \times \tilde{V}) \to \tilde{E}$ is the natural projection. Using (10.4.17), (10.2.14) and (10.4.12), as well as $(J^V)'_0 = 0$, we calculate

$$\begin{aligned}
&\big((\chi^{-1})^*\omega\big)_{[(\mathbb{1},0,0)]}\big(\pi'(A_1, \sigma_1, u_1), \pi'(A_2, \sigma_2, u_2)\big) \\
&\quad = \omega_m\big((A_{1*})_m + \lambda(\sigma_1, u_1), (A_{2*})_m + \lambda(\sigma_2, u_2)\big) \\
&\quad = \langle \sigma_1, A_2 \rangle - \langle \sigma_2, A_1 \rangle - \big\langle \mu, [A_1, B_1]\big\rangle + \omega^V(u_1, u_2) \\
&\quad = \tilde{\omega}_{[(\mathbb{1},0,0)]}\big(\pi'(A_1, \sigma_1, u_1), \pi'(A_2, \sigma_2, u_2)\big),
\end{aligned}$$

where $\mu = J(m)$. By invariance, then ω and $\chi^*\tilde{\omega}$ coincide on $G \cdot m$ and the Equivariant Darboux Theorem 8.6.3 yields that U can be shrunk and χ can be modified so that it becomes symplectic. Finally, we adjust U, $\tilde{\mathfrak{m}}^*$ and $\tilde{V}$ in such a way that χ becomes surjective and hence a symplectomorphism.

2. For $A \in \mathfrak{g}$, let A_* and $\tilde{A}_*$ denote the Killing vector fields generated by A under the actions Ψ and $\tilde{\Psi}$, respectively. Using Proposition 6.2.4/2, we calculate

$$A_* \lrcorner\, \omega = A_* \lrcorner \big(\chi^*\tilde{\omega}\big) = \chi^*\big((\chi_* A_*) \lrcorner\, \tilde{\omega}\big) = \chi^*(\tilde{A}_* \lrcorner\, \tilde{\omega}) = -\chi^* \mathrm{d}\tilde{J}_A = -\mathrm{d}(\tilde{J} \circ \chi)_A .$$

Since $\tilde{J} \circ \chi(m) = \tilde{J}([(\mathbb{1}, 0, 0)]) = \mu = J(m)$ and since the difference of two momentum mappings is locally constant, we conclude that $\tilde{J} \circ \chi = J$ provided U is connected, which can always be achieved if G is connected. □

Exercises

10.4.1 Write down the proof of Lemma 10.4.1 for the case $\mu = 0$.

10.4.2 Complete the proof of Lemma 10.4.1/1 by showing that the symplectic action of G_m on $(G \times \mathfrak{g}_m^* \times \mathfrak{m}^*, p_\mu^* \omega^{\mathscr{O}_\mu^-} + i_\mu^* \omega^{\mathrm{T}^*G})$ is Hamiltonian with momentum mapping induced from that of the symplectic action $\mathscr{R}$ on $(G \times \mathfrak{g}^*, \omega^{\mathrm{T}^*G})$ in the way described there.

10.5 Singular Symplectic Reduction

In this section, we generalize the Regular Reduction Theorem 10.3.1 to the case of a non-free group action. Here, as we know, several orbit types labelled by conjugacy classes of stabilizers can occur and, consequently, the quotient M_μ/G_μ is a union of strata. In what follows, we will describe the symplectic structure of these strata using the method of point reduction. The main tool is the Symplectic Tubular Neighbourhood Theorem 10.4.4. We will close this section with a brief comment on how the strata fit together to form a stratified symplectic space. Pioneering work in this field was done by Arms, Cushman and Gotay, see [10], and by Sjamaar and Lerman, see [275], where the case of the zero-level set was worked out. For a detailed discussion of all aspects of the general case, which would go beyond the scope of this book, we refer to the book of Ortega and Ratiu, see [232]. There, the reader can also find an exhaustive list of references. We also refer to Huebschmann, see [142], who has worked out singular reduction for the case of Kähler manifolds.

Let (M, ω, Ψ, J) be a Hamiltonian G-manifold with proper action and equivariant momentum mapping. Below, we use the notation introduced in Sect. 6.1. In particular, let $\hat{M} = M/G$ and let $\pi : M \to \hat{M}$ denote the natural projection. For a given coadjoint orbit $\mathscr{O} \subset \mathfrak{g}^*$, denote

$$\hat{M}_{\mathscr{O}} := \pi\big(J^{-1}(\mathscr{O})\big).$$

For a given orbit type $[H]$, the connected components of the subset

$$\hat{M}_{[H]} \cap \hat{M}_{\mathscr{O}}$$

of orbits of orbit type $[H]$ and momentum type $\mathscr{O}$ will be referred to as the orbit-momentum type strata of $\hat{M}$ or the reduced phase spaces. They will be denoted by $\hat{M}_\tau$, where the set of labels τ is denoted by $\mathtt{T}$. For given $\tau \in \mathtt{T}$ and μ being an element of the underlying coadjoint orbit, we define

$$M_{\tau,\mu} := \pi^{-1}(\hat{M}_\tau) \cap M_\mu,$$

where $M_\mu = J^{-1}(\mu)$ as before. Let $\iota_{\tau,\mu} : M_{\tau,\mu} \to M$ and $\pi_{\tau,\mu} : M_{\tau,\mu} \to \hat{M}_\tau$ denote the natural inclusion mapping and the natural projection induced by π, respectively. Note that, by equivariance of J, two points in $M_{\tau,\mu}$ are conjugate under G iff they are conjugate under G_μ. Consequently, $\pi_{\tau,\mu}$ means factorization with respect to G_μ. We also note that G_μ need not be connected and, thus, $M_{\tau,\mu}$ need not be connected as well. It can be obtained via the G_μ-action from one of its connected components though.

Lemma 10.5.1 *Let (M, ω, Ψ, J) be a Hamiltonian G-manifold with proper action and equivariant momentum mapping, let $m \in M$ and denote $\mu = J(m)$. Let $(\tilde{E}, \tilde{\omega}, \tilde{\Psi}, \tilde{J})$ be a symplectic tubular neighbourhood of the orbit $G \cdot m$ at m. There exists a G_μ-invariant open neighbourhood $\hat{E}$ of the orbit $G_\mu \cdot [(\mathbb{1}, 0, 0)]$ such that*

$$\tilde{J}^{-1}(\mu) \cap \hat{E} = \big\{\big[(g, 0, v)\big] \in \tilde{E} : g \in G_\mu,\ J^V(v) = 0\big\}. \tag{10.5.1}$$

Proof By (10.4.15), a point $[(g,\rho,v)] \in \tilde{E}$ belongs to $\tilde{J}^{-1}(\mu)$ iff

$$\rho + J^V(v) = \mathrm{Ad}^*\big(g^{-1}\big)\mu - \mu. \tag{10.5.2}$$

We choose a G_m-invariant scalar product in $\mathfrak{g}^*$ and use this to embed $\mathfrak{m}^*$, $\mathfrak{g}_m^*$ and $\mathfrak{g}_\mu^*$ into $\mathfrak{g}^*$ as the orthogonal complements of the annihilators $\mathfrak{m}^0$, $\mathfrak{g}_m^0$ and $\mathfrak{g}_\mu^0$. Then, $\mathfrak{g}_\mu^* = \mathfrak{m}^* \oplus \mathfrak{g}_m^*$ and hence the left hand side of (10.5.2) belongs to the subspace $\mathfrak{g}_\mu^*$. Thus, in order to prove that $\hat{E}$ exists, it suffices to show that there exists an open neighbourhood W of G_μ in G, invariant under left translation by G_μ and under right translation by G_m, such that for any $g \in W$, the condition $\mathrm{Ad}^*(g^{-1})\mu - \mu \in \mathfrak{g}_\mu^*$ implies $g \in G_\mu$. Indeed, then

$$\hat{E} := W \times_{G_m} \times \big(\tilde{\mathfrak{m}}^* \times \tilde{V}\big)$$

is G_μ-invariant and satisfies (10.5.1). To prove the existence of W, let $\mathrm{pr}^0 : \mathfrak{g}^* \to \mathfrak{g}_\mu^0$ denote orthogonal projection with respect to the G_m-invariant scalar product and consider the mapping

$$f : \mathscr{O}_\mu \to \mathfrak{g}_\mu^0, \qquad f(\eta) := \mathrm{pr}^0(\eta - \mu).$$

Using Formula (6.2.3) and $\mathrm{ad}^*(A)\mu \in \mathfrak{g}_\mu^0$, for $A \in \mathfrak{g}$ we calculate

$$f'_\mu(A_{*\mu}) = \mathrm{pr}^0(A_{*\mu}) = \mathrm{pr}^0\big(\mathrm{ad}^*(A)\mu\big) = \mathrm{ad}^*(A)\mu = A_{*\mu},$$

where A_* denotes the Killing vector field of the coadjoint action. Hence, f'_μ is injective. For dimensional reasons, it is bijective then and the Inverse Mapping Theorem 1.5.7 yields an open neighbourhood $\tilde{W}$ of μ in $\mathscr{O}_\mu$ where f is injective, that is, where $f(\eta) = 0$ implies $\eta = \mu$ for all $\eta \in \tilde{W}$. Since pr^0 is G_m-equivariant, $\tilde{W}$ can be chosen to be G_m-invariant. Then,

$$W := \big\{a \in G : \mathrm{Ad}^*\big(a^{-1}\big)\mu \in \tilde{W}\big\}$$

(the preimage of $\tilde{W}$ under the coadjoint orbit mapping of μ) has the desired properties. □

Remark 10.5.2 If we extend the twisted product to the case of topological spaces, we can rewrite (10.5.1) in the form

$$\tilde{J}^{-1}(\mu) \cap \hat{E} = G_\mu \times_{G_m} \big(\{0\} \times \big(\tilde{V} \cap \big(J^V\big)^{-1}(0)\big)\big).$$

Here, $(J^V)^{-1}(0)$ is not necessarily a manifold. We note that restriction to $\hat{E}$ is important, because the intersection of the coadjoint orbit of μ with the affine subspace $\mu + (\mathfrak{m}^* \oplus \mathfrak{g}_m^*)$ may contain points different from μ, so that (10.5.2) may have solutions with g lying outside the neighbourhood W of G_μ constructed in the above proof.

Lemma 10.5.3 *Under the assumptions of Lemma 10.5.1 we have*

$$\big(\tilde{J}^{-1}(\mu) \cap \hat{E}\big)_H = \mathrm{N}_{G_\mu}(H) \times_H \big(\{0\} \times \tilde{V}^H\big), \tag{10.5.3}$$

$$\big(\tilde{J}^{-1}(\mu) \cap \hat{E}\big)_{[H]} = G_\mu \times_H \big(\{0\} \times \tilde{V}^H\big), \tag{10.5.4}$$

with $H \equiv G_m$ and $\tilde{V}^H$ denoting the H-invariant elements of $\tilde{V}$.

Proof First, we show that

$$\big((J^V)^{-1}(0)\big)^H = V^H. \tag{10.5.5}$$

For that purpose, it suffices to show that $J^V(v) = 0$ for all elements $v \in V^H$. Indeed, for any element A of the Lie algebra of H, we have

$$\langle J^V(v), A\rangle = \frac{1}{2}\omega\big(v, \mathrm{Hess}_m(A_*)v\big) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0}\,\omega\big(v, (\Psi_{\exp(tA)})'_m v\big) = 0,$$

for any $v \in V^H$. To prove (10.5.3), let $[(g, 0, v)] \in \tilde{J}^{-1}(\mu) \cap \hat{E}$. If $G_{[(g,0,v)]} = H$, then $[(ag, 0, v)] = [(g, 0, v)]$ for all $a \in H$. This means that there exists an element $b \in H$ such that

$$(ag, 0, v) = \big(gb^{-1}, 0, (\Psi_b)'_m v\big).$$

We read off that $b = g^{-1}a^{-1}g$, that is, $g^{-1}Hg \subset H$ and hence $g \in \mathrm{N}_{G_\mu}(H)$. Moreover, by compactness of H, then $g^{-1}Hg = H$, so that b ranges through all of H and hence v is H-invariant. Conversely, if $g \in \mathrm{N}_{G_\mu}(H)$ and $v \in \tilde{V}^H$, then (10.5.5) implies $J^V(v) = 0$ and hence $[(g, 0, v)] \in \tilde{J}^{-1} \cap \hat{E}$. The inclusion $H \subset G_{[(g,0,v)]}$ is obvious. If, conversely, $a \in G_{[(g,0,v)]}$, then $ag = gb^{-1}$ for some $b \in H$. Since $g \in \mathrm{N}_{G_\mu}(H)$, then $a \in H$.

Equation (10.5.4) follows by a similar argument: if $G_{[(g,0,v)]} = aHa^{-1}$ for some $a \in G$, then for every $h \in H$ there exists $b \in H$ such that

$$\big(aha^{-1}g, 0, v\big) = \big(gb^{-1}, 0, (\Psi_b)'_m(v)\big).$$

Then, $b = g^{-1}ah^{-1}a^{-1}g$, so that $g^{-1}aHa^{-1}g \subset H$ and compactness of H implies that b ranges through all of H. Hence, $v \in \tilde{V}^H$. Conversely, by (10.5.5), every point $[(g, 0, v)]$ of the right hand side belongs to $\tilde{J}^{-1}(\mu) \cap \hat{E}$. Moreover, $gHg^{-1} \subset G_{[(g,0,v)]}$. If, on the other hand, $a \in G_{[(g,0,v)]}$, then $ag = gb^{-1}$ for some $b \in H$ and hence $a \in gHg^{-1}$. □

Let $(\tilde{E}, \tilde{\omega}, \tilde{\Psi}, \tilde{J})$ be a symplectic tubular neighbourhood of $G \cdot m$ with tube symplectomorphism $\chi : U \to \tilde{E}$. Denoting $\hat{U} = \chi^{-1}(\hat{E})$, from (10.5.3) and (10.5.4), we read off

$$M_H \cap M_\mu \cap \hat{U} = \chi^{-1}\big(\mathrm{N}_{G_\mu}(H) \times_H \big(\{0\} \times \tilde{V}^H\big)\big) \cong \mathrm{N}_{G_\mu}(H)/H \times \tilde{V}^H, \tag{10.5.6}$$

$$M_{[H]} \cap M_\mu \cap \hat{U} = \chi^{-1}\big(G_\mu \times_H \big(\{0\} \times \tilde{V}^H\big)\big) \cong G_\mu/H \times \tilde{V}^H. \tag{10.5.7}$$

By Lemma 7.4.6, $V^H \subset V$ is a symplectic vector space with symplectic form given by the restriction of ω^V to V^H. By Formula (10.5.7), V^H is a model space for the strata of the G-action on $\tilde{E}$ of isotropy type H. This is the key observation for the proof of the following theorem.

Theorem 10.5.4 (Singular Reduction Theorem) *Let (M, ω, Ψ, J) be a Hamiltonian G-manifold with proper action and equivariant momentum mapping. Let $\tau \in \mathrm{T}$ and let $\mathscr{O}$ be the corresponding coadjoint orbit. For every $\mu \in \mathscr{O}$,*

1. *the subset* $M_{\tau,\mu}$ *is an embedded submanifold of* M,
2. *there exists a unique smooth manifold structure on* $\hat{M}_\tau$ *such that the natural projection* $\pi_{\tau,\mu} : M_{\tau,\mu} \to \hat{M}_\tau$ *is a submersion,*
3. *there exists a unique symplectic form* ω^τ *on* $\hat{M}_\tau$ *such that*

$$\pi^*_{\tau,\mu}\omega^\tau = \iota^*_{\tau,\mu}\omega. \tag{10.5.8}$$

The smooth structure and the symplectic structure so induced on $\hat{M}_\tau$ *do not depend on the choice of* μ.

Proof 1. Let $\mu \in \mathscr{O}$ be given. By the Symplectic Tubular Neighbourhood Theorem 10.4.4, for every $m \in M_{\tau,\mu}$, there exists a symplectic tubular neighbourhood $(\tilde{E}, \tilde{\omega}, \tilde{\Psi}, \tilde{J})$ of $G \cdot m$ and a tube symplectomorphism $\chi : U \to \tilde{E}$. Moreover, by Lemma 10.5.1, there exists an open neighbourhood $\hat{E}$ of the orbit $G_\mu \cdot [(\mathbb{1}, 0, 0)]$ in $\tilde{E}$ satisfying (10.5.1). Then, $\hat{U} := \chi^{-1}(\hat{E})$ is an open neighbourhood of $G_\mu \cdot m$ in M. We show that[14]

$$M_{\tau,\mu} \cap \hat{U} = \chi^{-1}\big(G_\mu \times_H \big(\{0\} \times \tilde{V}^H\big)\big). \tag{10.5.9}$$

Consider the set $W := \chi^{-1}(\{[(\mathbb{1}, 0, v)] : v \in \tilde{V}^H\})$. Since W is a connected subset of $M_{[H]} \cap M_\mu$ which intersects $M_{\tau,\mu}$ and since the latter consists of connected components of $M_{[H]} \cap M_\mu$, we have $W \subset M_{\tau,\mu} \cap \hat{U}$. Then, $G_\mu \cdot W \subset M_{\tau,\mu} \cap \hat{U}$. By (10.5.4), $G_\mu \cdot W = M_{[H]} \cap M_\mu \cap \hat{U}$ and hence $M_{\tau,\mu} \cap \hat{U} = M_{[H]} \cap M_\mu \cap \hat{U}$. Then, (10.5.9) follows from Eq. (10.5.4). Since $G_\mu \times_H (\{0\} \times \tilde{V}^H)$ is an embedded submanifold of $\hat{E}$, it follows that $M_{\tau,\mu} \cap \hat{U}$ is an embedded submanifold of M. Finally, since $M_{\tau,\mu}$ is obtained via the G_μ-action from one of its connected components and since along a connected component the dimensions of V and V^H cannot change, for all $m \in M_{\tau,\mu}$, the subspaces V^H have the same dimension. Then, the assertion follows from Remark 1.7.4.

2. For $m \in M_{\tau,\mu}$, by identifying $\tilde{V}$ with the subset $H \times_H (\{0\} \times \tilde{V})$ of $\tilde{E}$, we obtain a continuous mapping

$$\varphi_m : \tilde{V}^H \longrightarrow G_\mu \times_H \big(\{0\} \times \tilde{V}^H\big) \xrightarrow{\chi^{-1}} M_{\tau,\mu} \xrightarrow{\pi_{\tau,\mu}} \hat{M}_\tau .$$

Since J is equivariant, G-orbits in M intersect $M_{\tau,\mu}$ in G_μ-orbits. Hence, φ_m is injective. Since for an open subset W of $\tilde{V}^H$, the subset $\chi^{-1}(G_\mu \times_H (\{0\} \times W))$ is open in $M_{\tau,\mu}$ and since $\pi_{\tau,\mu}$ is open, φ_m is also open and hence a homeomorphism onto its image. By composing its inverse with a chosen linear isomorphism $V^H \to \mathbb{R}^r$, where $r = \dim(V^H)$, we obtain a local chart κ_m on $\hat{M}_\tau$. Up to the diffeomorphism

$$G_\mu \times_H \big(\{0\} \times \tilde{V}^H\big) \cong G_\mu/H \times \tilde{V}^H, \tag{10.5.10}$$

given by (10.5.7), the transition mapping between the charts κ_{m_1} and κ_{m_2} is a composition of linear transformations with the restriction of the natural projection to the

[14] As before, for simplicity we denote $H \equiv G_m$.

second factor of $G_\mu/H \times \tilde{V}^H$ to a certain submanifold. Hence, the local charts κ_m, $m \in M_{\tau,\mu}$, define a smooth structure on $\hat{M}_\tau$. Since via χ and (10.5.10), the natural projection $\pi_{\tau,\mu} : M_{\tau,\mu} \to \hat{M}_\tau$ corresponds locally to a natural projection in a direct product, it is a submersion.

3. Since $\pi_{\tau,\mu}$ is a surjective submersion, any two local 2-forms on $\hat{M}_\tau$ satisfying (10.5.8) must coincide on their common domain. Therefore, it suffices to prove that ω^τ exists locally. Thus, let $m \in M_{\tau,\mu}$ and consider a symplectic tubular neighbourhood of $G \cdot m$. In the notation introduced under point 1, we will construct ω^τ on the open neighbourhood $\pi_{\tau,\mu}(M_{\tau,\mu} \cap \hat{U})$ of $\pi_{\tau,\mu}(m)$ in $\hat{M}_\tau$. Via χ and the isomorphism (10.5.10), this neighbourhood is identified with $\tilde{V}^H$, the subset $M_{\tau,\mu} \cap \hat{U}$ is identified with $G_\mu/H \times \tilde{V}^H$, the projection $M_{\tau,\mu} \cap \hat{U} \to \pi_{\tau,\mu}(M_{\tau,\mu} \cap \hat{U})$ corresponds to the natural projection $\mathrm{pr}_2 : G_\mu/H \times \tilde{V}^H \to \tilde{V}^H$ and the restriction of $\iota^*_{\tau,\mu}\omega$ to $M_{\tau,\mu} \cap \hat{U}$ corresponds to $j^*\tilde{\omega}$, where

$$j : G_\mu/H \times \tilde{V}^H \to \tilde{E}, \qquad j\big([g], v\big) := \big[(g, 0, v)\big].$$

Let $\pi_H : G_\mu \to G_\mu/H$ and $\tilde{\pi} : G \times \tilde{\mathfrak{m}}^* \times \tilde{V} \to \tilde{E}$ be the natural projections. Writing tangent vectors of $G_\mu/H \times \tilde{V}^H$ at $([g], v)$ in the form $(\pi'_H \circ \mathrm{L}'_g A, u)$ with $A \in \mathfrak{g}_\mu$ and $u \in V^H$ and using (10.4.12), we calculate

$$\begin{aligned}
&\big(j^*\tilde{\omega}\big)_{([g],v)}\big(\big(\pi'_H \circ \mathrm{L}'_g A_1, u_1\big), \big(\pi'_H \circ \mathrm{L}'_g A_2, u_2\big)\big)\\
&\quad = \tilde{\omega}_{[(g,0,v)]}\big(\tilde{\pi}'\big(\mathrm{L}'_g A_1, 0, u_1\big), \tilde{\pi}'\big(\mathrm{L}'_g A_2, 0, u_2\big)\big)\\
&\quad = \omega_m(u_1, u_2)\\
&\quad = \big(\mathrm{pr}_2^*\, \omega_m\big)_{([g],v)}\big(\big(\pi'_H \circ \mathrm{L}'_g A_1, u_1\big), \big(\pi'_H \circ \mathrm{L}'_g A_2, u_2\big)\big).
\end{aligned}$$

This shows that ω^τ exists on the open subset $\pi_{\tau,\mu}(M_{\tau,\mu} \cap \hat{U})$ of $\hat{M}_\tau$, where under the above identification of this subset with $\tilde{V}^H$ it is given by the restriction of ω_m to the subspace V^H of V. Since, by Lemma 7.4.6, this subspace is symplectic, ω^τ is symplectic.

It remains to show that the smooth structure on $\hat{M}_\tau$ and the form ω^τ so constructed do not depend on the choice of μ. For every $\tilde{\mu} \in \mathscr{O}$, there exists $g \in G$ such that $\mu = \mathrm{Ad}^*(g)\tilde{\mu}$. By restriction, Ψ_g induces a diffeomorphism $\varphi : M_{\tau,\tilde{\mu}} \to M_{\tau,\mu}$ which projects to the identical mapping $\mathrm{id}_{\hat{M}_\tau}$ of $\hat{M}_\tau$. Since $\pi_{\tau,\mu}$ and $\pi_{\tau,\tilde{\mu}}$ are submersions with respect to the smooth structure on $\hat{M}_\tau$ induced from $M_{\tau,\mu}$ and $M_{\tau,\tilde{\mu}}$, respectively, it follows that $\mathrm{id}_{\hat{M}_\tau}$ is a diffeomorphism. Hence, these smooth structures coincide. Moreover, using $\pi_{\tau,\mu} \circ \varphi = \pi_{\tau,\tilde{\mu}}$ and $\iota_{\tau,\mu} \circ \varphi = \Psi_g \circ \iota_{\tau,\tilde{\mu}}$, we obtain

$$\pi^*_{\tau,\tilde{\mu}}\omega^\tau = \iota^*_{\tau,\tilde{\mu}}\omega.$$

It follows that ω^τ coincides with the 2-form induced from $M_{\tau,\tilde{\mu}}$. □

Remark 10.5.5

1. Let us note the following consequences of the proof of Theorem 10.5.4.

(a) Via a symplectic tubular neighbourhood, every $m \in M_{\tau,\mu}$ induces a local symplectomorphism from an open neighbourhood of $\pi(m)$ in $(\hat{M}_\tau, \omega^\tau)$ onto $(\tilde{V}^H, \omega_m)$.
(b) The natural projection $\pi_{\tau,\mu} : M_{\tau,\mu} \to \hat{M}_\tau$ is a locally trivial fibre bundle with typical fibre G_μ/H, where $m \in M_{\tau,\mu}$ is arbitrary but fixed.

2. The argument of the proof of Theorem 10.5.4 shows that $M_H \cap M_\mu$ is a union of embedded submanifolds of M and that $\hat{M}_{[H]} \cap \hat{M}_{\mathscr{O}}$ is a union of symplectic manifolds. In both cases, these manifolds may have different dimensions, cf. Remark 6.6.2/1. Since, in practice, the orbit-momentum type strata are hard to find directly, the method of choice is to analyse the subsets $M_H \cap M_\mu$ for several combinations of H and μ in order to determine $\hat{M}_{[H]} \cap \hat{M}_{\mathscr{O}}$ and to read off the corresponding orbit-momentum type strata afterwards. In this context, let us note that, for given H and μ, the subset $\pi(M_H \cap M_\mu)$ of $\hat{M}$ is a union of connected components of $\hat{M}_{[H]} \cap \hat{M}_{\mathscr{O}}$ but need not coincide with the latter. It coincides under the condition that every subgroup of G_μ which is conjugate to H in G is also conjugate to H in G_μ.
3. We encourage the reader to write down the above proofs for the special case $\mu = 0$. This leads to some structural simplification.

Now, we are able to discuss singular reduction of a G-invariant Hamiltonian system (M, ω, h) with equivariant momentum mapping.

Proposition 10.5.6 *Let (M, ω, Ψ, J) be a Hamiltonian G-manifold with proper action and equivariant momentum mapping and let $h \in C^\infty(M)^G$. Let $\tau \in \mathbb{T}$, let $\mathscr{O}$ be the corresponding coadjoint orbit and let $\mu \in \mathscr{O}$.*

1. *$M_{\tau,\mu}$ is invariant under the flow of X_h and X_h restricts to a vector field $X_h^{\tau,\mu}$ on $M_{\tau,\mu}$ which is $\iota_{\tau,\mu}$-related to X_h,*
$$\iota'_{\tau,\mu} \circ X_h^{\tau,\mu} = X_h \circ \iota_{\tau,\mu}.$$
2. *The Hamiltonian h defines a smooth function h_τ on $\hat{M}_\tau$ by*
$$h_\tau \circ \pi_{\tau,\mu} = h \circ \iota_{\tau,\mu}.$$
The corresponding Hamiltonian vector field X_{h_τ} is $\pi_{\tau,\mu}$-related to $X_h^{\tau,\mu}$,
$$\pi'_{\tau,\mu} \circ X_h^{\tau,\mu} = X_{h_\tau} \circ \pi_{\tau,\mu}.$$

As a consequence, an invariant Hamiltonian system (M, ω, h) induces a uniquely determined reduced Hamiltonian system $(\hat{M}_\tau, \omega^\tau, \hat{h}^\tau)$ for each orbit-momentum type stratum τ of $\hat{M}$.

Proof By (6.7.3), the flow of X_h leaves invariant the stabilizers. Together with the Noether Theorem 10.1.9, this yields invariance of $M_{\tau,\mu}$. The rest of the proof is completely analogous to that of the regular case, cf. Proposition 10.3.3, and is therefore left to the reader. □

In the remainder of this section we will prove that the singular strata $\hat{M}_\tau$ can be obtained via regular reduction.[15] For that purpose, let H be a stabilizer, let Σ_H be a connected component of M_H and let $\iota_{\Sigma_H} : \Sigma_H \to M$ denote the natural inclusion mapping. Recall from Remark 6.6.2/6 that Σ_H is an embedded submanifold of M acted upon properly and freely by

$$\Gamma_{\Sigma_H} = \mathrm{N}^{\Sigma_H}/H,$$

where

$$\mathrm{N}^{\Sigma_H} := \mathrm{N}_{G^{\Sigma_H}}(H), \qquad G^{\Sigma_H} := \big\{g \in G : \Psi_g(\Sigma_H) \subset \Sigma_H\big\},$$

and that the quotient manifold $\Sigma_H/\Gamma_{\Sigma_H}$ may be naturally identified with a certain orbit type stratum $\hat{M}_\sigma$. Let $\mathfrak{n}$ denote the Lie algebra of $\mathrm{N}_G(H)$. Since $\mathrm{N}_G(H)_0 \subset \mathrm{N}^{\Sigma_H} \subset \mathrm{N}_G(H)$, this is also the Lie algebra of N^{Σ_H}. Recall that the dual mapping of the natural projection from $\mathfrak{n}$ onto the Lie algebra $\mathfrak{n}/\mathfrak{h}$ of Γ_{Σ_H} yields a natural identification of the dual vector space $(\mathfrak{n}/\mathfrak{h})^*$ with the annihilator $\mathfrak{h}^0$ of $\mathfrak{h}$ in $\mathfrak{n}$.[16]

Lemma 10.5.7 *Let (M, ω, Ψ, J) be a Hamiltonian G-manifold with proper action and equivariant momentum mapping, let H be a stabilizer and let Σ_H be a connected component of M_H. Let $j : \mathfrak{n} \to \mathfrak{g}$ denote the natural inclusion mapping.*

1. *$(\Sigma_H, \iota_{\Sigma_H}^*\omega)$ is a symplectic submanifold of (M, ω).*
2. *The action of Γ_{Σ_H} on Σ_H is symplectic. For every $\mu \in J(\Sigma_H)$, the mapping*

$$J^{H,\mu} : \Sigma_H \to \mathfrak{n}^*, \qquad J^{H,\mu}(m) := j^{\mathrm{T}} \circ \big(J \circ \iota_{\Sigma_H}(m) - \mu\big),$$

takes values in $\mathfrak{h}^0 \equiv (\mathfrak{n}/\mathfrak{h})^*$ and defines a momentum mapping[17] for this action.

Proof 1. According to Remark 6.6.2/6, Proposition 6.6.1/1 yields $\mathrm{T}_m\Sigma_H = (\mathrm{T}_m M)^H$ for all $m \in \Sigma_H$. Hence, the assertion follows from Proposition 7.4.6.

2. That the action is symplectic is obvious. For the proof of the assertions about $J^{H,\mu}$, we ignore the natural inclusion mapping $\mathfrak{h} \to \mathfrak{n}$. To see that the mapping $J^{H,\mu}$ takes values in $\mathfrak{h}^0$, let $A \in \mathfrak{h}$. Using the obvious identity

$$\big(j^{\mathrm{T}} \circ J\big)_A = J_{j(A)} \tag{10.5.11}$$

and the fact that $j(A)_*$ vanishes on Σ_H, we obtain

$$\mathrm{d}J_A^{H,\mu} = \mathrm{d}(J_{j(A)} \circ \iota_{\Sigma_H}) = -\iota_{\Sigma_H}^*\big(j(A)_* \lrcorner\, \omega\big) = 0.$$

Hence, $J_A^{H,\mu}$ is constant on Σ_H. Since $\mu \in J(\Sigma_H)$, there exists $m_0 \in \Sigma_H$ such that $J \circ \iota_{\Sigma_H}(m_0) = \mu$. Then, $J_A^{H,\mu}(m_0) = 0$ and hence

$$\big\langle J^{H,\mu}(m), A\big\rangle = J_A^{H,\mu}(m) = 0$$

---

[15] The authors of [232] call this statement Sjamaar's principle, because it first appeared in the thesis of Sjamaar [274] in the context of compact group actions and zero level reduction.

[16] Note that $\mathfrak{h}^0$ coincides with the subspace of H-invariant elements of the annihilator of $\mathfrak{h}$ in $\mathfrak{g}$.

[17] Which need not be equivariant, see Remark 10.5.8/2.

for all $m \in \Sigma_H$. Thus, $J^{H,\mu}$ takes values in $\mathfrak{h}^0$, indeed. To see that $J^{H,\mu}$ is a momentum mapping, let $A \in \mathfrak{n}$. First, we observe that the Killing vector field on Σ_H generated by $[A] \in \mathfrak{n}/\mathfrak{h}$ coincides with the Killing vector field $A_*^{\Sigma_H}$ generated by A under the action of N^{Σ_H} on Σ_H. Second, we observe that j is the Lie algebra homomorphism induced by the natural inclusion mapping $\mathrm{N}^{\Sigma_H} \to G$. Since the latter, together with ι_{Σ_H}, establishes a morphism of Lie group actions, Proposition 6.2.4/1 implies

$$\iota'_{\Sigma_H} \circ A_*^{\Sigma_H} = j(A)_* \circ j.$$

Using this and (10.5.11), we obtain

$$[A]_* \lrcorner \left(\iota_{\Sigma_H}^* \omega\right) = A_*^{\Sigma_H} \lrcorner \left(\iota_{\Sigma_H}^* \omega\right) = \iota_{\Sigma_H}^* \left(j(A)_* \lrcorner\, \omega\right) = -\mathrm{d}(J_{j(A)} \circ \iota_{\Sigma_H}) = -\mathrm{d}J_A^{H,\mu}.$$

□

Remark 10.5.8

1. By definition of $J^{H,\mu}$, $\mu \in J(\Sigma_H)$, we have
$$\left(J^{H,\mu}\right)^{-1}(0) = \Sigma_H \cap M_\mu. \tag{10.5.12}$$
2. The momentum mapping $J^{H,\mu}$ is not equivariant. Using the equivariance of J and j, we calculate the 1-cocycle defined by (10.1.6). Omitting the natural inclusion mapping $\iota_{\Sigma_H} : \Sigma_H \to M$, for $g \in \mathrm{N}^{\Sigma_H}$ and $m \in \Sigma_H$ we obtain
$$\begin{aligned}\sigma\left([g]\right) &= J^{H,\mu}\left(\Psi_g(m)\right) - \mathrm{Ad}^*\left([g]\right)\left(J^{H,\mu}(m)\right)\\ &= j^{\mathrm{T}}\left(J\left(\Psi_g(m)\right) - \mu\right) - \mathrm{Ad}^*\left([g]\right)\left(j^{\mathrm{T}}\left(J(m) - \mu\right)\right)\\ &= j^{\mathrm{T}}\left(\mathrm{Ad}^*(g)\mu - \mu\right).\end{aligned}$$

By Lemma 10.5.7 and Remark 10.5.8/2, if one modifies the coadjoint action of Γ_{Σ_H} on $(\mathfrak{n}/\mathfrak{h})^* \equiv \mathfrak{h}^0$ in the sense of (10.1.11), the momentum mapping $J^{H,\mu}$ will be equivariant with respect to this modified action. Note that the stabilizer of the origin of $(\mathfrak{n}/\mathfrak{h})^*$ under this modified action coincides with

$$(\Gamma_{\Sigma_H})_\mu = \left\{[g] \in \Gamma_{\Sigma_H} : \mathrm{Ad}^*(g)\mu = \mu\right\} \equiv \mathrm{N}_\mu^{\Sigma_H}/H, \qquad \mathrm{N}_\mu^{\Sigma_H} := G_\mu \cap \mathrm{N}^{\Sigma_H}.$$

Then, the assumptions entering the Regular Reduction Theorem 10.3.1 are fulfilled. In particular, $(J^{H,\mu})^{-1}(0)$ is an embedded submanifold of Σ_H and the reduced phase space

$$\left(J^{H,\mu}\right)^{-1}(0)/(\Gamma_{\Sigma_H})_\mu \equiv (\Sigma_H \cap M_\mu)/\mathrm{N}_\mu^{\Sigma_H} \tag{10.5.13}$$

carries a unique symplectic structure. In the remainder of this section, we will relate this symplectic manifold to the symplectic manifolds $\hat{M}_\tau$ constructed before.

Let $\Sigma_{H,\mu}$ be a connected component of $\Sigma_H \cap M_\mu$, let $\hat{\Sigma}_{H,\mu}$ denote the corresponding connected component of the reduced phase space $(\Sigma_H \cap M_\mu)/\mathrm{N}_\mu^{\Sigma_H}$ and let $\pi_{H,\mu} : \Sigma_{H,\mu} \to \hat{\Sigma}_{H,\mu}$ denote the natural projection. Since Σ_H is a connected component of M_H, $\Sigma_{H,\mu}$ is also a connected component of $M_H \cap M_\mu$. Therefore, $\Sigma_{H,\mu} \subset M_{\tau,\mu}$ for some orbit-momentum type stratum $\tau \in \mathbb{T}$. Since, by construction, two points of $\Sigma_{H,\mu}$ are conjugate under G iff they are conjugate under

$\mathrm{N}_\mu^{\Sigma_H}$, the natural inclusion mapping $\iota : \Sigma_{H,\mu} \to M_{\tau,\mu}$ induces an injective mapping $\varphi : \hat{\Sigma}_{H,\mu} \to \hat{M}_\tau$ by the following commutative diagram:

$$\begin{array}{ccc} \Sigma_{H,\mu} & \xrightarrow{\iota} & M_{\tau,\mu} \\ {\scriptstyle \pi_{H,\mu}}\downarrow & & \downarrow{\scriptstyle \pi_{\tau,\mu}} \\ \hat{\Sigma}_{H,\mu} & \xrightarrow{\varphi} & \hat{M}_\tau . \end{array} \tag{10.5.14}$$

Proposition 10.5.9 *The mapping φ defined by (10.5.14) is a symplectomorphism.*

For the proof we need

Lemma 10.5.10 *Let $m \in M_H \cap M_\mu$ and let $(\tilde{E}, \tilde{\omega}, \tilde{\Psi}, \tilde{J})$ be a symplectic tubular neighbourhood of the orbit $G \cdot m$ at m with tube symplectomorphism $\chi : U \to \tilde{E}$. Let $\hat{E}$ be the open G_μ-invariant neighbourhood of $G_\mu \cdot [(\mathbb{1}, 0, 0)]$ provided by Lemma 10.5.1 and denote $\hat{U} := \chi^{-1}(\hat{E})$. Assume that $\tilde{V}^H$ is connected.*

1. *If $m \in \Sigma_{H,\mu}$, then $\Sigma_{H,\mu} \cap \hat{U} = \chi^{-1}(\mathrm{N}_\mu^{\Sigma_H} \times_H (\{0\} \times \tilde{V}^H))$.*
2. *If $\Sigma_{H,\mu} \cap \hat{U} \neq \varnothing$, then $(G_\mu \cdot \Sigma_{H,\mu}) \cap \hat{U} = M_{\tau,\mu} \cap \hat{U}$.*

Proof of Lemma 10.5.10 1. Consider the set $W := \chi^{-1}(\{[(\mathbb{1}, 0, v)] : v \in \tilde{V}^H\})$. Since W is a connected subset of $M_H \cap M_\mu$ which intersects $\Sigma_{H,\mu}$ and since the latter is a connected component of $M_H \cap M_\mu$, we have $W \subset \Sigma_{H,\mu} \cap \hat{U}$. Since $\Sigma_{H,\mu} \cap \hat{U}$ is invariant under $\mathrm{N}_\mu^{\Sigma_H}$, then

$$\mathrm{N}_\mu^{\Sigma_H} \cdot W = \chi^{-1}\big(\mathrm{N}_\mu^{\Sigma_H} \times_H \big(\{0\} \times \tilde{V}^H\big)\big) \subset \Sigma_{H,\mu} \cap \hat{U}. \tag{10.5.15}$$

By (10.5.6), and since two points of $\Sigma_{H,\mu}$ are conjugate under G iff they are conjugate under $\mathrm{N}_\mu^{\Sigma_H}$, we have in fact equality in (10.5.15).

2. By point 1, if $\Sigma_{H,\mu} \cap \hat{U} \neq \varnothing$, then it contains a point $\chi^{-1}([(g, 0, v)])$ with $g \in \mathrm{N}_\mu^{\Sigma_H}$ and $v \in \tilde{V}^H$. Let W denote the subset defined under point 1. By the same argument, applied to $\Psi_g(W)$, we find that $\Psi_g(W) \subset \Sigma_{H,\mu} \cap \hat{U}$. Then, since $\mathrm{N}_\mu^{\Sigma_H} \subset G_\mu$,

$$G_\mu \cdot W \subset (G_\mu \cdot \Sigma_{H,\mu}) \cap \hat{U}.$$

By (10.5.7), $G_\mu \cdot W$ coincides with $M_{\tau,\mu} \cap \hat{U}$. Since $(G_\mu \cdot \Sigma_{H,\mu}) \cap \hat{U}$ is contained in $M_{\tau,\mu} \cap \hat{U}$, the assertion follows. □

Proof of Proposition 10.5.9 For every $m \in \Sigma_{H,\mu}$, we find a symplectic tubular neighbourhood $(\tilde{E}, \tilde{\omega}, \tilde{\Psi}, \tilde{J})$ of the orbit $G \cdot m$ at m with tube symplectomorphism $\chi : U \to \tilde{E}$. Using point 1 of Lemma 10.5.10 and the argument of the proof of point 3 of the Singular Reduction Theorem 10.5.4, one can show that χ induces a local symplectomorphism from an open neighbourhood of $\pi_{H,\mu}(m)$ onto $\tilde{V}^H$. Combining this with the corresponding result for $\hat{M}_\tau$, stated in Remark 10.5.5/1, we find that φ is given locally by the identical mapping of $\tilde{V}^H$. Hence, it is a symplectomorphism from $\hat{\Sigma}_{H,\mu}$ onto an open subset of $\hat{M}_\tau$. It remains to show that φ is surjective. For that purpose, for every $m \in M_{\tau,\mu}$, we choose a tube diffeomorphism $\chi_m : U_m \to \tilde{E}_m$ and an open subset $\hat{E}_m$ according to Lemma 10.5.1 and denote $\hat{U}_m = \chi_m^{-1}(\hat{E}_m)$. Let B denote the subset of $M_{\tau,\mu}$ of points which are not G_μ-conjugate to a point in $M_H \cap M_\mu$. We can cover $\hat{M}_\tau$ by the (possibly empty) open subsets $\pi_{\tau,\mu}(\hat{U}_m \cap M_{\tau,\mu})$ with $m \in (M_H \cap M_{\tau,\mu}) \cup B$. Since the fibres of $\pi_{\tau,\mu}$ are the G_μ-orbits in $M_{\tau,\mu}$, we have

$$\pi_{\tau,\mu}(\hat{U}_{m_1} \cap M_{\tau,\mu}) \cap \pi_{\tau,\mu}(\hat{U}_{m_2} \cap M_{\tau,\mu}) = \varnothing$$

whenever $m_1 \in M_H \cap M_{\tau,\mu}$ and $m_2 \in B$. In particular, since $\Sigma_{H,\mu} \subset M_H \cap M_{\tau,\mu}$, we have $\varphi(\hat{\Sigma}_{H,\mu}) \cap \pi_{\tau,\mu}(\hat{U}_{m_2} \cap M_{\tau,\mu}) = \varnothing$ for all $m_2 \in B$. Combining this with Lemma 10.5.10/2, we obtain that $\hat{M}_\tau \setminus \varphi(\hat{\Sigma}_{H,\mu})$ coincides with the union of the open subsets $\pi_{\tau,\mu}(\hat{U}_m \cap M_{\tau,\mu})$ over all points m in $((M_H \cap M_{\tau,\mu}) \setminus \Sigma_{H,\mu}) \cup B$. This shows that the image of φ is also closed in $\hat{M}_\tau$. Since $\hat{M}_\tau$ is connected, it follows that φ is surjective. □

Remark 10.5.11

1. Proposition 10.5.9 implies that for every connected component $\Sigma_{H,\mu}$ of $M_H \cap M_\mu$ there exists a unique orbit-momentum type stratum $\tau \in \mathrm{T}$, with underlying coadjoint orbit $\mathscr{O} = \mathscr{O}_\mu$, such that
$$M_{\tau,\mu} = G_\mu \cdot \Sigma_{H,\mu}. \tag{10.5.16}$$
This observation clarifies the relation between our presentation of singular point reduction and that in [232]. There, the reduction procedure is applied to the subsets
$$(G_\mu \cdot \Sigma_H) \cap M_\mu \equiv G_\mu \cdot (\Sigma_H \cap M_\mu),$$
where, as before, Σ_H is a connected component of M_H. In view of (10.5.16), Proposition 10.5.9 implies that the reduced phase space obtained in [232] is symplectomorphic to a union of certain orbit-momentum type strata $\hat{M}_\tau$ contained in $\hat{M}_{[H]} \cap \hat{M}_{\mathscr{O}}$. It is symplectomorphic to $\hat{M}_{[H]} \cap \hat{M}_{\mathscr{O}}$ under the condition that every subgroup of G_μ which is conjugate to H in G is also conjugate to H in G_μ, cf. Remark 10.5.5/2.
2. As in the regular case, see Proposition 10.3.7, one can perform singular orbit reduction, see [33] for first steps in this direction and [232], Sect. 8.4, for an exhaustive discussion. Ortega and Ratiu show that, provided one endows $J^{-1}(\mathscr{O}_\mu)$ with the appropriate topology,[18] singular orbit reduction yields a stratified symplectic space and that the latter is homeomorphic to the one obtained by point reduction.

[18]Here, a subtlety occurs: whereas $J^{-1}(\mu)$ can be simply viewed as a topological subspace of M, $J^{-1}(\mathscr{O}_\mu)$ has to be endowed with the initial topology induced by the mapping J. If μ is regular and G acts freely on M, this is consistent with endowing $J^{-1}(\mathscr{O}_\mu)$ with the initial submanifold structure provided by the Transversal Mapping Theorem 1.8.2.

3. We comment on the structure of the reduced phase space as a whole. For details we refer to [232] and [238], see also [275].

   For a given coadjoint orbit $\mathscr{O}$, the subset $\hat{M}_{\mathscr{O}}$ of $\hat{M}$ can be identified with the topological quotient M_μ/G_μ for every $\mu \in \mathscr{O}$. It decomposes as

$$\hat{M}_{\mathscr{O}} = \bigcup_\tau \hat{M}_\tau,$$

   where the union is over all orbit-momentum type strata whose underlying coadjoint orbit coincides with $\mathscr{O}$. This decomposition has certain topological and geometric properties which are summarized in the notion of Whitney stratification. In particular, $\hat{M}_{\mathscr{O}}$ has the local structure of a cone space. Accordingly, the symplectic manifolds $\hat{M}_\tau$ are usually referred to as the symplectic strata of $\hat{M}_{\mathscr{O}}$.

   Due to the G-invariance of ω, the subset $C^\infty(M)^G$ is a Poisson subalgebra of $C^\infty(M)$. As a consequence of Noether's Theorem, the subset

$$I = \left\{ f \in C^\infty(M)^G : f_{\restriction M_\mu} = 0 \right\}$$

   is a Poisson ideal of $C^\infty(M)^G$. Thus, we can form the quotient to obtain a Poisson algebra of continuous functions on $\hat{M}_{\mathscr{O}} = M_\mu/G_\mu$:

$$C^\infty(\hat{M}_{\mathscr{O}}) := C^\infty(M)^G/I.$$

   This way, $\hat{M}_{\mathscr{O}}$ is equipped with the structure of a Poisson space.[19] By construction, the inclusion mappings $\hat{M}_\tau \to \hat{M}_{\mathscr{O}}$ are Poisson space morphisms, that is, the natural mappings

$$C^\infty(\hat{M}_{\mathscr{O}}) \to C^\infty(\hat{M}_\tau)$$

   induced by pull-back are Poisson algebra homomorphisms. The whole structure is summarized in the statement that $\hat{M}_{\mathscr{O}}$, together with the Poisson algebra $C^\infty(\hat{M}_{\mathscr{O}})$ and the family of strata $\{\hat{M}_\tau\}$, constitutes a stratified symplectic space.

10.6 Examples from Classical Mechanics

In this section, we perform symplectic reduction for the following examples: the geodesic flow on S^3, the Kepler problem, the Euler top and the spherical pendulum. Apart from the geodesic flow, these examples will be taken up again in Sect. 10.8 and in Chap. 11. At this point, we would like to draw the attention of the reader to the book of Cushman and Bates [69], where the main focus is on examples of the above type, with the theoretical foundations being given in appendices.

Recall that the Euclidean scalar product of two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is denoted by $\mathbf{x} \cdot \mathbf{y}$ and that the corresponding norm is denoted by $\|\mathbf{x}\|$. In this notation, the canonical 1-form θ on $\mathrm{T}\mathbb{R}^n \cong \mathrm{T}^*\mathbb{R}^n$ takes the form

$$\theta(\mathbf{x}, \mathbf{y}) = \mathbf{y} \cdot \mathrm{d}\mathbf{x}. \tag{10.6.1}$$

[19] A Poisson space is a topological space together with a Poisson algebra of continuous functions.

Example 10.6.1 (Geodesic flow on S^3) Let $M = \mathrm{T}\mathbb{R}^4 \cong \mathbb{R}^4 \times \mathbb{R}^4$. The natural action ψ of SO(4) on $\mathbb{R}^4$ and the lift Ψ of this action to $\mathrm{T}\mathbb{R}^4$ are given by

$$\psi : \mathrm{SO}(4) \times \mathbb{R}^4 \to \mathbb{R}^4, \qquad \psi(a, \mathbf{x}) = a\mathbf{x}.$$
$$\Psi : \mathrm{SO}(4) \times \mathrm{T}\mathbb{R}^4 \to \mathrm{T}\mathbb{R}^4, \qquad \Psi\big(a, (\mathbf{x}, \mathbf{y})\big) = (a\mathbf{x}, a\mathbf{y}).$$

By Proposition 8.3.6, the canonical 1-form θ is invariant under Ψ. The Killing vector field generated by $A \in \mathfrak{so}(4)$ under this action is given by

$$A_*(\mathbf{x}, \mathbf{y}) = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} \Psi_{\exp(tA)}(\mathbf{x}, \mathbf{y}) = (A\mathbf{x}, A\mathbf{y}). \tag{10.6.2}$$

Hence, for the equivariant momentum mapping (10.1.23), we obtain

$$J : \mathrm{T}\mathbb{R}^4 \to \mathfrak{so}(4), \quad \big\langle J(\mathbf{x}, \mathbf{y}), A\big\rangle = (A\mathbf{x}) \cdot \mathbf{y}, \tag{10.6.3}$$

that is, $J_A(\mathbf{x}, \mathbf{y}) = (A\mathbf{x}) \cdot \mathbf{y}$. Thus, the tuple $(\mathrm{T}\mathbb{R}^4, \mathrm{d}\theta, \Psi, J)$ defines a Hamiltonian SO(4)-manifold. Let us consider the invariant Hamiltonian

$$H(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\|\mathbf{x}\|^2\|\mathbf{y}\|^2 - \frac{1}{2}(\mathbf{x} \cdot \mathbf{y})^2. \tag{10.6.4}$$

A brief computation (Exercise 10.6.1) yields the Hamilton equations

$$\dot{\mathbf{x}} = \|\mathbf{x}\|^2\mathbf{y} - (\mathbf{x} \cdot \mathbf{y})\mathbf{x}, \tag{10.6.5}$$
$$\dot{\mathbf{y}} = -\|\mathbf{y}\|^2\mathbf{x} + (\mathbf{x} \cdot \mathbf{y})\mathbf{y}, \tag{10.6.6}$$

from which we read off that the projection to the configuration space $\mathbb{R}^4$ of the integral curve with initial condition $\mathbf{x}_0$, $\mathbf{y}_0$ lies on the sphere of radius $\|\mathbf{x}_0\|$. Thus, the dynamics can be restricted to the submanifold

$$\mathrm{TS}^3 = \big\{(\mathbf{x}, \mathbf{y}) \in \mathrm{T}\mathbb{R}^4 : \|\mathbf{x}\|^2 = 1, \ \mathbf{x} \cdot \mathbf{y} = 0\big\}, \tag{10.6.7}$$

of $\mathrm{T}\mathbb{R}^4$, cf. Remark 2.1.4/2. One can check that via the natural inclusion mapping, θ pulls back to the canonical 1-form $\tilde{\theta}$ on TS^3, which is inherited from $\mathrm{T}^*\mathrm{S}^3$ via the isomorphism defined by the Euclidean metric. Thus, by restriction, the Hamiltonian system $(\mathrm{T}\mathbb{R}^4, \mathrm{d}\theta, H)$ induces a Hamiltonian system $(\mathrm{TS}^3, \mathrm{d}\tilde{\theta}, \tilde{H})$, where $\tilde{H} = H_{\restriction \mathrm{TS}^3}$. From (10.6.4)–(10.6.6) we read off

$$\tilde{H}(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\|\mathbf{y}\|^2$$

and the Hamilton equations

$$\dot{\mathbf{x}} = \mathbf{y}, \qquad \dot{\mathbf{y}} = -\|\mathbf{y}\|^2\mathbf{x}. \tag{10.6.8}$$

Since $\|\mathbf{y}\|^2 = 2\tilde{H}$ is a constant of motion, this system of equations can be trivially integrated:

$$\begin{bmatrix} \mathbf{x}(t) \\ \mathbf{y}(t) \end{bmatrix} = \begin{bmatrix} \cos(t\sqrt{2h}) & \frac{1}{\sqrt{2h}}\sin(t\sqrt{2h}) \\ -\sqrt{2h}\sin(t\sqrt{2h}) & \cos(t\sqrt{2h}) \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{bmatrix},$$

where $h = \tilde{H}(\mathbf{x}_0, \mathbf{y}_0)$. Every integral curve with $h > 0$ is periodic and its projection to the configuration space S^3 is a great circle whose position is defined by the initial conditions $\mathbf{x}_0, \mathbf{y}_0$. The Hamiltonian vector field $X_{\tilde{H}}$ is called the geodesic vector field on S^3 and its flow is called the geodesic flow on S^3. We leave it to the reader to check that this is indeed the geodesic flow on S^3 with respect to the Riemannian structure induced from $\mathbb{R}^4$ as discussed in Example 9.2.1.

Since the submanifold TS^3 given by (10.6.7) is invariant under the action Ψ, the latter restricts to an action

$$\tilde{\Psi}: \mathrm{SO}(4) \times \mathrm{TS}^3 \to \mathrm{TS}^3, \qquad \Psi\big(a, (\mathbf{x}, \mathbf{y})\big) = (a\mathbf{x}, a\mathbf{y}),$$

and the equivariant momentum mapping (10.6.3) restricts to an equivariant momentum mapping $\tilde{J}$ for this action. Thus, $(\mathrm{TS}^3, \mathrm{d}\tilde{\theta}, \tilde{\Psi}, \tilde{J})$ is a Hamiltonian SO(4)-manifold. Let us study its structure. For that purpose, from Remark 5.4.11/2 we recall that the Killing form of $\mathfrak{so}(4)$ induces an equivariant vector space isomorphism $\mathfrak{so}(4)^* \cong \mathfrak{so}(4)$. In turn, we can identify the vector space underlying $\mathfrak{so}(4)$ with $\bigwedge^2 \mathbb{R}^4$ via the mapping

$$A \mapsto A^{ij} \mathbf{e}_i \wedge \mathbf{e}_j$$

(summation convention), where $\mathbf{e}_i$ denote the standard basis vectors of $\mathbb{R}^4$. Since

$$\big\langle J(\mathbf{x}, \mathbf{y}), A \big\rangle = A_{ij} x^j y^i = \frac{1}{2}\big(x^j y^i - x^i y^j\big) A_{ij},$$

under these two identifications we find

$$J(\mathbf{x}, \mathbf{y}) = \mathbf{x} \wedge \mathbf{y}. \tag{10.6.9}$$

It follows that J satisfies the Plücker equation

$$J \wedge J = 0. \tag{10.6.10}$$

In coordinates, it reads

$$J_{12} J_{34} + J_{13} J_{42} + J_{14} J_{23} = 0. \tag{10.6.11}$$

In what follows, we use the natural scalar product

$$\langle \mathbf{u} \wedge \mathbf{v}, \mathbf{x} \wedge \mathbf{y} \rangle := (\mathbf{u} \cdot \mathbf{x})(\mathbf{v} \cdot \mathbf{y}) - (\mathbf{u} \cdot \mathbf{y})(\mathbf{v} \cdot \mathbf{x})$$

in $\bigwedge^2 \mathbb{R}^4$. Denoting the corresponding norm by $\| \cdot \|$, we obtain

$$\big\| J(\mathbf{x}, \mathbf{y}) \big\|^2 = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - (\mathbf{x} \cdot \mathbf{y})^2$$

and hence

$$\big\| \tilde{J}(\mathbf{x}, \mathbf{y}) \big\|^2 = \|\mathbf{y}\|^2 = 2\tilde{H}(\mathbf{x}, \mathbf{y}). \tag{10.6.12}$$

Let us determine the level set $\tilde{J}^{-1}(A)$ for $A \in \mathfrak{so}(4)$. In the case $A = 0$, according to (10.6.12), $\tilde{J}^{-1}(0)$ coincides with the zero section s_0 of TS^3. From the Hamilton equations (10.6.8) we read off that in this case the integral curves are points, that is, the dynamics is trivial. In the case $A \neq 0$, we can restrict $\tilde{J}$ to the subbundle

$$\mathrm{T}^+\mathrm{S}^3 = \mathrm{TS}^3 \setminus s_0\big(\mathrm{S}^3\big).$$

The restriction will be denoted by J^+. By the Noether Theorem 10.1.9, the level set $(J^+)^{-1}(A)$ is a union of integral curves of the geodesic vector field $X_{\tilde{H}}$. We show that it consists in fact of a single integral curve of $X_{\tilde{H}}$. This implies, in particular, that the image of J^+ may be interpreted as the space of integral curves. According to (10.6.10),

$$\operatorname{im} J^+ = \{A \in \mathfrak{so}(4) : A \wedge A = 0,\ A \neq 0\} \tag{10.6.13}$$

(Exercise 10.6.2). This is a 5-dimensional submanifold of $\mathfrak{so}(4)$. We show that J^+ is a submersion. Let $(\mathbf{x}, \mathbf{y}) \in \mathrm{T}^+\mathrm{S}^3$. Since

$$\dim \mathrm{T}_{(\mathbf{x},\mathbf{y})}(\mathrm{T}^+\mathrm{S}^3) = \dim\ker(J^+)'_{(\mathbf{x},\mathbf{y})} + \dim\operatorname{im}(J^+)'_{(\mathbf{x},\mathbf{y})},$$

it is enough to show $\dim\ker(J^+)'_{(\mathbf{x},\mathbf{y})} = 1$. Let $X = (\dot{\mathbf{x}}, \dot{\mathbf{y}}) \in \mathrm{T}_{(\mathbf{x},\mathbf{y})}(\mathrm{T}^+\mathrm{S}^3)$ and let $t \mapsto (\mathbf{x}(t), \mathbf{y}(t))$ be a curve through $(\mathbf{x}, \mathbf{y})$ representing X. Then,

$$(J^+)'_{(\mathbf{x},\mathbf{y})}(X) = \frac{\mathrm{d}}{\mathrm{d}t}\restriction_0 \mathbf{x}(t) \wedge \mathbf{y}(t) = \dot{\mathbf{x}} \wedge \mathbf{y} + \mathbf{x} \wedge \dot{\mathbf{y}}.$$

The condition $(J^+)'_{(\mathbf{x},\mathbf{y})}(X) = 0$ is fulfilled for multiples of the geodesic vector field,

$$X = s(\mathbf{y}, -\|\mathbf{y}\|^2\mathbf{x}), \quad s \in \mathbb{R},$$

and one can show that these are the only solutions (Exercise 10.6.3). Thus, J^+ is a submersion, indeed. Now, the Level Set Theorem 1.8.3 yields that for every $A \in \operatorname{im} J^+$, the level set $(J^+)^{-1}(A)$ is a one-dimensional submanifold of TS^3 satisfying

$$\mathrm{T}_{(\mathbf{x},\mathbf{y})}(J^+)^{-1}(A) = \ker(J^+)'_{(\mathbf{x},\mathbf{y})}.$$

Since $\ker(J^+)'_{(\mathbf{x},\mathbf{y})}$ is spanned by $X_{\tilde{H}}(\mathbf{x}, \mathbf{y})$, every connected component of $(J^+)^{-1}(A)$ coincides with an integral curve of $X_{\tilde{H}}$. Below, we will see that the level sets of J^+ are in fact 1-spheres, cf. Remark 10.6.2. This yields the assertion.

To summarize, due to the high degree of symmetry in this example, the preimages of the momentum mapping already yield the integral curves, so that a further symplectic reduction is not necessary.

Remark 10.6.2 We show that $\operatorname{im} J^+$ is foliated by the generic orbits of the coadjoint representation. This way, we will be able to describe the geometry of the energy surfaces. We use the isomorphism

$$\mathfrak{so}(4) \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2)$$

of Example 5.2.15/5 to represent J^+ by a pair (K, I) of $\mathfrak{su}(2)$-valued mappings on TS^3. According to (5.2.11), under the identification $\mathbb{R}^4 \cong \mathbb{H}$ of Example 5.1.11 and $\mathfrak{su}(2) \cong \mathfrak{sp}(1)$ of Example 5.2.8, J^+, K and I are related by

$$J^+\mathbf{q} = K\mathbf{q} - \mathbf{q}I \tag{10.6.14}$$

for all $\mathbf{q} \in \mathbb{H}$. Decomposing $K = \sum_i K_i \mathrm{I}_i^{\mathbb{C}}$ and $I = \sum_i I_i \mathrm{I}_i^{\mathbb{C}}$ in the basis of Example 5.2.8, we read off

$$K_1 = \frac{1}{2}(J^+_{23} + J^+_{14}), \qquad K_2 = \frac{1}{2}(J^+_{31} + J^+_{24}), \qquad K_3 = \frac{1}{2}(J^+_{12} + J^+_{34}),$$
$$I_1 = \frac{1}{2}(J^+_{23} - J^+_{14}), \qquad I_2 = \frac{1}{2}(J^+_{31} - J^+_{24}), \qquad I_3 = \frac{1}{2}(J^+_{12} - J^+_{34}).$$

Using

$$\|K\|^2 = K_1^2 + K_2^2 + K_3^2, \qquad \|I\|^2 = I_1^2 + I_2^2 + I_3^2$$

one finds that in terms of K and I, the Plücker equation (10.6.10) takes the form

$$\|K\|^2 = \|I\|^2 \tag{10.6.15}$$

and that the identity (10.6.12) reads

$$\|K(\mathbf{x}, \mathbf{y})\|^2 = \frac{1}{4}\|\mathbf{y}\|^2 = \|I(\mathbf{x}, \mathbf{y})\|^2 \tag{10.6.16}$$

(Exercise 10.6.4). Thus, the image im J^+ can be characterized as follows:

$$\operatorname{im} J^+ = \big\{(A, B) \in \mathfrak{su}(2) \times \mathfrak{su}(2) : \|A\| = \|B\| \neq 0\big\}.$$

According to (8.4.8), for every $(A, B) \in \operatorname{im} J^+$ the coadjoint orbit $\mathscr{O}_{(A,B)}$ is diffeomorphic to $\mathrm{S}^2_{\|A\|} \times \mathrm{S}^2_{\|A\|}$. From (10.6.16) we read off that the preimages of these orbits under J^+ have the topology of $\mathrm{S}^3 \times \mathrm{S}^2$ and that they coincide with the energy level sets in $\mathrm{T}^+\mathrm{S}^3$. For every fixed energy level

$$h = \frac{1}{2}\|\mathbf{y}\|^2 = 2\|A\|^2 = 2\|B\|^2 \neq 0,$$

J^+ yields a locally trivial fibre bundle

$$J_h^+ : \tilde{H}^{-1}(h) \cong \mathrm{S}^3 \times \mathrm{S}^2_{\sqrt{2}\|A\|} \longrightarrow \mathrm{S}^2_{\|A\|} \times \mathrm{S}^2_{\|A\|}. \tag{10.6.17}$$

The connected components of the fibres of this bundle are the integral curves of the geodesic vector field corresponding to the energy level h and the base manifold $\mathrm{S}^2_{\|A\|} \times \mathrm{S}^2_{\|A\|}$ is the space of integral curves of energy h.

To show that the fibres are in fact 1-spheres, and thus coincide with the integral curves of X_H, we study this bundle using the original parameterization by pairs $(\mathbf{x}, \mathbf{y})$ of orthogonal vectors. For fixed $\mathbf{y}$ and for given $\mathbf{x} \wedge \mathbf{y}$, the vector $\mathbf{x}$ runs exactly once through the great circle obtained by the intersection of S^3 with the 2-dimensional surface defined by the 2-vector $\mathbf{x} \wedge \mathbf{y}$. We conclude that, as a fibre bundle, (10.6.17) is isomorphic to the Stiefel bundle $\mathrm{S}_{\mathbb{R}}(2, 4) \to \tilde{\mathrm{G}}_{\mathbb{R}}(2, 4)$, where

$$\mathrm{S}_{\mathbb{R}}(2, 4) = \mathrm{SO}(4)/\mathrm{SO}(2) \cong \mathrm{S}^3 \times \mathrm{S}^2$$

is the space of orthonormal 2-frames in $\mathbb{R}^4$ (the Stiefel manifold) and

$$\tilde{\mathrm{G}}_{\mathbb{R}}(2, 4) = \mathrm{SO}(4)/\big(\mathrm{SO}(2) \times \mathrm{SO}(2)\big) \cong \mathrm{S}^2 \times \mathrm{S}^2$$

denotes the space of oriented 2-dimensional surfaces in $\mathbb{R}^4$ (the 2-fold covering space of the Graßmann manifold $\mathrm{G}_{\mathbb{R}}(2, 4)$), cf. Examples 5.7.5 and 5.7.6. For the proof, see Exercise 10.6.5.

Example 10.6.3 (Kepler problem) To start with, we summarize some well-known facts from classical mechanics. The phase space of the Kepler problem is

$$M = \mathrm{T}^*\big(\mathbb{R}^3 \setminus \{0\}\big) \cong \big(\mathbb{R}^3 \setminus \{0\}\big) \times \mathbb{R}^3.$$

In the standard bundle coordinates $(\mathbf{q}, \mathbf{p})$, the symplectic form is given by

$$\omega = \mathrm{d}p_i \wedge \mathrm{d}q^i,$$

and the Hamiltonian is

$$H(\mathbf{q}, \mathbf{p}) = \frac{\mathbf{p}^2}{2} - \frac{1}{q}, \qquad q = \|\mathbf{q}\|.$$

Here, the reduced mass and the coupling constant have been put equal to 1. With this choice, we consider the case of an attractive force. The Hamiltonian vector field generated by H is

$$X_H = p_i \partial_{q^i} - \frac{q^i}{q^3} \partial_{p_i}.$$

The Hamiltonian H is invariant under the lift to M of the natural action of SO(3) on $\mathbb{R}^3$. From Example 10.1.23 we know that under the identifications $\mathfrak{so}(3)^* \cong \mathfrak{so}(3)$ induced by the Killing form and $\mathfrak{so}(3) \cong \mathbb{R}^3$ given by (5.2.6), the associated equivariant momentum mapping $J: M \to \mathfrak{so}(3)^*$ coincides with angular momentum,

$$J(\mathbf{q}, \mathbf{p}) = \mathbf{q} \times \mathbf{p} = \mathbf{L}.$$

Thus, we have a left Hamiltonian SO(3)-manifold with invariant Hamiltonian H. Let us discuss symplectic reduction for this theory. Note that the coadjoint orbits of SO(3) are labelled by $l = \|\mathbf{L}_0\|$. We have to distinguish between two cases.

(a) The case $l > 0$. Here, $\mathbf{L}_0$ is a regular value of J and the level set

$$J^{-1}(\mathbf{L}_0) = \big\{(\mathbf{q}, \mathbf{p}) \in M : \mathbf{q} \times \mathbf{p} = \mathbf{L}_0\big\}$$

is an embedded submanifold of M, acted upon properly and freely by the stabilizer $\mathrm{SO}(3)_{\mathbf{L}_0} \cong \mathrm{SO}(2)$. Thus, the Regular Reduction Theorem applies, yielding the reduced phase space

$$\hat{M}_l = J^{-1}(\mathbf{L}_0)/\mathrm{SO}(3)_{\mathbf{L}_0}.$$

To describe $\hat{M}_l$ explicitly, we note that the constant of motion $\mathbf{L}_0$ reduces the dynamics to the plane orthogonal to $\mathbf{L}_0$. Without loss of generality, we can assume $\mathbf{L}_0 = (0, 0, l)$. Then, this plane coincides with the q^1–q^2-plane and $J^{-1}(\mathbf{L}_0)$ is contained in the phase space

$$\tilde{M} = \mathrm{T}^*\big(\mathbb{R}^2 \setminus \{0\}\big) \cong \big(\mathbb{R}_+ \times \mathrm{SO}(2)\big) \times \mathbb{R}^2.$$

Since $\mathbf{q} \neq 0$, we can use polar coordinates (q, ϕ) in the q^1–q^2-plane. We denote

$$p_q = p_1 \cos\phi + p_2 \sin\phi, \qquad p_\phi = -q p_1 \sin\phi + q p_2 \cos\phi$$

and observe that $(q^1, q^2, p_1, p_2) \mapsto (q, \varphi, p_q, p_\varphi)$ is a canonical transformation of $\tilde{M}$. In terms of the new coordinates, we have

$$J^{-1}(\mathbf{L}_0) = \big\{(q, \phi, p_q, p_\phi) \in \tilde{M} : p_\phi = l\big\} \cong \big(\mathbb{R}_+ \times \mathrm{SO}(2)\big) \times \mathbb{R}$$

and

$$H_{\restriction J^{-1}(\mathbf{L}_0)} = \frac{1}{2}\left(p_q^2 + \frac{l^2}{q^2}\right) - \frac{1}{q}.$$

Moreover, the action of $\mathrm{SO}(3)_{\mathbf{L}_0}$ is given by $\phi \mapsto \phi + \alpha$ and the functions q, p_q, p_φ and $H_{\restriction J^{-1}(\mathbf{L}_0)}$ are invariant under this action. Thus, the reduced phase space $\hat{M}_l$ is isomorphic to $\mathbb{R}_+ \times \mathbb{R}$ endowed with the symplectic form $\mathrm{d}p_q \wedge \mathrm{d}q$ and the reduced Hamiltonian is given by

$$H_l = \frac{1}{2}\left(p_q^2 + \frac{l^2}{q^2}\right) - \frac{1}{q}. \tag{10.6.18}$$

As a result, we arrive at a one-dimensional problem with effective potential

$$U_l(r) = \frac{l^2}{2q^2} - \frac{1}{q}.$$

From classical mechanics we know that the orbits in configuration space are ellipses for $h < 0$, hyperbolas for $h > 0$ and parabolas for $h = 0$. Below, we will use the hidden symmetry provided by the Lenz-Runge vector to prove these facts by purely algebraic arguments.

(b) The case $l = 0$, that is, $\mathbf{L}_0 = 0$. This is not a regular value of J. The stabilizer is $\mathrm{SO}(3)_0 = \mathrm{SO}(3)$. We find

$$J^{-1}(0) = \big\{(\mathbf{q}, \mathbf{p}) \in M : \mathbf{p} = \alpha \mathbf{q},\ \alpha \in \mathbb{R}\big\} \cong \big(\mathbb{R}^3 \setminus \{0\}\big) \times \mathbb{R},$$

that is, $J^{-1}(0)$ is a trivial line bundle over the configuration space $\mathbb{R}^3 \setminus \{0\}$. Let $m_0 = (\mathbf{q}_0, \alpha\mathbf{q}_0) \in J^{-1}(0)$ be a chosen point. Its stabilizer $\mathrm{SO}(3)_{m_0} \cong \mathrm{SO}(2)$ consists of the rotations about the axis defined by $\mathbf{q}_0$. The connected component Σ of $M_{\mathrm{SO}(3)_{m_0}}$ containing m_0 is

$$\Sigma = \big\{(\lambda\mathbf{q}_0, \lambda\alpha\mathbf{q}_0) \in M : \alpha, \lambda \in \mathbb{R},\ \lambda > 0\big\} \cong \mathbb{R}_+ \times \mathbb{R}. \tag{10.6.19}$$

We observe that $\mathrm{SO}(3) \cdot \Sigma = J^{-1}(0)$. Hence, we obtain $\Sigma \cap J^{-1}(0) = \Sigma$, that is, the intersection $\Sigma \cap J^{-1}(0)$ consists of a single connected component. The subset $\hat{M}_{[\mathrm{SO}(3)_{m_0}]} \cap \hat{M}_{\mathscr{O}_0}$ of the orbit space $\hat{M} = M/\mathrm{SO}(3)$ consists of a single orbit-momentum type stratum $\hat{M}_0$. The corresponding submanifold $M_{\tau,\mu}$ is given by $\mathrm{SO}(3) \cdot \Sigma = J^{-1}(0)$. Thus,

$$\hat{M}_0 = J^{-1}(0)/\mathrm{SO}(3),$$

which may be identified with Σ. Under this identification, the reduction of the canonical 1-form yields

$$p_i\, \mathrm{d}q^i = \alpha q_i\, \mathrm{d}q^i = \frac{\alpha}{2}\, \mathrm{d}q^2 = p_q\, \mathrm{d}q, \qquad p_q = \alpha q,$$

and the reduced Hamiltonian reads

$$H_0 = \frac{p_q^2}{2} - \frac{1}{q}. \tag{10.6.20}$$

Obviously, the orbits are located on rays in $\mathbb{R}^3$. For $h < 0$ they are bounded from above by the value $q = -\frac{1}{h}$.

Now, let us recall that there are three additional constants of motion in the Kepler problem, given by the Lenz-Runge vector

$$\mathbf{R} = \mathbf{p} \times \mathbf{L} - \frac{\mathbf{q}}{q}, \tag{10.6.21}$$

which fulfils

$$\mathbf{L} \cdot \mathbf{R} = 0. \tag{10.6.22}$$

This fact enhances the symmetry of the problem. We provide the discussion for the case $h < 0$. Thus, in what follows, let $\Sigma_- \subset M$ be the open subset of points with negative energy. We define

$$\mathbf{K} := -\frac{1}{2}\left(\mathbf{L} - \frac{1}{\sqrt{-2H}}\mathbf{R}\right), \qquad \mathbf{I} := -\frac{1}{2}\left(\mathbf{L} + \frac{1}{\sqrt{-2H}}\mathbf{R}\right). \tag{10.6.23}$$

For the Poisson brackets, we find

$$\{K_i, K_j\} = \varepsilon_{ij}{}^k K_k, \qquad \{I_i, I_j\} = \varepsilon_{ij}{}^k I_k, \qquad \{K_i, I_j\} = 0. \tag{10.6.24}$$

We read off that $\{K_i\}$ and $\{I_i\}$ constitute bases in the Lie algebra $\mathfrak{su}(2)$. Therefore, via the isomorphism $\mathfrak{so}(4) \cong \mathfrak{su}(2) \times \mathfrak{su}(2)$ used before, together they constitute a basis

$$e_i := K_i, \qquad e_{3+i} := I_i, \quad i = 1, 2, 3,$$

of the Lie algebra $\mathfrak{so}(4)$.[20] It is easy to see that the linear mapping

$$\psi : \mathfrak{so}(4) \to \mathfrak{X}(M), \qquad \psi(e_i) := X_{e_i} \tag{10.6.25}$$

defines a right symplectic action of $\mathfrak{so}(4)$ on M. Since, by construction, every vector field $\psi(A)$ with $A \in \mathfrak{so}(4)$ is Hamiltonian, Proposition 10.1.3 implies that ψ is Hamiltonian with momentum mapping

$$J : M \to \mathfrak{so}(4)^*, \qquad J(m) := e_i(m)e^{*i},$$

where e^{*i} denotes the dual basis in $\mathfrak{so}(4)^*$. Thus, (M, ω, ψ, J) is a right Hamiltonian $\mathfrak{so}(4)$-manifold. If we identify $\mathfrak{so}(4)^* \cong \mathfrak{so}(4)$ via $e_i \mapsto e^{*i}$, we obtain

$$J_{e_i} = e_i, \qquad X_{J_{e_i}} = X_{e_i} = e_{i*} \tag{10.6.26}$$

and, consequently,

$$\{J_{e_i}, J_{e_j}\} = \{e_i, e_j\} = c^k_{ij} e_k = c^k_{ij} J_{e_k} = J_{c^k_{ij} e_k} = J_{\{e_i, e_j\}},$$

where c^k_{ij} are the structure constants of $\mathfrak{so}(4)$ with respect to the basis $\{e_i\}$. Thus, the system is strongly Hamiltonian. By (10.6.22) and (10.6.23),

$$\|K\|^2 = \|I\|^2. \tag{10.6.27}$$

[20] More precisely, we have faithful representations of these Lie algebras in the Poisson algebra of smooth functions on phase space.

An explicit calculation (Exercise 10.6.6) yields

$$\|K\|^2 = \frac{1}{-2H} = \|I\|^2. \tag{10.6.28}$$

Thus, as for the geodesic flow on S^3, the image of J is given by the 5-dimensional submanifold

$$\operatorname{im} J = \big\{(A, B) \in \mathfrak{su}(2) \times \mathfrak{su}(2) : \|A\| = \|B\| \neq 0\big\}$$

of $\mathfrak{so}(4)$, the coadjoint orbits fulfil $\mathscr{O}_{(A,B)} \cong S^2_{\|A\|} \times S^2_{\|A\|}$ and their preimages under J coincide with the energy level sets defined by (10.6.28), cf. Remark 10.6.2. As in that example, one can show that $J : \Sigma_- \to \operatorname{im} J$ is a submersion and that the level sets coincide with the integral curves of the Hamiltonian vector field. This means that the values of the constants of motion $\mathbf{L}$ and $\mathbf{R}$ uniquely determine the integral curves. As before, one has to distinguish between the following two cases.

(a) The case $l \neq 0$. Here, the integral curve in Σ_- projects to an oriented ellipse in the configuration space $\mathbb{R}^3 \setminus \{0\}$. To see this, recall that the motion takes place in the plane orthogonal to the angular momentum vector $\mathbf{L}$. Due to (10.6.22), the Lenz-Runge vector $\mathbf{R}$ defines an axis in this plane. Using (10.6.21) and denoting $\varepsilon := \|\mathbf{R}\|$, one finds

$$\varepsilon^2 = 2hl^2 + 1, \qquad \mathbf{R} \cdot \mathbf{q} = l^2 - q. \tag{10.6.29}$$

Let ϕ be the angle between $\mathbf{R}$ and $\mathbf{q}$. Then, (10.6.29) implies

$$\mathbf{R} \cdot \mathbf{q} = \varepsilon q \cos\phi = l^2 - q.$$

It follows that

$$q = \frac{l^2}{\varepsilon \cos\phi + 1} \tag{10.6.30}$$

with ε being given by (10.6.29). This is the polar equation of an ellipse in the plane orthogonal to $\mathbf{L}$. For the semimajor and semiminor axes we obtain

$$a = \frac{l^2}{1-\varepsilon^2} = \frac{1}{-2h}, \qquad b = \frac{l^2}{\sqrt{1-\varepsilon^2}} = \frac{l}{\sqrt{-2h}}, \tag{10.6.31}$$

respectively. The direction of the semimajor axis is determined by $\mathbf{R}$.

(b) The case $l = 0$. Here, the orbits in the configuration space $\mathbb{R}^3 \setminus \{0\}$ are half-open intervals

$$\left\{\left(s\mathbf{e}, \pm\sqrt{2\left(h + \frac{1}{s}\right)}\mathbf{e}\right) : s \in \left(0, -\frac{1}{h}\right]\right\},$$

where $\mathbf{e}$ is an arbitrary vector of unit length in $\mathbb{R}^3$. Note that for the initial conditions $(-\frac{1}{h}\mathbf{e}, 0)$, the integral curve falls into the origin after the finite time

$$T = \frac{\pi}{2}(-2h)^{-\frac{3}{2}}.$$

This makes explicit that the Hamiltonian vector field of the Kepler problem is not complete.

Remark 10.6.4

1. For $h > 0$, one can solve the problem by deriving commutation relations analogous to (10.6.24). This leads to the Lie algebra $\mathfrak{so}(3,1)$. For $h = 0$, the expression $\frac{1}{\sqrt{-2h}}$ does not make sense. In this case, one has to use L_i and R_i as generators, which leads to the Lie algebra of the group of Euclidean motions $\mathrm{SO}(3) \ltimes \mathbb{R}^3$ (Exercise 10.6.7).

   For the history of the Lenz-Runge vector and the resulting symmetries, see [108, 109].
2. In 1970, Moser found a deep relation between the Kepler problem and the geodesic flow on S^3. The latter provides a certain regularization of the Kepler problem [222].

   In the above notation, we consider the following extension of the stereographic projection to the tangent bundle of S^3:
$$\Phi : \mathrm{T}^+\mathrm{S}^3_{\mathrm{p}} \to \mathrm{T}\mathbb{R}^3_0, \qquad (\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{u}, \mathbf{v}), \tag{10.6.32}$$
   defined by
$$u_k := \frac{x_k}{1 - x_4}, \qquad v_k := (1 - x_4)y_k + y_4 x_k, \quad k = 1, 2, 3.$$
   Here, $\mathrm{S}^3_{\mathrm{p}} := \mathrm{S}^3 \setminus (0, 0, 0, 1)$ denotes the 3-sphere with the north pole removed (the punctured sphere). One can check that Φ is a symplectomorphism and that the pull-back $F = (\Phi^{-1})^* \tilde{H}$ of the Hamiltonian by Φ^{-1} is given by
$$F(\mathbf{u}, \mathbf{v}) = \frac{1}{8} \|\mathbf{v}\|^2 \big(\|\mathbf{u}\|^2 + 1\big)^2$$
   (Exercise 10.6.8). The corresponding Hamilton equations read
$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}s} = \nabla_{\mathbf{v}} F, \qquad \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}s} = -\nabla_{\mathbf{u}} F.$$
   In what follows, we restrict ourselves to the level set defined by $\tilde{H} = \frac{1}{2}$. Then, every transformation $F \to G(F)$ of the Hamiltonian with the property $G'(\frac{1}{2}) = 1$ yields a flow whose restriction to the level set coincides with the flow generated by F. We choose
$$G(\mathbf{u}, \mathbf{v}) := \sqrt{2F(\mathbf{u}, \mathbf{v})} - 1 = \frac{1}{2} \|\mathbf{v}\| \big(\|\mathbf{u}\|^2 + 1\big) - 1.$$
   For the Hamiltonian G, the Hamilton equations, restricted to the level set, have the same form as the equations defined by F and the level set is now given by $G = 0$. Let us define a new flow parameter, $s \mapsto t(s)$, by
$$\frac{\mathrm{d}t}{\mathrm{d}s} = \|\mathbf{v}\|.$$
   This yields a scaling of the Hamiltonian vector field X_G by $\|\mathbf{v}\|^{-1}$ and induces the following transformation of the Hamilton equations:
$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \|\mathbf{v}\|^{-1} \nabla_{\mathbf{v}} G, \qquad \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\|\mathbf{v}\|^{-1} \nabla_{\mathbf{u}} G.$$

We set

$$K(\mathbf{u},\mathbf{v}) := \|\mathbf{v}\|^{-1} G(\mathbf{u},\mathbf{v}) - \frac{1}{2} = \frac{1}{2}\|\mathbf{u}\|^2 - \|\mathbf{v}\|^{-1}.$$

On the level set, we have $K = -\frac{1}{2}$ and the Hamilton equations for K have the canonical form. If we now perform the canonical transformation

$$(\mathbf{u},\mathbf{v}) \mapsto \big(\mathbf{q}(\mathbf{u},\mathbf{v}), \mathbf{p}(\mathbf{u},\mathbf{v})\big) := (-\mathbf{v}, \mathbf{u}),$$

we obtain the Hamiltonian of the Kepler problem:

$$K(\mathbf{q},\mathbf{p}) = \frac{1}{2}\mathbf{p}^2 - \frac{1}{q}.$$

Finally, let us study the behaviour of the flows in the neighbourhood of the north pole $(0,0,0,1)$: For $x_4 \to 1$, we have $\|\mathbf{u}\| \to \infty$, $\|\mathbf{v}\| \to 0$ and hence

$$\|\mathbf{p}\| \to \infty, \quad q \to 0.$$

In this limit, we obtain the singular integral curves of the Kepler problem. More precisely, for $K = -\frac{1}{2}$ and $\mathbf{L} = 0$ we obtain the integral curves

$$\left\{ \left(s\mathbf{e}, \pm\sqrt{-1 + \frac{2}{s}}\,\mathbf{e} \right) : s \in (0, 2] \right\}.$$

Under the above diffeomorphism, running through the orbit of the Kepler problem from $(\mathbf{e}, 0)$ to the singularity $(0, \infty)$ corresponds to running through a great circle from the south pole to the north pole. Moving on the same great circle back to the south pole corresponds to moving through the Kepler orbit back to the starting point. Instead of falling into the singularity, the moving particle is getting reflected. In this sense, dynamics has been regularized. Mathematically speaking, the space of momenta has been compactified.

Let us summarize. Every energy level set corresponding to a negative value of h is diffeomorphically mapped by the above discussed transformation onto the unit sphere bundle over the 3-sphere with the north pole removed (the pointed 3-sphere). After an appropriate reparameterization of the time variable, this diffeomorphism maps the flow of the Kepler problem to the geodesic flow of the pointed 3-sphere. This way, the singular Kepler orbits are mapped to regular orbits of the geodesic flow through the north pole.

The above ideas have attracted much attention. There is a whole bunch of related regularization procedures, see e.g. [67, 175, 176, 185] and [300]. The article of Vivarelli [300] contains an overview with a large number of references. We also refer to Cushman and Bates [69], where some regularization methods are contained in the form of exercises to Chap. II, and to Guillemin and Sternberg [118], where in particular the relation to the conformal group is discussed in detail.

Example 10.6.5 (The Euler top) We use the results of Examples 8.4.5/3, 9.2.1 and 9.2.2. We consider a top, that is, a rigid body, which is fixed at one distinguished point. For the description of the dynamics of rigid bodies one uses two reference frames: a fixed inertial frame in 3-dimensional Euclidean space, given by an orthonormal frame $\{\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3\}$, and a non-inertial, so called body frame, which is fixed to the body, described by an orthonormal frame $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. The corresponding coordinates are called space and body coordinates, respectively. If we put the origins of both reference frames into the distinguished point, then the position of the body is uniquely determined by a rotation matrix a, which transforms the frame $\{\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3\}$ into $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, that is,

$$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = (a\mathbf{n}_1, a\mathbf{n}_2, a\mathbf{n}_3),$$

or

$$\mathbf{e}_i = a^j{}_i \mathbf{n}_j.$$

Thus, the configuration space of a top is the group manifold of the rotation group SO(3) and the phase space is $\mathrm{T}^*\mathrm{SO}(3) \cong \mathrm{SO}(3) \times \mathfrak{so}(3)^*$, where we use the trivialization by left translation, cf. (8.3.6). That is, the top belongs to the class discussed in Example 9.2.2. As in the previous example, we use the identifications $\mathfrak{so}(3)^* \cong \mathfrak{so}(3) \cong \mathbb{R}^3$ and the corresponding identifications of the adjoint and the coadjoint representations of SO(3), cf. (5.2.6) and Remark 5.4.11/2. For an element $\mathbf{x} \in \mathbb{R}^3$ we shall write x when viewed as an element of $\mathfrak{so}(3)$ or of $\mathfrak{so}(3)^*$.

In what follows, we will rather use the body frame. For every curve $t \mapsto a(t)$ in the configuration space and every t, the tangent vector $\dot{a}(t)$ defines an element

$$\omega(t) = a^{-1}\dot{a}(t)$$

of the Lie algebra. The corresponding vector $\boldsymbol{\omega}$ in $\mathbb{R}^3$ is the vector of angular velocity. The kinetic energy of a rigid body is

$$T = \frac{1}{2}\boldsymbol{\omega} \cdot (\Theta\boldsymbol{\omega}),$$

with Θ denoting the inertia tensor. Here, we restrict ourselves to the Euler top, that is, we assume that there are no external forces acting on the body. Then, the generalized momentum coincides with the vector of angular momentum,

$$\mathbf{L} = \frac{\partial T}{\partial \boldsymbol{\omega}} = \Theta\boldsymbol{\omega},$$

and the Hamiltonian is obtained from T by a Legendre transformation:

$$H(a, \mathbf{L}) = \frac{1}{2}\big(\Theta^{-1}\mathbf{L}\big) \cdot \mathbf{L}. \tag{10.6.33}$$

Under the above identifications, it is given by the function

$$H : \mathrm{SO}(3) \times \mathfrak{so}(3) \to \mathbb{R}, \qquad H(a, L) = \frac{1}{2}\Theta^{-1}(L, L), \tag{10.6.34}$$

where Θ^{-1} is the constant, symmetric, positive-definite covariant tensor of rank 2 defined by

$$\Theta^{-1}(L, L) := \big(\Theta^{-1}\mathbf{L}\big) \cdot \mathbf{L}.$$

The tensor Θ constitutes a Riemannian metric on the configuration space SO(3) which is invariant under the lift Ψ of the action of SO(3) on itself by left translation, cf. Example 10.1.24. Thus, H belongs to the class of Hamiltonians discussed in Example 9.2.1, that is, the dynamics of the Euler top is described by the geodesic flow of Θ on the group manifold SO(3). Since H is invariant under Ψ, the Hamilton equations can be read off immediately from (9.2.9):[21]

$$a^{-1}\dot{a} = \Theta^{-1}L, \qquad \dot{L} = -\mathrm{ad}\big(\Theta^{-1}L\big)L. \tag{10.6.35}$$

The first equation restates that L corresponds to angular momentum. It yields $\Theta^{-1}L = \omega$. Then, the second equation takes the form[22] $\dot{L} = [L, \omega]$. Applying the isomorphism (5.2.6), we obtain

$$\dot{\mathbf{L}} + \boldsymbol{\omega} \times \mathbf{L} = 0. \tag{10.6.36}$$

These are the Euler equations for the force-free top. Equation (10.6.36) yields angular momentum conservation in the body frame.

Let us discuss symplectic reduction for this model. From Example 10.1.24 we know that the Hamiltonian system under consideration possesses a natural momentum mapping. Under the above identifications, it is given by

$$J : \mathrm{SO}(3) \times \mathfrak{so}(3) \to \mathfrak{so}(3), \qquad J(a, L) = \mathrm{Ad}(a)L, \tag{10.6.37}$$

cf. (10.1.27). Let $\mathrm{SO}(3)_L$ denote the stabilizer of L under the adjoint representation. According to Example 5.4.7, the adjoint representation of SO(3) is isomorphic to the fundamental representation of SO(3) on $\mathbb{R}^3$. Thus, for every $L \neq 0$, $\mathrm{SO}(3)_L$ is isomorphic to the subgroup $\mathrm{SO}(2) \subset \mathrm{SO}(3)$ of rotations about the axis defined by $\mathbf{L}$. For $L = 0$, we have $\mathrm{SO}(3)_L = \mathrm{SO}(3)$. Theorem 10.3.5 yields that $J^{-1}(L)$ is diffeomorphic to the group manifold SO(3) and that the symplectic manifold $(J^{-1}(L)/\mathrm{SO}(3)_L, \omega^L)$ provided by the Regular Reduction Theorem 10.3.1 is isomorphic to the adjoint orbit $\mathscr{O}_L$ through L, endowed with the negative Kirillov form. Thus,

$$J^{-1}(0)/\mathrm{SO}(3)_0 \cong \mathscr{O}_0 = \{0\}$$

and

$$J^{-1}(L)/\mathrm{SO}(3)_L \cong \mathscr{O}_L = \mathrm{S}^2_{\|L\|}, \quad L \neq 0. \tag{10.6.38}$$

Moreover, from (10.6.37) we read off that the bundle

$$\pi_L : J^{-1}(L) \to J^{-1}(L)/\mathrm{SO}(3)_L$$

coincides with the bundle of oriented 2-frames in $\mathbb{R}^3$.

Now, for a fixed nonzero element L_0 of $\mathfrak{so}(3)$, let us describe the reduced Hamiltonian system

$$\big(J^{-1}(L_0)/\mathrm{SO}(3)_{L_0}, \omega^{L_0}, H_{L_0}\big)$$

[21] We use the simplified notation $\Theta^{-1}L$ for the element of $\mathfrak{so}(3)$ corresponding to the vector $\Theta^{-1}\mathbf{L}$.

[22] A pair (L, ω), fulfilling this differential equation, is called a Lax pair and the equation is called Lax equation, see Sect. 11.2.

provided by Proposition 10.3.3. For that purpose, we identify $J^{-1}(L_0)/\mathrm{SO}(3)_{L_0}$ with the adjoint orbit $\mathscr{O}_{L_0}$ through L_0, that is, with the sphere $\mathrm{S}^2_{\|L_0\|}$ of radius $\|L_0\|$ in $\mathfrak{so}(3)$. Elements of $J^{-1}(L_0)$ are pairs (a, L) such that $L = \mathrm{Ad}(a^{-1})L_0$. Consequently, tangent vectors of $J^{-1}(L_0)$ at (a, L) are given by pairs $(\mathrm{L}'_a A, \mathrm{ad}(L)A)$ with $A \in \mathfrak{so}(3)$. Since the stabilizer $\mathrm{SO}(3)_{L_0}$ acts by $(h, (a, L)) \mapsto (ha, L)$, the natural projection $\pi_{L_0} : J^{-1}(L_0) \to \mathrm{S}^2_{\|L_0\|}$ and its tangent mapping are given by

$$\pi_{L_0}(a, L) = L, \qquad (\pi_{L_0})'_{(a,L)}\big(\mathrm{L}'_a A, \mathrm{ad}(L)A\big) = \mathrm{ad}(L)A,$$

respectively. Thus, from (10.6.35) we read off the reduced Hamiltonian and its Hamiltonian vector field:

$$H_{L_0}(L) = \frac{1}{2}\Theta^{-1}(L, L), \qquad (X_{H_{L_0}})_L = \mathrm{ad}(L)\big(\Theta^{-1}L\big), \quad L \in \mathrm{S}^2_{\|L_0\|}. \quad (10.6.39)$$

The corresponding Hamilton equations are

$$\dot{L} = \big[L, \Theta^{-1}L\big], \quad L \in \mathrm{S}^2_{\|L_0\|}, \qquad (10.6.40)$$

or, in vector notation,

$$\dot{\mathbf{L}} = \mathbf{L} \times \big(\Theta^{-1}\mathbf{L}\big), \qquad \|\mathbf{L}\| = \|\mathbf{L}_0\|.$$

Up to left translation by the stabilizer $\mathrm{SO}(3)_{L_0}$, every solution $t \mapsto L(t)$ of the reduced system (10.6.40) defines a solution $t \mapsto a(t)$ such that $t \mapsto (a(t), L(t))$ is a curve in $J^{-1}(L_0)$.

Remark 10.6.6 Using the above results, together with the fact that the Hamiltonian is a constant of motion, one can discuss the qualitative behaviour of integral curves. Assume that the body frame is chosen in such a way that the basis vectors $\mathbf{e}_i$ coincide with the principal axes of inertia. Then, $\Theta = \mathrm{diag}(\Theta_1, \Theta_2, \Theta_3)$ and the constants of motion $\mathbf{L}^2$ and H take the form

$$\mathbf{L}^2 = L_1^2 + L_2^2 + L_3^2, \qquad H = \frac{L_1^2}{2\Theta_1} + \frac{L_2^2}{2\Theta_2} + \frac{L_3^2}{2\Theta_3}.$$

Consequently, the integral curves $t \mapsto \mathbf{L}(t)$ of the reduced system are given by the lines of intersection of the sphere with radius $\|\mathbf{L}\|$ with the ellipsoid defined by H. For an exhaustive discussion we refer to [69]. There, the reader will also find a solution of the reduced Euler equations in terms of Jacobi's elliptic functions.

Example 10.6.7 (The spherical pendulum) Let $M = \mathrm{T}^*\mathbb{R}^3 \cong \mathrm{T}\mathbb{R}^3 \cong \mathbb{R}^3 \times \mathbb{R}^3$ endowed with the standard symplectic structure. Consider the restriction of the natural action of $\mathrm{SO}(3)$ on $\mathbb{R}^3$ to the subgroup $\mathrm{SO}(2)$ given by rotations about the axis of $\mathbf{e}_3$. In Example 10.1.23 we have seen that the Killing vector field generated by $A \in \mathfrak{so}(2)$ under the lift Ψ of this action to $\mathrm{T}\mathbb{R}^3$ is given by

$$A_*(\mathbf{x}, \mathbf{y}) = (A\mathbf{x}, A\mathbf{y})$$

and that the mapping

$$J : \mathrm{T}\mathbb{R}^3 \to \mathfrak{so}(2), \qquad \big\langle J(\mathbf{x}, \mathbf{y}), A\big\rangle = \mathbf{y} \cdot (A\mathbf{x}) \equiv (\mathbf{x} \times \mathbf{y}) \cdot \mathbf{A}, \quad (10.6.41)$$

is an equivariant momentum mapping. As before, the Lie algebra element A is identified with the vector $\mathbf{A} = \|\mathbf{A}\|\mathbf{e}_3$ in $\mathbb{R}^3$ via the isomorphism (5.2.6). Thus, according to (10.6.41), the momentum mapping is given by the projection of angular momentum to the $\mathbf{e}_3$-axis. To summarize, the tuple $(\mathrm{T}\mathbb{R}^3, \mathrm{d}\theta, \Psi, J)$ is a Hamiltonian SO(2)-manifold. Now, consider the SO(2)-invariant Hamiltonian

$$H(\mathbf{x},\mathbf{y}) = \frac{1}{2}\|\mathbf{y}\|^2 + \mathbf{x}\cdot\mathbf{e}_3 + \frac{1}{2}\big(\|\mathbf{y}\|^2 - \mathbf{x}\cdot\mathbf{e}_3\big)\big(\|\mathbf{x}\|^2 - 1\big) - \frac{1}{2}\mathbf{x}\cdot\mathbf{y}.$$

The corresponding Hamilton equations read

$$\dot{\mathbf{x}} = \mathbf{y} + \mathbf{y}\big(\|\mathbf{x}\|^2 - 1\big) - (\mathbf{x}\cdot\mathbf{y})\mathbf{x}, \tag{10.6.42}$$

$$\dot{\mathbf{y}} = -\mathbf{e}_3 + \frac{1}{2}\mathbf{e}_3\big(\|\mathbf{x}\|^2 - 1\big) - \big(\|\mathbf{y}\|^2 - \mathbf{x}\cdot\mathbf{e}_3\big)\mathbf{x} + (\mathbf{x}\cdot\mathbf{y})\mathbf{y}. \tag{10.6.43}$$

Since $\mathbf{x}\cdot\dot{\mathbf{x}} = 0$, the projection of the integral curve with initial conditions $\mathbf{x}_0$, $\mathbf{y}_0$ to the configuration space $\mathbb{R}^3$ is located on the sphere of radius $\|\mathbf{x}_0\|$. Thus, the Hamiltonian system $(\mathrm{T}\mathbb{R}^3, \omega, H)$ restricts to a Hamiltonian system on the submanifold

$$\mathrm{TS}^2 = \big\{(\mathbf{x},\mathbf{y}) \in \mathrm{T}\mathbb{R}^3 : \|\mathbf{x}\|^2 = 1,\ \mathbf{x}\cdot\mathbf{y} = 0\big\}, \tag{10.6.44}$$

cf. Remark 2.1.4/2. As in Example 10.1.23, one can check that the canonical 1-form θ of $\mathrm{T}\mathbb{R}^3$ pulls back to the canonical 1-form $\tilde{\theta}$ on TS^2. The restriction of H to TS^2 is given by

$$\tilde{H}(\mathbf{x},\mathbf{y}) = \frac{1}{2}\|\mathbf{y}\|^2 + \mathbf{x}\cdot\mathbf{e}_3 \tag{10.6.45}$$

and the Hamilton equations (10.6.42) and (10.6.43) become

$$\dot{\mathbf{x}} = \mathbf{y}, \qquad \dot{\mathbf{y}} = -\mathbf{e}_3 - \big(\|\mathbf{y}\|^2 - \mathbf{x}\cdot\mathbf{e}_3\big)\mathbf{x}. \tag{10.6.46}$$

We note that TS^2 is the phase space and $\tilde{H}$ is the Hamiltonian of the spherical pendulum.[23] Since Ψ leaves the submanifold TS^2 invariant, it induces a symplectic action

$$\tilde{\Psi} : \mathrm{SO}(2)\times\mathrm{TS}^2 \to \mathrm{TS}^2, \qquad \tilde{\Psi}(\mathbf{x},\mathbf{y}) = (a\cdot\mathbf{x}, a\cdot\mathbf{y}),$$

and the restriction $\tilde{J} = J_{\restriction \mathrm{TS}^2}$ yields an equivariant momentum mapping. We find

$$\tilde{J}(\mathbf{x},\mathbf{y}) = x_1 y_2 - x_2 y_1. \tag{10.6.47}$$

Thus, we obtain a Hamiltonian SO(2)-manifold $(\mathrm{TS}^2, \mathrm{d}\tilde{\theta}, \tilde{\Psi}, \tilde{J})$ with Hamiltonian $\tilde{H}$. The Noether Theorem 10.1.9 implies that $\tilde{J}$ is a constant of motion.

Let us discuss symplectic reduction for this system. First, observe that $\tilde{\Psi}$ is not free, because it has the fixed points

$$m_0^{\pm} = (\pm\mathbf{e}_3, 0).$$

Thus, we have to perform singular reduction. Since the adjoint action of the Abelian group SO(2) is trivial, for all values $j \in \mathfrak{so}(2)$ we have $\mathscr{O}_j = \{j\}$ and $G_j = \mathrm{SO}(2)$.

[23] With the gravitational acceleration set equal to 1.

The points $m_0^{\pm}$ have the stabilizer SO(2), whereas all other point of TS^2 have the trivial stabilizer $\{\mathbb{1}\}$. Consequently, the submanifolds of the above isotropy types are given by

$$M_{\{\mathbb{1}\}} = \mathrm{TS}^2 \setminus \{m_0^+ \cup m_0^-\}, \qquad M_{\mathrm{SO}(2)} = \{m_0^+\} \cup \{m_0^-\}.$$

For the intersections with $M_j = J^{-1}(j)$ we obtain

$$M_{\mathrm{SO}(2)} \cap M_j = \varnothing, \qquad M_{\{\mathbb{1}\}} \cap M_j = \{(\mathbf{x}, \mathbf{y}) \in \mathrm{TS}^2 : (\mathbf{x} \times \mathbf{y}) \cdot \mathbf{e}_3 = j\}$$

for $j \neq 0$ and

$$M_{\mathrm{SO}(2)} \cap M_0 = \{m_0^+\} \cup \{m_0^-\}, \qquad M_{\{\mathbb{1}\}} \cap M_0 = \{(\mathbf{x}, 0) \in \mathrm{TS}^2 : \mathbf{x} \neq \pm\mathbf{e}_3\}$$

for $j = 0$. Hence, in the first case, $\hat{M}_{\mathscr{O}_j}$ consists of the single orbit-momentum type stratum

$$\hat{M}_j = \{(\mathbf{x}, \mathbf{y}) \in \mathrm{TS}^2 : (\mathbf{x} \times \mathbf{y}) \cdot \mathbf{e}_3 = j\}/\mathrm{SO}(2)$$

and in the second case it consists of the three orbit-momentum type strata

$$\hat{M}_0 = \{(\mathbf{x}, 0) \in \mathrm{TS}^2 : \mathbf{x} \neq \pm\mathbf{e}_3\}/\mathrm{SO}(2), \qquad \hat{M}_{0+} = \{m_0^+\}, \qquad \hat{M}_{0-} = \{m_0^-\},$$

forming the principal stratum and the two secondary strata, respectively.

Remark 10.6.8

1. Let us analyse the structure of the strata found above by means of classical invariant theory, see [247] and [312]. The algebra $\mathbb{R}[\mathbf{x}, \mathbf{y}]^{\mathrm{SO}(2)}$ of invariant polynomials in the variables $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ is generated by the invariants

$$\tau_1 = x_3, \qquad \tau_2 = y_3, \qquad \tau_3 = y_1^2 + y_2^2,$$
$$\tau_4 = x_1 y_1 + x_2 y_2, \qquad \tau_5 = x_1^2 + x_2^2, \qquad \tau_6 = x_1 y_2 - x_2 y_1$$

   which fulfil the relations

$$\tau_4^2 + \tau_6^2 = \tau_3 \tau_5, \quad \tau_3 \geq 0, \ \tau_5 \geq 0. \tag{10.6.48}$$

   They define the so-called Hilbert mapping

$$\tau : \mathrm{T}\mathbb{R}^3 \to \mathbb{R}^6, \quad \tau(\mathbf{x}, \mathbf{y}) := \big(\tau_1(\mathbf{x}, \mathbf{y}), \ldots, \tau_6(\mathbf{x}, \mathbf{y})\big).$$

   In invariant theory it is shown that the topological quotient $\mathrm{T}\mathbb{R}^3/\mathrm{SO}(2)$ is homeomorphic to the image of τ, that is, to the subset of $\mathbb{R}^6$ defined by the relations (10.6.48).[24] In the present case, this can be checked by direct inspection, see [69]. Let us pass to new invariants σ_i defined by

$$\sigma_3 := \tau_3 + \tau_2^2, \qquad \sigma_i := \tau_i, \ i \neq 3.$$

   Restriction to TS^2 yields the additional relations

$$\sigma_5 + \sigma_1^2 = 1, \qquad \sigma_4 + \sigma_1 \sigma_2 = 0,$$

[24] A subset of $\mathbb{R}^n$ defined by equations and inequalities is said to be semialgebraic.

which allow for elimination of σ_4 and σ_5. Therefore, $\mathrm{TS}^2/\mathrm{SO}(2)$ can be identified with the image of the mapping

$$\sigma : \mathrm{TS}^2 \to \mathbb{R}^4, \qquad \sigma(\mathbf{x},\mathbf{y}) := \big(\sigma_1(\mathbf{x},\mathbf{y}), \sigma_2(\mathbf{x},\mathbf{y}), \sigma_3(\mathbf{x},\mathbf{y}), \sigma_6(\mathbf{x},\mathbf{y})\big),$$

which is the subset of $\mathbb{R}^4$ defined by the relations

$$\sigma_3\big(1-\sigma_1^2\big) - \sigma_2^2 - \sigma_6^2 = 0, \qquad \sigma_3 - \sigma_2^2 \geq 0, \qquad |\sigma_1| \leq 1.$$

According to (10.6.47), we have $\sigma_6 = \tilde{J}$. Therefore, $\hat{M}_{\mathscr{O}_j} = \tilde{J}^{-1}(j)/\mathrm{SO}(2)$ can be identified with the subset of $\mathbb{R}^3$ defined by the relations

$$\sigma_3\big(1-\sigma_1^2\big) - \sigma_2^2 - j^2 = 0, \qquad \sigma_3 - \sigma_2^2 \geq 0, \qquad |\sigma_1| \leq 1. \tag{10.6.49}$$

In case $j \neq 0$, this subset coincides with the graph of the function

$$\sigma_3(\sigma_1,\sigma_2) = \frac{j^2+\sigma_2^2}{1-\sigma_1^2}, \qquad |\sigma_1| < 1.$$

Hence, in this case we obtain a 2-dimensional smooth manifold diffeomorphic to $\mathbb{R}^2$, representing the orbit-momentum type stratum $\hat{M}_j$. In case $j = 0$, the subset (10.6.49) cannot be the graph of a function, because it contains the vertical lines $\{(\pm 1, 0, \sigma_3) \in \mathbb{R}^3, \sigma_3 \geq 0\}$. The endpoints $(\pm 1, 0, 0)$ of these lines correspond to the secondary strata $\hat{M}_{0\pm}$ of $\hat{M}_{\mathscr{O}_0}$ and all the rest corresponds to the principal stratum $\hat{M}_0$. Topologically, this subset is also isomorphic to $\mathbb{R}^2$.

2. From the above construction we can read off the preimages of the mapping

$$\pi_j \colon \tilde{J}^{-1}(j) \to \hat{M}_j. \tag{10.6.50}$$

For the point $(\pm 1, 0, 0)$, the preimage consists of the single point $m_0^{\pm}$. For all the other points, it is a 1-sphere.

3. Using the invariants, one can compute the Poisson brackets and, thus, determine the symplectic structure on the strata of the reduced phase space, see [69] and Exercise 10.6.9. One obtains

$$\{\sigma_i, \sigma_k\}_j = \sum_l \varepsilon_{ikl} \frac{\partial F_j}{\partial \sigma_l},$$

where

$$F_j(\sigma_1,\sigma_2,\sigma_3) = \sigma_3\big(1-\sigma_1^2\big) - \sigma_2^2 - j^2.$$

The reduced Hamiltonian can be read off from (10.6.45),

$$H_j = \frac{1}{2}\sigma_3 + \sigma_1, \tag{10.6.51}$$

and the Hamilton equations are given by

$$\dot{\sigma}_i = \{H_j, \sigma_i\}_j.$$

All these equation have to be viewed as equations on $\tilde{J}^{-1}(j)/\mathrm{SO}(2) = \sigma(\tilde{J}^{-1}(j))$. In particular, the relation $F_j(\sigma_1,\sigma_2,\sigma_3) = 0$ has to be taken into account.

Exercises

10.6.1 Prove the formulae (10.6.5) and (10.6.6). Show that the projections of integral curves of this system to the configuration space $\mathbb{R}^4$ are located on 3-spheres.

10.6.2 Verify Eq. (10.6.13).

10.6.3 In Example 10.6.1, show that the kernel of $\tilde{J}'_{(\mathbf{x},\mathbf{y})}$ is spanned by the geodesic vector field $X_{\tilde{H}}$.

10.6.4 Prove the formulae (10.6.15) and (10.6.16).

10.6.5 Prove that the fibre bundles given by (10.6.17) are isomorphic to the Stiefel bundle $\mathrm{S}_{\mathbb{R}}(2,4) \to \tilde{\mathrm{G}}_{\mathbb{R}}(2,4)$.
Hint. Find the preimages under the covering homomorphism (5.1.10) of the SO(2)-subgroups of SO(4) occurring in $\mathrm{S}_{\mathbb{R}}(2,4)$ and $\tilde{\mathrm{G}}_{\mathbb{R}}(2,4)$.

10.6.6 Prove Formula (10.6.28).

10.6.7 Show that for positive values of the energy, the hidden symmetry of the Kepler problem is given by the Lie algebra of SO(3, 1) and that for vanishing energy it is given by the Lie algebra of the group of Euclidean motions $\mathrm{SO}(3) \ltimes \mathbb{R}^3$, cf. Remark 10.6.4. Find the corresponding orbits.

10.6.8 Prove that the mapping (10.6.32) defines a symplectomorphism. Compute the pull-back of the Hamiltonian under this mapping.

10.6.9 Compute the Poisson brackets of Remark 10.6.8/2. Show that the symplectic form on the reduced phase space is given by

$$\omega = \frac{2}{\sigma_1^2 - 1} \mathrm{d}\sigma_1 \wedge \mathrm{d}\sigma_2.$$

10.7 A Model from Gauge Theory

In this section, we discuss a model arising from lattice approximation of SU(3)-gauge theory in the Hamiltonian approach. For a detailed presentation we refer to [92] and further references therein.[25]

Let Λ be a finite regular cubic lattice in $\mathbb{R}^3$. Let us denote the sets of oriented i-dimensional elements of Λ (sites, links, plaquettes and cubes) by Λ^i. The gauge group is $G = \mathrm{SU}(3)$ and its Lie algebra is $\mathfrak{g} = \mathfrak{su}(3)$. The $\mathfrak{g}$-valued gauge potential will be approximated on links by its G-valued parallel transporter:

$$\Lambda^1 \to G, \qquad (\mathbf{x}, \mathbf{y}) \mapsto a_{(\mathbf{x},\mathbf{y})}.$$

Thus, the configuration space and the phase space of the system are given by

$$Q = G^{\Lambda^1}, \qquad M = \mathrm{T}^* G^{\Lambda^1},$$

[25] This is part of a research program which aims at developing a non-perturbative approach to quantum gauge theory in the Hamiltonian framework, with special attention paid to the role of non-generic gauge orbit strata [144, 161–164], see also [254–256].

respectively. Using the natural identification $\mathrm{T}^*G^{\Lambda^1} \cong (\mathrm{T}^*G)^{\Lambda^1}$, the left trivialization $\mathrm{T}^*G \cong G \times \mathfrak{g}^*$ given by (8.3.6) and the isomorphism $\mathfrak{g}^* \cong \mathfrak{g}$ induced by the Ad-invariant scalar product $\langle A, B\rangle = -\operatorname{tr}(AB)$ on $\mathfrak{g}$, we may identify M with $G^{\Lambda^1} \times \mathfrak{g}^{\Lambda^1}$. Under this identification, the momentum canonically conjugate to the gauge potential (the colour-electric field) is given by a mapping

$$\Lambda^1 \to \mathfrak{g}, \qquad (\mathbf{x},\mathbf{y}) \mapsto A_{(\mathbf{x},\mathbf{y})}.$$

Local gauge transformations are approximated by mappings

$$\Lambda^0 \to G, \qquad \mathbf{x} \mapsto g_{\mathbf{x}},$$

acting on the parallel transporters by

$$a'_{(\mathbf{x},\mathbf{y})} = g_{\mathbf{x}} a_{(\mathbf{x},\mathbf{y})} g_{\mathbf{y}}^{-1}. \tag{10.7.1}$$

This defines an action of G^{Λ^0} on Q. Under the above identification, the lift of this action to $M = \mathrm{T}^*Q = G^{\Lambda^1} \times \mathfrak{g}^{\Lambda^1}$ is given by (10.7.1) and by

$$A'_{(\mathbf{x},\mathbf{y})} = \operatorname{Ad}(g_{\mathbf{x}}) A_{(\mathbf{x},\mathbf{y})}.$$

In lattice gauge theory, one considers the following gauge invariant Hamiltonian:

$$H = -\frac{\delta^3}{2} \sum_{(\mathbf{x},\mathbf{y})\in\Lambda^1} \operatorname{tr}\big(A^2_{(\mathbf{x},\mathbf{y})}\big) + \frac{1}{2\alpha^2\delta} \sum_{p\in\Lambda^2} \big(6 - \operatorname{tr}\big(a_p + a_p^\dagger\big)\big), \tag{10.7.2}$$

see [167] for the original source. Here, δ and α denote the lattice spacing and the coupling constant, respectively, and a_p is the parallel transporter around the plaquette $p = (\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u})$,

$$a_p = a_{(\mathbf{x},\mathbf{y})} a_{(\mathbf{y},\mathbf{z})} a_{(\mathbf{z},\mathbf{u})} a_{(\mathbf{u},\mathbf{x})}.$$

In what follows, we restrict ourselves to the simplest non-trivial case, where Λ consists of a single plaquette. Using a lattice tree, one can carry out an intermediate reduction which leaves one with a configuration space Q given by the group manifold G and with a phase space M given by the cotangent bundle $\mathrm{T}^*G \cong G \times \mathfrak{g}$. Let us denote the elements of $G \times \mathfrak{g}$ by (a, A) and let us write tangent vectors of $G \times \mathfrak{g}$ at (a, A) in the form $(\mathrm{L}'_a B, C)$ with $B, C \in \mathfrak{g}$. The group of local gauge transformations boils down to G itself and its action on $G \times \mathfrak{g}$ is given by

$$\Psi\big(g, (a, A)\big) = \big(gag^{-1}, \operatorname{Ad}(g)A\big). \tag{10.7.3}$$

According to (8.3.7), the canonical 1-form θ on $G \times \mathfrak{g}$ is given by

$$\theta_{(a,A)}\big(L'_a B, C\big) = \langle A, B\rangle \tag{10.7.4}$$

and for the equivariant momentum mapping (10.1.23), Example 10.1.25 yields

$$J : G \times \mathfrak{g} \to \mathfrak{g}, \qquad J(a, A) = aAa^{-1} - A. \tag{10.7.5}$$

Thus, $(M, \mathrm{d}\theta, \Psi, J)$ is a Hamiltonian G-manifold. Finally, the Hamiltonian (10.7.2) reduces to

$$H = -\frac{\delta^3}{2} \operatorname{tr}\big(A^2\big) + \frac{1}{2\alpha^2\delta}\big(6 - \operatorname{tr}\big(a + a^\dagger\big)\big). \tag{10.7.6}$$

In what follows, we shall discuss symplectic reduction for this system. From the point of view of physics, this corresponds to removing the (unphysical) gauge degrees of freedom. Therefore, the level set $M_0 = J^{-1}(0)$ is relevant, because $J(a, A) = 0$ is equivalent to the Gauß law.[26] Thus, we consider the reduced phase space

$$\hat{P} := \hat{M}_{\mathscr{O}_0} \equiv M_0/G. \tag{10.7.7}$$

From (10.7.5) we read off

$$M_0 = \big\{(a, A) \in G \times \mathfrak{g} : aA = Aa\big\}.$$

In particular, M_0 contains the subset $T \times \mathfrak{t}$, where $T \subset \mathrm{SU}(3)$ denotes the subgroup of diagonal matrices (maximal torus) and $\mathfrak{t}$ denotes its Lie algebra. By restriction, the natural projection $M_0 \to \hat{P}$ induces a mapping

$$\lambda : T \times \mathfrak{t} \to \hat{P}. \tag{10.7.8}$$

Let $(a, A) \in M_0$. Since a and A commute, they possess a common eigenbasis. Since a is unitary and A is anti-Hermitian, the eigenbasis can be chosen to be orthonormal. Hence, (a, A) is conjugate to an element of $T \times \mathfrak{t}$ under Ψ, that is, every G-orbit in M_0 intersects the submanifold $T \times \mathfrak{t}$. Hence, λ is surjective. Since two elements of $T \times \mathfrak{t}$ are conjugate under G iff they differ by a simultaneous permutation of their entries, λ descends to a bijection

$$(T \times \mathfrak{t})/\mathrm{S}_3 \to \hat{P},$$

where S_3 denotes the symmetric group on 3 symbols. Standard arguments ensure that this is in fact a homeomorphism.[27] Thus, we can use λ to describe $\hat{P}$.

Next, we determine the stabilizers and the orbit types of the elements of $T \times \mathfrak{t}$ under the actions of S_3 and of $\mathrm{SU}(3)$. The basic observation is that in both cases the stabilizer of (a, A) depends on the number of entries which simultaneously coincide for both a and A. This number can be 0, 2 or 3. Denote the corresponding subsets of $T \times \mathfrak{t}$ by, respectively, $(T \times \mathfrak{t})_2$, $(T \times \mathfrak{t})_1$ and $(T \times \mathfrak{t})_0$ and define

$$\hat{P}_i := \lambda\big((T \times \mathfrak{t})_i\big), \quad i = 2, 1, 0.$$

By restriction, λ induces mappings

$$\lambda_i : (T \times \mathfrak{t})_i \to \hat{P}_i, \quad i = 2, 1, 0, \tag{10.7.9}$$

which descend to homeomorphisms from $(T \times \mathfrak{t})_i/\mathrm{S}_3$ onto $\hat{P}_i$, $i = 2, 1, 0$. To determine the subsets $(T \times \mathfrak{t})_i$ explicitly, as in Example 6.6.6, let $Z \cong \mathbb{Z}_3$ denote the centre of $G = \mathrm{SU}(3)$, let T^j, $j = 1, 2, 3$, denote the subset of T consisting of the elements whose entries other than the jth one coincide and define $\mathfrak{t}^j$ analogously. For $j = 1, 2, 3$ let

$$(T \times \mathfrak{t})^j := \big(T^j \times \mathfrak{t}^j\big) \setminus \big(Z \times \{0\}\big).$$

---

[26] See [161].

[27] In particular, $\hat{P}$ is an orbifold.

We find

$$
\begin{aligned}
(T\times\mathfrak{t})_0 &= Z\times\{0\},\\
(T\times\mathfrak{t})_1 &= \bigcup_{j=1}^{3}(T\times\mathfrak{t})^j,\\
(T\times\mathfrak{t})_2 &= (T\times\mathfrak{t})\setminus\big((T\times\mathfrak{t})_1\cup(T\times\mathfrak{t})_0\big).
\end{aligned}
\tag{10.7.10}
$$

Obviously, all the subsets $(T\times\mathfrak{t})_i$ are embedded submanifolds of $T\times\mathfrak{t}$. Moreover, all the connected components $(T\times\mathfrak{t})^j$ of $(T\times\mathfrak{t})_1$ project to the same subset of $(T\times\mathfrak{t})/\mathrm{S}_3$. Under the action of S_3, the elements of $(T\times\mathfrak{t})_2$ have the trivial stabilizer $\{\mathbb{1}\}$ and hence the orbit type $[\{\mathbb{1}\}]$, the elements of $(T\times\mathfrak{t})_0$ have the stabilizer S_3 and hence the orbit type $[\mathrm{S}_3]$ and the elements of $(T\times\mathfrak{t})^j$ have the subgroup of order 2 generated by the transposition of the two coinciding entries as their stabilizer. Since the latter subgroups are conjugate in S_3, all the elements of $(T\times\mathfrak{t})_1$ have the same orbit type $[\mathrm{S}_2]$. To summarize, we have found that, under the action of S_3,

$$
(T\times\mathfrak{t})_2=(T\times\mathfrak{t})_{[\{\mathbb{1}\}]},\qquad (T\times\mathfrak{t})_1=(T\times\mathfrak{t})_{[\mathrm{S}_2]},\qquad (T\times\mathfrak{t})_0=(T\times\mathfrak{t})_{[\mathrm{S}_3]}.
$$

On the other hand, under the action of $\mathrm{SU}(3)$, the elements of $(T\times\mathfrak{t})_2$ have stabilizer T and hence orbit type $[T]$, the elements of $(T\times\mathfrak{t})_0$ have stabilizer $\mathrm{SU}(3)$ and hence orbit type $[\mathrm{SU}(3)]$ and the elements of $(T\times\mathfrak{t})^j$ have stabilizer $\mathrm{U}(2)^j$, where the subgroups $\mathrm{U}(2)^j$ are given by (6.6.3). Since the latter subgroups are conjugate in $\mathrm{SU}(3)$, all elements of $(T\times\mathfrak{t})_1$ have the same orbit type $[\mathrm{U}(2)]$. Thus,

$$
\hat{P}_2=\hat{P}_{[T]},\qquad \hat{P}_1=\hat{P}_{[\mathrm{U}(2)]},\qquad \hat{P}_0=\hat{P}_{[\mathrm{U}(3)]}.
$$

Recall that $\hat{P}=\hat{M}_{\mathscr{O}_0}$, so that with $\hat{P}_i$ we have determined the orbit-momentum type subsets $\hat{M}_{[H]}\cap\hat{M}_{\mathscr{O}}$ with coadjoint orbit $\mathscr{O}=\mathscr{O}_0$. From (10.7.10) we read off that $\hat{P}_2$ and $\hat{P}_1$ are connected and that $\hat{P}_0$ consists of three points labelled by the elements of Z. Thus, altogether there are five orbit-momentum type strata. By Theorem 10.5.4, each of these strata inherits the structure of a smooth symplectic manifold from $G\times\mathfrak{g}$. Taking into account that S_3 is finite, from (10.7.10) we can also read off the dimensions: $\hat{P}_2$ has dimension 4 and $\hat{P}_1$ has dimension 2.

One can use the mappings λ_i to describe the symplectic structure of the orbit-momentum type strata as follows. Since $\mathfrak{t}$ is the Lie subalgebra of $\mathfrak{su}(3)$ associated with the Lie subgroup T, the submanifold $T\times\mathfrak{t}$ is a symplectic submanifold of $G\times\mathfrak{g}$. Analogously, so are $(T\times\mathfrak{t})^j$, $j=1,2,3$. It follows that $(T\times\mathfrak{t})_2$ and $(T\times\mathfrak{t})_1$ are symplectic manifolds. For convenience, in the following we will view $(T\times\mathfrak{t})_0$ as a (trivial) symplectic manifold, too.

Proposition 10.7.1 *The mapping λ is Poisson. The mappings λ_k are local symplectomorphisms.*

Proof By definition, $C^\infty(\hat{P})$ is a quotient of $C^\infty(G\times\mathfrak{g})^G$, see Remark 10.5.11/3. Hence, the first assertion is a direct consequence of the fact that $T\times\mathfrak{t}$ is a symplectic submanifold of $G\times\mathfrak{g}$. The second assertion is trivial for $i=0$. For $i=1,2$ it

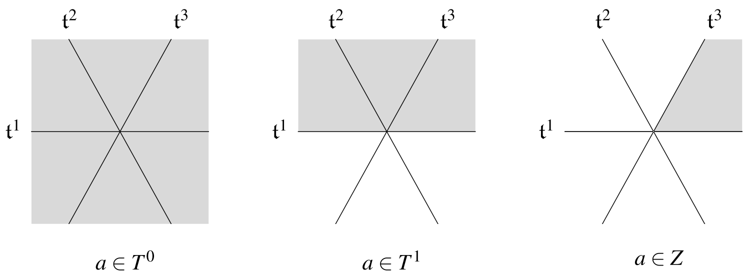


Fig. 10.1 The fibres $\hat{\pi}^{-1}([a])$ of the projection $\hat{\pi} : \hat{P} \to \hat{Q}$

follows by observing that any point of $\hat{P}_i$ has a representative (a, A) in $(T \times \mathfrak{t})_i$ and that a sufficiently small $G_{(a,A)}$-invariant neighbourhood W of (a, A) in $(T \times \mathfrak{t})_i$ generates a tube for the action of $G = G_0$ on $\pi^{-1}(\hat{P}_i) \cap M_0$, modelling the open neighbourhood $G \cdot W$ of the orbit $G \cdot (a, A)$ onto $G \times_{G_{(a,A)}} W \cong G/G_{(a,A)} \times W$. □

Remark 10.7.2

1. Since the submanifolds $(T \times \mathfrak{t})_i$ are symplectic and since S_3 is finite, the quotient $(T \times \mathfrak{t})/S_3$ naturally carries the structure of a stratified symplectic space, cf. Remark 10.5.11/3. In view of this, Proposition 10.7.1 states that λ descends to an isomorphism of stratified symplectic spaces from $(T \times \mathfrak{t})/S_3$ onto $\hat{P}$.
2. Dynamics on $\hat{P}$ with respect to an SU(3)-invariant Hamiltonian like (10.7.2) is thus given by the dynamics on $T \times \mathfrak{t}$ with respect to the corresponding S_3-invariant Hamiltonian and the symplectic form $\mathrm{d}\theta$, where θ is given by (10.7.4) with $a \in T$ and $A, B, C \in \mathfrak{t}$.

Now, let us analyse the projection

$$\hat{\pi} : \hat{P} \to \hat{Q}$$

induced by the cotangent bundle projection $\mathrm{T}^*Q \to Q$. We have the commutative diagram

$$\begin{array}{ccc} T \times \mathfrak{t} & \xrightarrow{\lambda} & \hat{P} \\ {\scriptstyle \mathrm{pr}_1}\downarrow & & \downarrow{\scriptstyle \hat{\pi}} \\ T & \longrightarrow & \hat{Q} \end{array}$$

where the lower horizontal arrow is defined by restriction of the natural projection $Q \to \hat{Q}$. Consequently, the fibre over $[a] \in \hat{Q}$ is given by

$$\hat{\pi}^{-1}\big([a]\big) = \mathfrak{t}/(S_3)_a,$$

where the representative a is chosen in T and where $(S_3)_a$ denotes the stabilizer of a under the action of S_3. There are 3 cases, illustrated in Fig. 10.1. As in Example 6.6.6, let T^0 denote the subset of T of elements with pairwise distinct entries.

(a) If $a \in T^0$, $(\mathrm{S}_3)_a$ is trivial, hence $\hat{\pi}^{-1}([a]) = \mathfrak{t}$. That is, the fibre is a full 2-plane and belongs to $\hat{P}_2$.
(b) If $a \in T^j \setminus Z$, $j = 1, 2, 3$, then $(\mathrm{S}_3)_a = \mathrm{S}_2$, acting by permutation of the coinciding entries. Hence, $\hat{\pi}^{-1}([a]) = \mathfrak{t}/\mathrm{S}_2$, acting by reflection about the subspace $\mathfrak{t}^j$. Therefore, the fibre may be identified with one of the two closed half-planes of $\mathfrak{t}$ cut out by $\mathfrak{t}^j$. Its interior belongs to $\hat{P}_2$, whereas the boundary $\mathfrak{t}^j$ belongs to $\hat{P}_1$.
(c) If $a \in Z$, then $(\mathrm{S}_3)_a = \mathrm{S}_3$. The action of S_3 on $\mathfrak{t}$ is generated by the reflections about the 3 subspaces $\mathfrak{t}^j$, $j = 1, 2, 3$. Hence, $\hat{\pi}^{-1}([a])$ may be identified with one of the six closed subsets of $\mathfrak{t}$ (the Weyl chambers) cut out by $\mathfrak{t}^j$, $j = 1, 2, 3$ (the walls of the Weyl chambers). The interior of the Weyl chamber chosen belongs to $\hat{P}_2$, the walls minus the origin belong to $\hat{P}_1$ and the origin belongs to $\hat{P}_0$.

Remark 10.7.3

1. Recall from Example 6.6.6 that the reduced configuration space $\hat{Q}$ coincides with the orbit space of the action of S_3 on the subgroup $T \subset \mathrm{SU}(3)$ of diagonal matrices by permuting the entries,
$$\hat{Q} = T/\mathrm{S}_3.$$
From the above discussion it is obvious that the projection $\hat{\pi} : \hat{P} \to \hat{Q}$ does not preserve the stratification, because the fibres over points in the secondary[28] orbit type strata of $\hat{Q}$, corresponding to the edges and the vertices of the 2-simplex structure of $\hat{Q}$, intersect more than one orbit-momentum type stratum of $\hat{P}$.
2. The description of the reduced data given here generalizes to an arbitrary compact semisimple Lie group in an obvious way: T and $\mathfrak{t}$ are replaced by a maximal torus in G and the corresponding Cartan subalgebra of $\mathfrak{g}$. $\hat{Q}$ is replaced by a Weyl alcove in T and S_3 is replaced by the Weyl group of G.
   In the simpler case $G = \mathrm{SU}(2)$, on the other hand, we obtain $\hat{P} \cong (\mathrm{S}^1 \times \mathbb{R})/\mathrm{S}_2$, where the generator of S_2 acts by complex conjugation in the first component and by multiplication by -1 in the second one. This way, the reduced phase space is getting identified with an orbifold usually referred to as the canoe, see Fig. 10.2. It consists of three orbit-momentum type strata: the two points $[(\pm\mathbb{1}, 0)]$ and the rest, which is a 2-dimensional symplectic manifold diffeomorphic to the 2-punctured 2-plane. For a detailed discussion of its symplectic structure, see [142]. From Example 10.6.7 it is clear that $\hat{P}$ coincides with the reduced phase space of the spherical pendulum with zero angular momentum.
3. As we know from Proposition 10.5.9, the orbit-momentum type strata of $\hat{P}$ can also be obtained via regular reduction of the connected components $\Sigma_{H,\mu}$ of the subsets $M_H \cap M_\mu$ by the quotient Lie groups $(\Gamma_{\Sigma_H})_\mu$. For $\mu = 0$ and for the choice of the subgroups T and $\mathrm{U}(2)^1$ as representatives of the corresponding orbit types, we find that

[28] Cf. Remark 6.6.2/5.

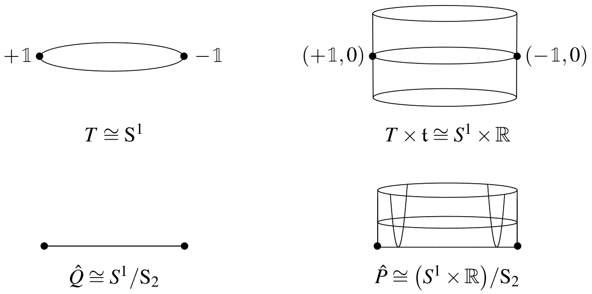


Fig. 10.2 Reduced configuration space $\hat{Q}$ and reduced phase space $\hat{P}$ for $G = \mathrm{SU}(2)$

(a) the subsets $M_H \cap M_\mu$ are given by

$$M_T \cap M_0 = (T \times \mathfrak{t})_2, \qquad M_{\mathrm{U}(2)^1} \cap M_0 = (T \times \mathfrak{t})^1,$$
$$M_{\mathrm{SU}(3)} \cap M_0 = Z \times \{0\},$$

where the last one decomposes into the connected components $\{(a,0)\}$, $a \in Z$;

(b) the groups $(\Gamma_{\Sigma_H})_0 \equiv \Gamma_{\Sigma_H}$ are given by

$$\Gamma_{(T\times\mathfrak{t})_2} \cong S_3, \qquad \Gamma_{(T\times\mathfrak{t})^1} \cong \{\mathbb{1}\}, \qquad \Gamma_{\{(a,0)\}} \cong \{\mathbb{1}\}, \quad a \in Z.$$

4. For a detailed description of $\hat{Q}$ and $\hat{P}$ in terms of invariants, including the stratification, we refer to [92]. Using standard results of the invariant theory of complex matrices one can show that the algebra of invariant real polynomials on M_0 is generated by

$$c_k := \mathrm{Re}\big(\mathrm{tr}\big(a(-iA)^k\big)\big), \quad k = 0,1,2,$$
$$d_k := \mathrm{Im}\big(\mathrm{tr}\big(a(-iA)^k\big)\big), \quad k = 0,1,2,$$
$$t_k := \mathrm{tr}\big((-iA)^k\big), \quad k = 2,3.$$

By setting $A = 0$ one obtains a set of generators for the invariant real polynomials on $Q = G$:

$$c_0 = \mathrm{Re}\,\mathrm{tr}(a), \qquad d_0 = \mathrm{Im}\,\mathrm{tr}(a).$$

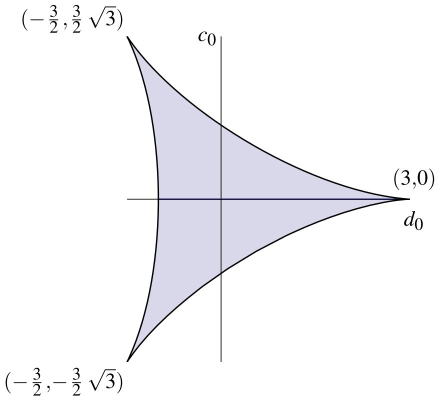
The associated Hilbert mappings

$$\rho_{\hat{P}} = (c_0, d_0, c_1, d_1, c_2, d_2, t_2, t_3) : M_0 \to \mathbb{R}^8, \qquad \rho_{\hat{Q}} = (c_0, d_0) : Q \to \mathbb{R}^2$$

induce homeomorphisms from $\hat{P}$ and $\hat{Q}$ onto the respective images, which are semialgebraic subsets of $\mathbb{R}^8$ and $\mathbb{R}^2$, respectively. Consider the following polynomials:

$$R_1 = \big(3 + c_0^2 + d_0^2\big)t_2 - 2\big(c_1^2 + d_1^2\big) - 4(c_0c_2 + d_0d_2),$$
$$R_2 = \left(3 - \frac{1}{3}\big(c_0^2 + d_0^2\big)\right)t_3 - 2(c_1c_2 + d_1d_2),$$
$$R_3 = c_0c_2 - d_0d_2 - 2c_0t_2 - c_1^2 + d_1^2 + 3c_2,$$

Fig. 10.3 The image of $\rho_{\hat{Q}}$ in $\mathbb{R}^2$



$$R_4 = c_0 d_2 + d_0 c_2 + 2d_0 t_2 - 2c_1 d_1 - 3d_2,$$
$$R_5 = \frac{1}{2}\big((c_0 - 1)d_1 + d_0 c_1\big)t_2 + \left(\frac{2}{3}c_0 d_0 + d_0\right)t_3 - c_1 d_2 - d_1 c_2,$$
$$P_1 = 27 - c_0^4 - 2c_0^2 d_0^2 - d_0^4 + 8c_0^3 - 24c_0 d_0^2 - 18c_0^2 - 18d_0^2,$$
$$P_2 = \frac{1}{2}t_2^3 - 3t_3^2,$$
$$P_3 = t_2^2 - c_2^2 - d_2^2.$$

The image of $\rho_{\hat{Q}}$ is given by the inequality $P_1 \geq 0$. It forms a region in $\mathbb{R}^2$ bounded by a hypocycloid, see Fig. 10.3. As is shown in [92], the defining equations and inequalities for the image of $\rho_{\hat{P}}$ are given by

$$R_i = 0, \quad i = 1, \ldots, 5, \qquad P_i \geq 0, \quad i = 1, 2, 3.$$

In a similar way, one can derive the defining relations for the images under $\rho_{\hat{P}}$ of the orbit-momentum type strata.

5. It turns out that the reduced configuration space is a deformation retract of the reduced phase space, see [95]. Therefore, to study the topology of the reduced phase space amounts to studying the topology of the reduced configuration space. For case studies in the context of lattice gauge theory, see [61, 62].

10.8 The Energy Momentum Mapping

Let (M, ω, Ψ, J) be a Hamiltonian G-manifold with G-invariant Hamiltonian H. As we know from the Noether Theorem 10.1.9, the functions J_A are constants of motion for all $A \in \mathfrak{g}$, that is, the level sets of H and of J are invariant under the flow of the Hamiltonian vector field X_H. Thus, the dynamics reduces to the level sets of the combined mapping

$$\mathscr{E}\colon M \to \mathbb{R} \times \mathfrak{g}^*, \quad \mathscr{E}(m) := \big(H(m), J(m)\big). \tag{10.8.1}$$

Studying the partition of M into these level sets is at the heart of a general programme formulated by Smale, see [277, 278], as well as [1] and [69]. In this context, the notion of the bifurcation set of a mapping plays a key role.

Definition 10.8.1 (Bifurcation set) A smooth mapping $\varphi: M \to N$ is called locally trivial at $p_0 \in N$ if there exists a neighbourhood $U \subset N$ of p_0, such that the following conditions are fulfilled:

1. The level set $\varphi^{-1}(p)$ is an embedded submanifold of M for all $p \in U$.
2. There exists a smooth mapping $\Phi: \varphi^{-1}(U) \to \varphi^{-1}(p_0)$, such that
$$\varphi \times \Phi: \varphi^{-1}(U) \to U \times \varphi^{-1}(p_0)$$
is a diffeomorphism.

The bifurcation set $\mathscr{B}_\varphi$ of φ is the set of all points in N, for which φ is not locally trivial.

Point 2 means that $\varphi_{\restriction \varphi^{-1}(U)}: \varphi^{-1}(U) \to U$ is a trivial fibre bundle over U with typical fibre $\varphi^{-1}(p_0)$. In particular, then $(\varphi \times \Phi)_{\restriction \varphi^{-1}(p)}$ yields a diffeomorphism between the level sets $\varphi^{-1}(p_0)$ and $\varphi^{-1}(p)$. As p runs through the bifurcation set, the topology of the level set $\varphi^{-1}(p)$ may change. The programme of Smale can be summarized as follows:

1. Determine the topological type of the level sets of $\mathscr{E}$ and its bifurcation set.
2. Determine the flow of X_H on each level set.
3. Investigate how the fibres $\mathscr{E}^{-1}(h, \mu)$ fit together as (h, μ) runs through the values of $\mathscr{E}$.

The following two observations are helpful for working out this programme in concrete cases. The first one is of general nature and concerns the bifurcation set. Let M_φ be the set of critical points of φ. Then, $\varphi(M_\varphi)$ is the set of its critical values. The proof of the following proposition is left to the reader (Exercise 10.8.1).

Proposition 10.8.2 *For a smooth mapping $\varphi: M \to N$, one has $\varphi(M_\varphi) \subset \mathscr{B}_\varphi$. If φ is proper, then $\varphi(M_\varphi) = \mathscr{B}_\varphi$.*

The second observation is a consequence of the specific structure of the energy-momentum mapping. By the Regular Reduction Theorem 10.3.3 for invariant Hamiltonian systems, the reduced Hamiltonian is given by
$$H_\mu \circ \pi_\mu = H \circ j_\mu,$$
see (10.3.5). Thus,
$$(H_{\restriction J^{-1}(\mu)})^{-1}(h) = \pi_\mu^{-1} \circ H_\mu^{-1}(h)$$
and, therefore,
$$\mathscr{E}^{-1}(h, \mu) = H^{-1}(h) \cap J^{-1}(\mu) = (H_{\restriction J^{-1}(\mu)})^{-1}(h) = \pi_\mu^{-1}\big(H_\mu^{-1}(h)\big). \quad (10.8.2)$$

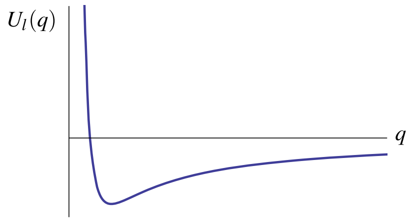


Fig. 10.4 The effective potential (10.8.4)

This identity reduces the study of the topology of the fibres $\mathscr{E}^{-1}(h,\mu)$ to computing the preimages of the level sets of the reduced Hamiltonian and taking the preimages under π_μ. An analogous statement holds for the singular case.

In what follows we discuss essential aspects of the programme of Smale for the Kepler problem, the Euler top and the spherical pendulum.

Example 10.8.3 (Kepler problem) According to Example 10.6.3, we have

$$M = \mathrm{T}\big(\mathbb{R}^3 \setminus \{0\}\big) \cong \big(\mathbb{R}^3 \setminus \{0\}\big) \times \mathbb{R}^3$$

and under the identification of $\mathfrak{so}(3)$ with $\mathbb{R}^3$ via the isomorphism (5.2.6), the energy-momentum mapping is given by

$$\mathscr{E} = (H, J) : \mathrm{T}\big(\mathbb{R}^3 \setminus \{0\}\big) \to \mathbb{R} \times \mathbb{R}^3$$

with

$$H(\mathbf{q},\mathbf{p}) = \frac{\mathbf{p}^2}{2} - \frac{1}{q}, \qquad J(\mathbf{q},\mathbf{p}) = \mathbf{q} \times \mathbf{p} = \mathbf{L}.$$

Let $(h,\mathbf{L}) \in \mathbb{R} \times \mathbb{R}^3$ and denote $l = \|\mathbf{L}\|$. By (10.8.2), we have

$$\mathscr{E}^{-1}(h,\mathbf{L}) = \pi_{\mathbf{L}}^{-1}\big(H_l^{-1}(h)\big),$$

with the reduced Hamiltonian H_l and the natural projection

$$\pi_{\mathbf{L}} : J^{-1}(\mathbf{L}) \to J^{-1}(\mathbf{L})/\mathrm{SO}(3)_{\mathbf{L}}. \tag{10.8.3}$$

As in Example 10.6.3, we have to distinguish the following two cases.

(a) In case $l > 0$, the reduced Hamiltonian H_l is given by (10.6.18) and we have a one-dimensional motion with effective potential

$$U_l(q) = \frac{l^2}{2q^2} - \frac{1}{q}, \tag{10.8.4}$$

see Fig. 10.4. The reduced Hamiltonian H_l possesses the single critical point $(q^0, p_q^0) = (l^2, 0)$ and hence the single critical value $h^0 = -\frac{1}{2l^2}$. Using

$$p_q = \pm\sqrt{2\big(h - U_l(q)\big)},$$

from the shape of the effective potential we can read off the topology of the level sets $H_l^{-1}(h)$, see Fig. 10.5: For $h^0 < h < 0$ we get a sphere S^1, which in

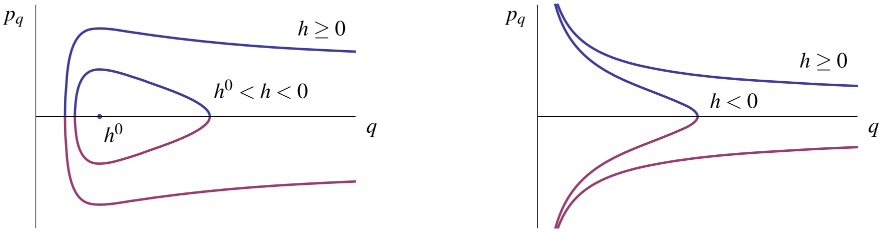


Fig. 10.5 Level sets $H_l^{-1}(h)$ in the Kepler problem for $l > 0$ (*left*) and $l = 0$ (*right*)

the limit $h \to h^0$ descends to the 1-point space. For $h \geq 0$ one obtains the real line $\mathbb{R}$. Since for $l > 0$, the preimages of the projection (10.8.3) are isomorphic to S^1, we obtain

$$\mathscr{E}^{-1}(h, \mathbf{L}) = \begin{cases} S^1 & \text{for } h = h^0, \\ T^2 & \text{for } h^0 < h < 0, \\ \mathbb{R} \times S^1 & \text{for } h \geq 0. \end{cases} \tag{10.8.5}$$

(b) In case $l = 0$, under the identification of the reduced phase space with the submanifold $\Sigma \subset M$ given by (10.6.19) with some chosen nonzero $\mathbf{q}_0 \in \mathbb{R}^3$, the integral curves and hence the level sets of the reduced Hamiltonian H_l are of the form

$$\left\{ \left(q\mathbf{e}, \pm\sqrt{2\left(h + \frac{1}{q}\right)}\mathbf{e} \right) : 0 < q \leq q_{\max} \right\}$$

with $\mathbf{e} = \frac{\mathbf{q}_0}{\|\mathbf{q}_0\|}$, see Fig. 10.5. For $h < 0$, we have $q_{\max} = -\frac{1}{h}$ and hence $H_l^{-1}(h) \cong \mathbb{R}$. For $h \geq 0$ we find $q_{\max} = \infty$ and hence $H_l^{-1}(h) \cong S^0 \times \mathbb{R}$. Since for $l = 0$ the preimages of the projection (10.8.3) are isomorphic to S^2, we obtain

$$\mathscr{E}^{-1}(h, 0) = \begin{cases} \mathbb{R} \times S^2 & \text{for } h < 0 \\ S^0 \times \mathbb{R} \times S^2 & \text{for } h \geq 0. \end{cases} \tag{10.8.6}$$

From (10.8.5) and (10.8.6) we read off that the bifurcation set is the preimage of the set

$$\left\{ \left(-\frac{1}{2l^2}, l \right) : l > 0 \right\} \cup \left\{ (0, l) : l \geq 0 \right\} \cup \left\{ (h, 0) : h \in \mathbb{R} \right\} \subset \mathbb{R}^2, \tag{10.8.7}$$

under the mapping

$$\mathbb{R} \times \mathfrak{so}(3) \to \mathbb{R}^2, \qquad (h, L) \mapsto \big(h, \|L\|\big), \tag{10.8.8}$$

see Fig. 10.6. We encourage the reader to study the Hamiltonian flow on the level sets given by (10.8.5) and (10.8.6) and to compare these results with what he knows from classical mechanics, see Exercise 10.8.2. In the next chapter, we will return to the case $h^0 < h < 0$, $\mathbf{L} \neq 0$, in the context of action and angle variables.

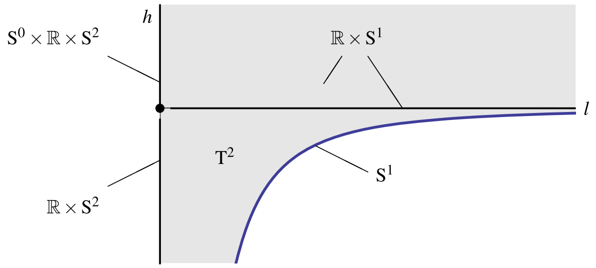


Fig. 10.6 Image of the energy-momentum mapping $\mathscr{E}$ of the Kepler problem, projected to the l–h-plane. The (projected) bifurcation set consists of the critical values of $\mathscr{E}$, represented by the hyperbola, and of the two coordinate axes. The regions are labelled by the topological type of the level sets of $\mathscr{E}$ according to (10.8.5) and (10.8.6)

Example 10.8.4 (Euler top) According to Example 10.6.5, we have

$$M = \mathrm{T}^*\mathrm{SO}(3) \cong \mathrm{SO}(3) \times \mathfrak{so}(3)$$

and the energy-momentum mapping is given by

$$\mathscr{E} = (H, J) \colon \mathrm{SO}(3) \times \mathfrak{so}(3) \to \mathbb{R} \times \mathfrak{so}(3)$$

with

$$H(a, L) = \frac{1}{2}\Theta^{-1}(L, L), \qquad J(a, L) = \mathrm{Ad}(a)L.$$

Let $(h, L_0) \in \mathbb{R} \times \mathfrak{so}(3)$ be a fixed value of $\mathscr{E}$ and denote $l := \|L_0\|$. Recall that the reduced phase space may be identified with the adjoint orbit $\mathscr{O}_{L_0}$ endowed with the negative Kirillov structure and that the latter coincides with the origin for $l = 0$ and with the sphere S^2_l of radius l in $\mathfrak{so}(3)$ otherwise. Moreover, in the latter case, the reduced Hamiltonian reads

$$H_l(L) = \frac{1}{2}\Theta^{-1}(L, L), \quad L \in S^2_l.$$

To find its level sets, we first have to find the critical points. For that purpose, we pass to vectors $\mathbf{L} \in S^2_l \subset \mathbb{R}^3$ and choose the body frame in such a way that

$$\Theta = \begin{bmatrix} \Theta_1 & 0 & \\ 0 & \Theta_2 & 0 \\ 0 & 0 & \Theta_3 \end{bmatrix}.$$

For simplicity we assume that

$$0 < \lambda_3 = \Theta_3^{-1} < \lambda_2 = \Theta_2^{-1} < \lambda_1 = \Theta_1^{-1} \tag{10.8.9}$$

and leave the discussion of the other cases to the reader (Exercise 10.8.3). The critical points of H_l can be determined from

$$\big(H_l'' + \alpha f''\big)(\mathbf{L}) = 0, \qquad f(\mathbf{L}) = \frac{1}{2}\big(\|\mathbf{L}\|^2 - l^2\big),$$

with $\mathbf{L} \in \mathbb{R}^3$ and α denoting a Lagrange multiplier. We obtain the equations

$$(\lambda_1 L_1, \lambda_2 L_2, \lambda_3 L_3) - \alpha(L_1, L_2, L_3) = 0, \qquad L_1^2 + L_2^2 + L_3^2 = l^2,$$

that is, $\mathbf{L}$ is located on the intersection of the eigenspaces of the matrix Θ^{-1} with S_l^2. This yields the critical points

$$\mathbf{L}_i^{\pm} = \pm l\mathbf{e}_i, \quad i = 1, 2, 3. \tag{10.8.10}$$

Remark 10.8.5 Since the critical points of H_l coincide with the critical points of the corresponding Hamiltonian vector field X_{H_l}, they can be read off from (10.6.39), too. Thus, on the level of $\mathfrak{so}(3)$, they are given by the solutions of the equation $[\Theta^{-1}L, L] = 0$ and hence by the eigenvectors of Θ^{-1}, indeed.

Next, we show that H_l is a Morse function. For that purpose, we observe that for every i, the tangent space

$$\mathrm{T}_{\mathbf{L}_i^{\pm}} \mathrm{S}_l^2 = \ker f'(\mathbf{L}_i^{\pm}) = \{\mathbf{X} \in \mathbb{R}^3 : \mathbf{X} \cdot \mathbf{L}_i^{\pm} = 0\}$$

is spanned by the two standard basis vectors $\mathbf{e}_j$ and $\mathbf{e}_k$ with $j, k \neq i$. We compute

$$\mathrm{Hess}_{\mathbf{L}_1^{\pm}}(H_l) = \begin{bmatrix} \lambda_2 - \lambda_1 & 0 \\ 0 & \lambda_3 - \lambda_1 \end{bmatrix},$$
$$\mathrm{Hess}_{\mathbf{L}_2^{\pm}}(H_l) = \begin{bmatrix} \lambda_1 - \lambda_2 & 0 \\ 0 & \lambda_3 - \lambda_2 \end{bmatrix},$$
$$\mathrm{Hess}_{\mathbf{L}_3^{\pm}}(H_l) = \begin{bmatrix} \lambda_1 - \lambda_3 & 0 \\ 0 & \lambda_2 - \lambda_3 \end{bmatrix}.$$

Since $\lambda_1 - \lambda_2 > 0$, $\lambda_1 - \lambda_3 > 0$ and $\lambda_2 - \lambda_3 > 0$, the critical points $\mathbf{L}_1^{\pm}, \mathbf{L}_2^{\pm}, \mathbf{L}_3^{\pm}$ are non-degenerate and have the Morse indices, respectively, 2, 1 and 0.

Remark 10.8.6 Using Proposition 9.7.1, from the Hessians we read off that the critical point $\mathbf{L}_i^{\pm}$ of the Hamiltonian vector field X_{H_l} on S_l^2 is stable for $i = 1, 3$ and unstable for $i = 2$. It follows that these points are also stable or unstable with respect to the projected flow on $\mathrm{T}^*\mathrm{SO}(3)/\mathrm{SO}(3) \cong \mathfrak{so}(3)$ as a whole, see [181, Thm. III.12.4]. Thus, the corresponding stationary motions in $\mathrm{T}^*\mathrm{SO}(3)$, given by rotation about the corresponding principal axes of inertia, are relatively SO(3)-stable or SO(3)-unstable, cf. Remark 6.8.5/4.

Proposition 10.8.7 *The level sets of H_l have the following topological structure.*

1. *For $l = 0$, one has $h = 0$ and $H_l^{-1}(h)$ consists of a single point.*
2. *For $l > 0$ and $h = \frac{1}{2}l^2\lambda_i$, $i = 1, 3$, the level set $H_l^{-1}(h)$ consists of two points.*
3. *For $l > 0$ and $h \in (\frac{1}{2}l^2\lambda_3, \frac{1}{2}l^2\lambda_2) \cup (\frac{1}{2}l^2\lambda_2, \frac{1}{2}l^2\lambda_1)$, one has $H_l^{-1}(h) = \mathrm{S}^1 \sqcup \mathrm{S}^1$.*
4. *For $l > 0$ and $h = \frac{1}{2}l^2\lambda_2$, the level set $H_l^{-1}(h)$ is the union of two distinct great circles on S^2.*

Proof 1. As noted above, in this case the reduced phase space consists of a single point.

2. For $h = \frac{1}{2}l^2\lambda_3$, we have

$$\lambda_1 L_1^2 + \lambda_2 L_2^2 + \lambda_3 L_3^2 = l^2\lambda_3, \qquad L_1^2 + L_2^2 + L_3^2 = l^2.$$

If we multiply the second equation by λ_3 and subtract it from the first one, we obtain

$$(\lambda_1 - \lambda_3)L_1^2 + (\lambda_2 - \lambda_3)L_2^2 = 0.$$

Under the assumption (10.8.9), this implies $L_1 = L_2 = 0$ and hence $\mathbf{L} = \mathbf{L}_3^{\pm}$.

3. The critical points $\mathbf{L}_3^{\pm}$ are non-degenerate minima with Morse index 0. Hence, according to the Morse Lemma 8.9.3, for values of h between $\frac{1}{2}l^2\lambda_3$ and $\frac{1}{2}l^2\lambda_3 + \varepsilon$ with ε small enough, the level set $H_l^{-1}(h)$ is the union of two 1-spheres in the neighbourhood of the points $\mathbf{L}_3^{+}$ and $\mathbf{L}_3^{-}$. An analogous statement is true for values of h between $\frac{1}{2}l^2\lambda_1 - \varepsilon$ and $\frac{1}{2}l^2\lambda_1$. Then, the assertion follows from the Morse Isotopy Lemma 8.9.6.

4. For $h = \frac{1}{2}l^2\lambda_2$, an analogous calculation as under point 2 yields the equation

$$(\lambda_2 - \lambda_3)L_3^2 - (\lambda_1 - \lambda_2)L_1^2 = 0.$$

This defines two planes $P_{\pm}$ intersecting in the subspace spanned by $\mathbf{e}_2$. Hence,

$$H_l^{-1}(h) = (P_+ \cup P_-) \cap \mathrm{S}_l^2.$$

This set consists of two circles which intersect in the points $\mathbf{L}_2^{\pm}$. □

In the following, we write $L_i^{\pm}$ for the elements of $\mathfrak{so}(3)$ corresponding to the critical points $\mathbf{L}_i^{\pm}$.

Proposition 10.8.8 *For* $L_0 \neq 0$, *the function* $H_{\restriction J^{-1}(L_0)}$ *is an* $\mathrm{SO}(3)_{L_0}$*-invariant Morse-Bott function with* 6 *non-degenerate critical* $\mathrm{SO}(3)_{L_0}$*-orbits*

$$\gamma_i^{\pm} = \pi_{L_0}^{-1}\left(L_i^{\pm}\right), \quad i = 1, 2, 3,$$

having Morse indices, respectively, 2, 1 *and* 0.

Proof Due to $H \circ j_{L_0} = H_l \circ \pi_{L_0}$, for any $p \in J^{-1}(L_0)$, we have

$$(H_{\restriction J^{-1}(L_0)})'_p = (H_l)'_{\pi_{L_0}(p)} \circ (\pi_{L_0})'_p.$$

Since π_{L_0} is a submersion, this implies that p is a critical point of $H_{\restriction J^{-1}(L_0)}$ iff $\pi_{L_0}(p)$ is a critical point of H_{L_0}. The subset of critical points of $H_{\restriction J^{-1}(L_0)}$ is thus the union of the subsets $\gamma_i^{\pm}$, $i = 1, 2, 3$. Obviously, every $\gamma_i^{\pm}$ is an $\mathrm{SO}(3)_{L_0}$-orbit. Since $\mathrm{SO}(3)_{L_0} \cong \mathrm{SO}(2)$, it is therefore diffeomorphic to a 1-sphere.

It remains to show that $H_{\restriction J^{-1}(L_0)}$ is a Morse function and to determine its Morse index at $\gamma_i^{\pm}$. For that purpose, we choose an open neighbourhood $U_i^{\pm}$ of the critical point $L_i^{\pm}$ in S_l^2 and a local trivialization

$$\chi_i^{\pm} : \pi_{L_0}^{-1}\left(U_i^{\pm}\right) \to U_i^{\pm} \times \mathrm{SO}(3)_{L_0}.$$

Using $\chi_i^\pm$, the $\mathrm{SO}(3)_{L_0}$-orbit $\gamma_i^\pm$ can be parameterized by

$$t \mapsto \left(\chi_i^\pm\right)^{-1}\left(L_i^\pm, \exp(tA)\right),$$

where $A \in \mathfrak{so}(3)_{L_0} \cong \mathfrak{so}(2)$, $A \neq 0$. For every fixed value t_0, the submanifold

$$\mathscr{S}_i^\pm = \left(\chi_i^\pm\right)^{-1}\left(U_i^\pm \times \exp(t_0 A)\right) \tag{10.8.11}$$

intersects $\gamma_i^\pm$ transversally in the point $(\chi_i^\pm)^{-1}(L_i^\pm, \exp(t_0 A))$. This shows that H_l is a Morse-Bott function. Since $(\pi_{L_0})_{\restriction \mathscr{S}_i^\pm} : \mathscr{S}_i^\pm \to U_i^\pm$ is a diffeomorphism, the Morse index of $H_{\restriction J^{-1}(L_0)}$ at $\gamma_i^\pm(t_0)$ coincides with the Morse index of H_l at $L_i^\pm$. □

Proposition 10.8.9 *The level sets of the energy-momentum mapping $\mathscr{E}$ for the Euler top have the following topological structure.*

1. *For $L_0 = 0$, one has $h = 0$ and $\mathscr{E}^{-1}(h, L_0) = \mathrm{SO}(3)$.*
2. *For $L_0 \neq 0$ and $h = \frac{1}{2}l^2\lambda_1$ or $h = \frac{1}{2}l^2\lambda_3$, one has $\mathscr{E}^{-1}(h, L_0) = \mathrm{S}^1 \sqcup \mathrm{S}^1$.*
3. *For $L_0 \neq 0$ and $h \in (\frac{1}{2}l^2\lambda_3, \frac{1}{2}l^2\lambda_2) \cup (\frac{1}{2}l^2\lambda_2, \frac{1}{2}l^2\lambda_1)$, $\mathscr{E}^{-1}(h, L_0) = \mathrm{T}^2 \sqcup \mathrm{T}^2$.*
4. *For $L_0 \neq 0$ and $h = \frac{1}{2}l^2\lambda_2$, the level set $\mathscr{E}^{-1}(h, L_0)$ is a union of two tori intersecting along two circles.*

Proof 1. This is obvious.

3. Proposition 10.8.7 implies $H_l^{-1}(h) = \mathrm{S}^1 \sqcup \mathrm{S}^1$. Each of these two 1-spheres is the boundary of a 2-disc D_i^2, $i = 1, 2$, which is contractible in S_l^2. It follows that the bundle $\pi_{L_0} : J^{-1}(L_0) \to \mathrm{S}_l^2$ is trivial over D_i^2. Thus, $\pi_{L_0}^{-1}(\mathrm{D}_i^2) \cong \mathrm{D}_i^2 \times \mathrm{S}^1$ and hence

$$\mathscr{E}^{-1}(h, L_0) = \pi_{L_0}^{-1}\left(H_l^{-1}(h)\right) \cong \left(\partial \mathrm{D}_1^2 \times \mathrm{S}^1\right) \sqcup \left(\partial \mathrm{D}_2^2 \times \mathrm{S}^1\right).$$

2. This is a limiting case of point 3. According to Proposition 10.8.7, here $H_l^{-1}(h)$ consists of two points. Over each of them we have a fibre S^1.

4. According to Proposition 10.8.7, in this case, $H_l^{-1}(h)$ is the union of two circles S_i^1 on S_l^2, which intersect in the points $L_i^\pm$. As under point 3, these circles are the boundaries of two 2-discs and we obtain 2-tori, which intersect along the critical curves $\gamma_i^\pm$ of Proposition 10.8.8. □

According to Proposition 10.8.9, the bifurcation set of the energy-momentum mapping $\mathscr{E}$ of the Euler top is the union of three paraboloids,

$$\mathscr{B}_{\mathscr{E}} = \bigcup_{i=1}^{3} \left\{ (h, L) \in \mathbb{R} \times \mathfrak{so}(3) : \frac{1}{2}\lambda_i \|L\|^2 = h \right\},$$

see Fig. 10.7. We encourage the reader to find out how the level sets of $\mathscr{E}$ fit together to build $J^{-1}(L_0)$ for a fixed L_0, see Exercise 10.8.4.

Example 10.8.10 (Spherical pendulum) According to Example 10.6.7, we have

$$M = \mathrm{TS}^2 = \left\{(\mathbf{x}, \mathbf{y}) \in \mathrm{T}\mathbb{R}^3 : \|\mathbf{x}\| = 1,\ \mathbf{x} \cdot \mathbf{y} = 0\right\} \tag{10.8.12}$$

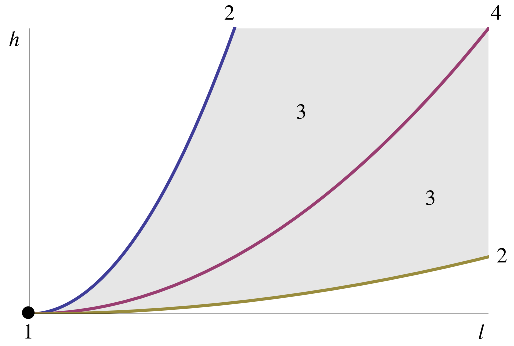


Fig. 10.7 Image of the energy-momentum mapping $\mathscr{E}$ of the Euler top, projected to the l–h-plane via the mapping (10.8.8) (shaded region). The (projected) bifurcation set consists of the three parabolas corresponding to λ_1 (narrow), λ_2 (medium) and λ_3 (wide). The numbers refer to the cases 1–4 of Proposition 10.8.9

and the energy momentum mapping is given by

$$\mathscr{E} = (H, J) : \mathrm{TS}^2 \to \mathfrak{so}(2) \cong \mathbb{R}$$

with[29]

$$H(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\|\mathbf{y}\|^2 + \mathbf{x} \cdot \mathbf{e}_3, \qquad J(\mathbf{x}, \mathbf{y}) = x_1 y_2 - x_2 y_1. \tag{10.8.13}$$

To find the critical points of $\mathscr{E}$, we compute the Hamiltonian vector fields (Exercise 10.8.5):

$$\begin{aligned} X_H(\mathbf{x}, \mathbf{y}) &= \big(\mathbf{y}, -\mathbf{e}_3 - \big(\|\mathbf{y}\|^2 - \mathbf{x} \cdot \mathbf{e}_3\big)\mathbf{x}\big), \\ X_J(\mathbf{x}, \mathbf{y}) &= \big((-x_2, x_1, 0), (-y_2, y_1, 0)\big). \end{aligned} \tag{10.8.14}$$

The critical points of X_H are

$$m_\pm^0 = (0, 0, \pm 1, 0, 0, 0). \tag{10.8.15}$$

At these points, we have

$$H\big(m_\pm^0\big) = \pm 1, \qquad J\big(m_\pm^0\big) = 0, \qquad X_J\big(m_\pm^0\big) = 0.$$

The points where X_J and X_H are parallel are determined by the equations

$$\begin{aligned} -x_2 &= \lambda y_1, & -y_2 &= -\lambda\big(\|\mathbf{y}\|^2 - x_3\big)x_1, \\ x_1 &= \lambda y_2, & y_1 &= -\lambda\big(\|\mathbf{y}\|^2 - x_3\big)x_2, \\ 0 &= \lambda y_3, & 0 &= -1 - \big(\|\mathbf{y}\|^2 - x_3\big)x_3, \end{aligned}$$

with $\lambda \in \mathbb{R}$. First, we conclude that $\lambda^2 = -x_3$, that is, every $\lambda \in [-1, 1]$ defines a horizontal plane intersecting S^2 in the critical point m_-^0 for $\lambda = \pm 1$ and in a 1-sphere otherwise. In spherical coordinates

$$x_1 = \cos\phi \sin\vartheta, \qquad x_2 = \sin\phi \sin\vartheta, \qquad x_3 = \cos\vartheta,$$

---

[29]Resetting $H = \tilde{H}$ and $J = \tilde{J}$.

this 1-sphere is parameterized by the angle ϕ. In these coordinates, the solution of the above system of equations reads

$$y_1 = \pm \frac{\sin\vartheta \sin\phi}{\sqrt{-\cos\vartheta}}, \qquad y_2 = \pm \frac{\sin\vartheta \cos\phi}{\sqrt{-\cos\vartheta}}, \qquad y_3 = 0. \tag{10.8.16}$$

The corresponding values of the constants of motion are

$$H(\vartheta) = \frac{3}{2}\cos\vartheta - \frac{1}{2\cos\vartheta}, \qquad J(\vartheta) = \pm \frac{1-\cos^2\vartheta}{\sqrt{-\cos\vartheta}}. \tag{10.8.17}$$

This is the parameter presentation of a curve $s \mapsto \gamma(s)$, $s = \cos\vartheta$, in the space of values of the energy-momentum mapping $\mathscr{E}$. The union of this curve with the isolated point $(H(m_+^0), J(m_+^0)) = (1, 0)$ yields the set of critical values of $\mathscr{E}$, see Fig. 10.8. The isolated point is obtained for the parameter value $\cos\vartheta = 1$. The curve γ can be parameterized by the value j of J or by the value h of H. Let h_j and j_h denote the corresponding value of the other constant of motion. While the function $j \mapsto h_j$ cannot be expressed in terms of elementary functions, for the function $h \mapsto j_h$ one finds

$$j_h = \pm \frac{2}{9}\left(3 - h^2 + h\sqrt{h^2+3}\right)\sqrt{h + \sqrt{h^2+3}} \tag{10.8.18}$$

(Exercise 10.8.6). It remains to determine the set of regular values of $\mathscr{E}$. For that purpose, we note that H can be written in the form

$$H = \frac{y_3^2}{2\sin^2\vartheta} + \frac{J^2}{2\sin^2\vartheta} + \cos\vartheta.$$

For given j, the critical points of $H_{\restriction J^1(j)}$ are determined by the equations

$$y_3 = 0, \qquad j^2\cos\vartheta + \sin^4\vartheta = 0.$$

Comparison with (10.8.16) and (10.8.17) shows that the latter equations define a critical point with critical value h_j. From $H \to \infty$ for $\cos\vartheta \to \pm 1$ we conclude that this critical point is a minimum. Thus, the regular points of $\mathscr{E}$ are located above the curve γ and, therefore, the image of the energy-momentum mapping is given by

$$\mathscr{E}\left(\mathrm{TS}^2\right) = \left\{(h, j) \in \mathbb{R}^2 : h \geq h_j,\ j \in \mathbb{R}\right\}.$$

Proposition 10.8.11 *The level sets of the energy-momentum mapping $\mathscr{E}$ for the spherical pendulum have the following topological structure.*

1. *For $h > -1$, $|j| < |j(h)|$ and $(h, j) \neq (1, 0)$, the level set $\mathscr{E}^{-1}(h, j)$ is a 2-torus.*
2. *For $h > -1$ and $j = \pm j(h)$, the level set $\mathscr{E}^{-1}(h, j)$ is a 1-sphere.*
3. *For $(h, j) = (-1, 0)$, the level set $\mathscr{E}^{-1}(h, j)$ consists of a single point.*
4. *For $(h, j) = (1, 0)$, the level set $\mathscr{E}^{-1}(h, j)$ is a 2-torus, in which a longitudinal circle has been contracted to a point.*

The results are summarized in Fig. 10.8.

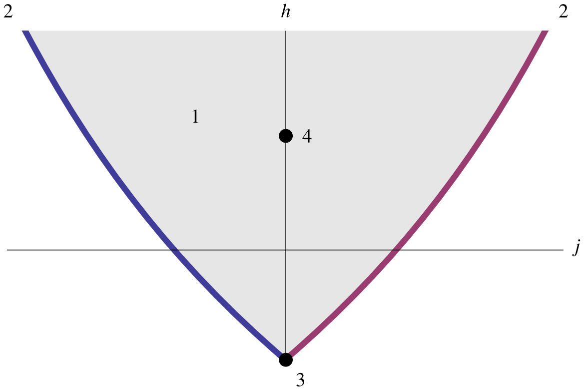


Fig. 10.8 Image of the energy-momentum mapping $\mathscr{E}$ of the spherical pendulum (shaded region). The bifurcation set $\mathscr{B}_{\mathscr{E}}$ consists of the boundary and the point $(j,h)=(0,1)$. The numbers refer to the cases 1–4 of Proposition 10.8.11

Proof 1. By Remark 10.6.8/2, the fibres of the natural projection $\pi_j\colon J^{-1}(j)\to \hat{M}_{\mathscr{O}_j}$ are 1-spheres and $\hat{M}_{\mathscr{O}_j}$ may be identified with the subset of $\mathbb{R}^3$ defined by the relations (10.6.49). Expressing the reduced Hamiltonian (10.6.51) in terms of the invariant polynomials σ_1, σ_3,

$$H_j = \frac{1}{2}\sigma_3 + \sigma_1,$$

we find that these relations can be written in the form

$$2(H_j - \sigma_1)\big(1-\sigma_1^2\big) - \sigma_2^2 - j^2 = 0, \quad H_j \geq \frac{\sigma_2^2}{2} + \sigma_1, \ |\sigma_1| \leq 1. \tag{10.8.19}$$

This means that for every regular value h, the level set H_j is the 1-sphere which forms the boundary of the 2-disc defined by the two inequalities in (10.8.19). Since this implies that the bundle $\mathscr{E}^{-1}(h,j)\to H_j^{-1}(h)$ is trivial, (10.8.2) yields the assertion.

2. By (10.8.16) and (10.8.17), the level set can be parameterized by ϕ as follows:

$$\phi \mapsto \left(\sin\vartheta\cos\phi, \sin\vartheta\sin\phi, \cos\vartheta, \mp\frac{\sin\vartheta\sin\phi}{\sqrt{-\cos\vartheta}}, \pm\frac{\sin\vartheta\cos\phi}{\sqrt{-\cos\vartheta}}, 0\right).$$

3. This is obtained from point 2 by taking the limit $\sin\vartheta\to 0$.

4. According to Remark 10.6.8/2, for regular points of $\hat{M}_0$, the preimages of the projection

$$\pi_0\colon J^{-1}(0)\to \hat{M}_0$$

are 1-spheres. For the two singular points $p_\pm = (\pm 1, 0, 0)$ of $\hat{M}_0$, they consist of the corresponding single points $m_0^\pm = (0,0,\pm 1, 0,0,0)$. Since $h = 1 = \frac{1}{2}\sigma_3 + \sigma_1$, it follows that only p_+ lies in the image of π_0. This completes the proof. □

To conclude, we stress that the occurrence of tori in the above models is no accident, because all of them are integrable systems. This will be the topic of the next chapter.

Exercises

10.8.1 Prove Proposition 10.8.2.

10.8.2 Study the properties of the Hamiltonian flow of the Kepler problem on the level sets given by Eqs. (10.8.5) and (10.8.6).

10.8.3 Study the topological structure of the energy-momentum mapping of the Euler top for the case $\lambda_3 < \lambda_2 = \lambda_1$.

10.8.4 Discuss the structure of the foliation of the level set $J^{-1}(L_0)$ for the Euler top. Find out how the fibres fit together to constitute $J^{-1}(L_0)$ for a fixed value L_0.

Hint. One possible strategy consists in using $J^{-1}(L_0) \cong \mathrm{SO}(3)$, together with the fact that the group manifold of SO(3) can be viewed as a ball of radius π with identified antipodes.

10.8.5 Compute the Hamiltonian vector fields generated by the Hamiltonian H and the momentum mapping J of Example 10.8.10.

10.8.6 Verify Formula (10.8.18).

10.8.7 Calculate the critical points of the Hamiltonian (10.8.13) and the corresponding Morse indices.

Chapter 11
Integrability

In this chapter, we study the concept of integrability of Hamiltonian systems in a systematic way. We start with the very notion of an integrable system and with a number of examples: the two-body problem, the two-centre problem, the top, the spherical pendulum and the Toda lattice. In Sect. 11.2, we analyze Lax pairs in the context of Hamiltonian systems on coadjoint orbits. In particular, we show that the Toda lattice can be understood in this framework. In Sects. 11.3 and 11.4 we analyze the local geometric structure of integrable systems. We prove the Arnold Theorem, discuss the relation with symplectic reduction and present the construction of local action and angle variables in detail. Thereafter, we construct action and angle variables for a number of examples. We also show that action and angle variables are very well adapted to the study of small perturbations of integrable systems. In this context, we meet another application of KAM theory. In Sect. 11.7 we give an introduction to global aspects in the spirit of Nekhoroshev and Duistermaat, with some emphasis on monodromy. This phenomenon will be illustrated in detail for the case of the spherical pendulum. Finally, we present a generalization to the concept of so-called non-commutative integrability in the sense of Mishchenko and Fomenko. We prove the Mishchenko-Fomenko Theorem and illustrate it for the case of the Euler top.

11.1 Basic Notions and Examples

Definition 11.1.1 A Hamiltonian system (M, ω, H) of dimension $2n$ is called integrable[1] if there exist n constants of motion $H_1 = H, H_2, \ldots, H_n$ satisfying the following conditions:

1. The functions H_i are in involution, that is, $\{H_i, H_j\} = 0$ for all $i, j = 1, \ldots, n$.
2. The subset of regular points of the mapping $\mathscr{H} = (H_1, \ldots, H_n)\colon M \to \mathbb{R}^n$ is dense in M.

[1] Or completely integrable.

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7_11,

For an integrable system we will write $(M, \omega, H_1, \dots, H_n)$ or $(M, \omega, \mathscr{H})$. Recall that the set of singular points of the mapping $\mathscr{H}$ is denoted by $M_{\mathscr{H}}$ and that the set of regular points is denoted by $M^{\mathscr{H}}$. Thus, we have the disjoint decomposition $M = M^{\mathscr{H}} \cup M_{\mathscr{H}}$, where $M^{\mathscr{H}}$ is an open and dense submanifold of M.

Remark 11.1.2

1. The Hamiltonian vector fields X_{H_i} span a distribution on M, which will be denoted by $D^{\mathscr{H}}$. Since $[X_{H_i}, X_{H_j}] = X_{\{H_i, H_j\}} = 0$, this distribution is involutive. By Proposition 3.2.13/2, its rank is constant along the integral curves of the Hamiltonian vector fields X_{H_i}. Hence, Theorem 3.5.10 implies that $D^{\mathscr{H}}$ is integrable and Proposition 3.5.21 yields that the integral manifolds form a foliation of M. Since $\omega(X_{H_i}, X_{H_j}) = \{H_i, H_j\} = 0$, the distribution $D^{\mathscr{H}}$ is isotropic and hence its integral manifolds are isotropic submanifolds of M.
2. The open submanifold $M^{\mathscr{H}}$ of M consists of the points where the differentials $\mathrm{d}H_i$ and hence the Hamiltonian vector fields X_{H_i} are linearly independent, that is, where $D^{\mathscr{H}}$ has the maximal rank n. Since the rank of $D^{\mathscr{H}}$ is constant along the integral curves of the X_{H_i}, $M^{\mathscr{H}}$ is invariant under their flows. In particular, the dynamics reduces to $M^{\mathscr{H}}$. Moreover, the restriction
$$\mathscr{H}_r : M^{\mathscr{H}} \to \mathbb{R}^n \tag{11.1.1}$$
of $\mathscr{H}$ to $M^{\mathscr{H}}$ is a submersion.[2]
3. The restriction $D^{\mathscr{H}_r}$ of $D^{\mathscr{H}}$ to $M^{\mathscr{H}}$ is a regular distribution of rank n. Since it is isotropic, it is Lagrange. Moreover, due to $X_{H_j}(H_j) = \{H_j, H_i\} = 0$, we have $D^{\mathscr{H}_r} \subset \ker \mathscr{H}_r{}'$ and by counting dimensions, we find that equality holds. Hence, according to Example 3.5.4/4, the maximal integral manifolds of $D^{\mathscr{H}_r}$ are given by the level set components of $\mathscr{H}_r$. In particular, the latter are Lagrangian submanifolds of $M^{\mathscr{H}}$. To summarize, the level set components of $\mathscr{H}_r$ form a Lagrangian foliation of $M^{\mathscr{H}}$ which is generated by the Lagrangian distribution $D^{\mathscr{H}_r}$ and which is invariant under the flows of the Hamiltonian vector fields X_{H_i}. In particular, the dynamics reduces to these level set components.

Obviously, every autonomous system with one degree of freedom is integrable. Moreover, by Proposition 9.1.10, in an autonomous system, every constant of motion is in involution with H. In particular, an autonomous system with two degrees of freedom is integrable if only it admits a second constant of motion, functionally independent from H. In higher dimensions, integrability is rare. We start with a number of examples.

Example 11.1.3 (Two-Body Problem) The two-body problem has $n = 6$ degrees of freedom given by the position vectors $\mathbf{q}_1$ and $\mathbf{q}_2$ of the two bodies, that is, the phase

[2]Beware that the set of regular values of $\mathscr{H}$ is contained in the set of values of $\mathscr{H}_r$ but need not coincide with the latter.

space is $M = \mathrm{T}^*(\mathbb{R}^3 \times \mathbb{R}^3 \setminus \Delta_{\mathbb{R}^3})$, where $\Delta_{\mathbb{R}^3}$ denotes the diagonal. For definiteness, we consider the Hamiltonian[3]

$$H(\mathbf{p}_1, \mathbf{p}_2, \mathbf{q}_1, \mathbf{q}_2) = \frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} - \frac{k}{\|\mathbf{q}_1 - \mathbf{q}_2\|}. \tag{11.1.2}$$

By Galilei invariance, we have 10 constants of motion: the total momentum $\mathbf{P}$, the total angular momentum $\mathbf{L}$, the total energy E and the vector $M\mathbf{R} - t\mathbf{P}$, with M denoting the total mass and $\mathbf{R}$ being the position vector of the centre of mass. By separating out the motion of the centre of mass, for the relative motion we obtain the phase space $\mathrm{T}^*(\mathbb{R}^3 \setminus \{0\})$ and the constants of motion

$$H_{\mathrm{rel}} = \frac{\mathbf{p}^2}{2\mu} - \frac{k}{\|\mathbf{q}\|}, \qquad \mathbf{L}_{\mathrm{rel}} = \mathbf{q} \times \mathbf{p}.$$

Here,

$$\mu = \frac{m_1 m_2}{m_1 + m_2}, \qquad \mathbf{q} = \mathbf{q}_1 - \mathbf{q}_2, \qquad \mathbf{p} = \mu \dot{\mathbf{q}}$$

are, respectively, the reduced mass, the relative position and the relative momentum. One easily checks that

$$\mathbf{P}, \quad H_{\mathrm{rel}}, \quad \mathbf{L}_{\mathrm{rel}}^2, \quad (L_{\mathrm{rel}})_3 \tag{11.1.3}$$

are constants of motion in involution and that their differentials are independent on a dense subset of M (Exercise 11.1.1). Thus, the 2-body problem is integrable.

Example 11.1.4 (Two-Centre Problem) It turns out that already the 3-body problem is in general not integrable. Let us restrict our attention to the case where the mass of one of the three bodies is assumed to be so small that it does not influence the motion of the other two bodies. This approximation leads to the so-called two-centre problem,[4] where one studies the motion of a body in the field of two fixed force centres. The Hamiltonian of this model is given by[5]

$$H = \frac{1}{2}\mathbf{p}^2 + \frac{k_1}{r_1} + \frac{k_2}{r_2}, \tag{11.1.4}$$

with $\mathbf{p}$ denoting the momentum and r_i denoting the distances of the body from the two centres. For simplicity, we have set the mass of the body equal to one. Coulson and Joseph have shown that the two-centre problem can be treated in the

---

[3]The Kepler potential can be replaced by any central force potential.

[4]Some authors call this the Three-body problem of Euler, because Euler solved it first and published it in his memoirs in 1760.

[5]It turns out that this is a good approximation provided the electromagnetic forces between the particles dominate. It is not a good approximation if the gravitational forces dominate, see Sect. 4.3 in [286] for a discussion of this point. We also refer to §47 of [18] for an application to the motion of the moon in the gravitational field of the earth.

same way for any number n of degrees of freedom [68]. Following these authors, we choose Cartesian coordinates q_i such that the two centres are located at the points $\mathbf{a} = (0, \dots, a)$ and $-\mathbf{a}$. Then, the configuration space is $Q = \mathbb{R}^n \setminus \{\mathbf{a}, -\mathbf{a}\}$ and the distances of the body from the centres are given by

$$r_1^2 = \sum_{i=1}^{n-1} q_i^2 + (q_n - a)^2, \qquad r_2^2 = \sum_{i=1}^{n-1} q_i^2 + (q_n + a)^2. \tag{11.1.5}$$

Angular momentum is represented by the skew-symmetric n-dimensional matrix

$$L_{ij} = q_i p_j - q_j p_i. \tag{11.1.6}$$

We define

$$L^2 = \sum_{i<j} L_{ij}^2, \qquad L_\perp^2 = \sum_{i<j<n} L_{ij}^2. \tag{11.1.7}$$

By direct inspection, one can check the following (Exercise 11.1.2).

(a) The angular momentum components L_{ij} with $i, j < n$ are constants of motion.
(b) The quantity

$$A = \frac{1}{2}\left(L^2 + a^2 p_n^2\right) + a q_n \left(\frac{k_1}{r_1} - \frac{k_2}{r_2}\right) \tag{11.1.8}$$

is a constant of motion which Poisson-commutes with all components L_{ij}, $i, j < n$, and hence with $L_\perp^2$. The reader can find the relation of A to the Lenz-Runge vector of the Kepler problem by sending one of the centres to infinity.

We conclude that the quantities H, A and $L_\perp^2$ are constants of motion in involution. For $n \geq 3$, their differentials are linearly independent on a dense subset of T^*Q. For $n = 2$, the so-called planar two-centre problem, this still holds for H and A but one has $L_\perp = 0$. As a result, for $n = 2$ or 3, the two-centre problem is integrable.

Example 11.1.5 (Top) We take up Example 10.6.5. Recall from there that the phase space of a top is

$$\mathrm{T}^*\mathrm{SO}(3) \cong \mathrm{SO}(3) \times \mathfrak{so}(3)$$

and that in the special case of the Euler top, the equations of motion have the form of a Lax equation for the Lax pair (L, ω). We continue to identify $\mathfrak{so}(3)$ with $\mathbb{R}^3$ via the isomorphism (5.2.6), denoting the vector corresponding to a Lie algebra element A by $\mathbf{A}$. Here, we consider the more general case of a top in a constant[6] external force field $\mathbf{f}$. The most important example of this type is, of course, the gravitational field $\mathbf{f} = -mg\mathbf{n}_3$ of the earth. Before turning to the discussion of integrability, let us derive the Hamilton equations and the Poisson structure.

[6]With respect to the inertial frame.

In the body frame, the external field is given by the vector

$$\mathbf{F} = a^{-1}\mathbf{f}$$

and the potential has the form

$$V(a) = -\mathbf{F} \cdot \mathbf{S},$$

where $\mathbf{S}$ denotes the position vector of the centre of mass. On the level of the Lie algebra, the relation between F and f reads $F = \mathrm{Ad}(a^{-1})f$ and the Hamiltonian function takes the form

$$H(a, L) = \frac{1}{2}\Theta^{-1}(L, L) - \langle F, S \rangle, \tag{11.1.9}$$

where $\langle \cdot, \cdot \rangle$ denotes the Ad-invariant scalar product defined by

$$\langle F, S \rangle = -\frac{1}{2}\,\mathrm{tr}(FS) \equiv \mathbf{F} \cdot \mathbf{S}.$$

The Hamilton equations for this system are given by (9.2.8). To find their explicit form, under the identification of $\mathfrak{so}(3)^*$ with $\mathfrak{so}(3)$ induced by $\langle \cdot, \cdot \rangle$ we compute

$$\big\langle \big(\mathrm{L}'_a\big)^{\mathrm{T}}\big((H_L)'_a\big), A \big\rangle = (H_L)'_a\big(\mathrm{L}'_a A\big) = -\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0}\big\langle \mathrm{Ad}\big(\exp(-tA)\big)F, S \big\rangle = \big\langle \mathrm{ad}(F)S, A \big\rangle.$$

Hence,

$$\big(\mathrm{L}'_a\big)^{\mathrm{T}}\big((H_L)'_a\big) = \mathrm{ad}(F)S$$

and the Hamilton equations are given by

$$a^{-1}\dot{a} = \Theta^{-1}L, \qquad \dot{L} = -\mathrm{ad}\big(\Theta^{-1}L\big)L - \mathrm{ad}(F)S. \tag{11.1.10}$$

The first equation is called the Poisson equation and the second one is called the Euler equation. In vector notation it is given by

$$\dot{\mathbf{L}} = \mathbf{L} \times \boldsymbol{\omega} + \mathbf{S} \times \mathbf{F}. \tag{11.1.11}$$

Since $\mathbf{S} \times \mathbf{F} \equiv \mathbf{N}$ is the torque acting on the system, the Euler equation is equivalent to the angular momentum balance equation in the body frame,

$$\dot{\mathbf{L}} + \boldsymbol{\omega} \times \mathbf{L} = \mathbf{N}. \tag{11.1.12}$$

To find constants of motion in involution, we determine the Poisson structure. For that purpose, we view L and F as mappings $\mathrm{SO}(3) \times \mathfrak{so}(3) \to \mathfrak{so}(3)$, given by

$$L(a, A) = A, \qquad F(a, A) = \mathrm{Ad}\big(a^{-1}\big)f.$$

Let L_i and F_i denote the corresponding coefficient functions with respect to the basis $\{e_i \equiv \mathrm{I}_i^{\mathbb{R}}\}$ of Example 5.2.8. Using the method of Example 9.2.2, we compute the Hamiltonian vector fields

$$X_{L_i}(a, A) = \big(L'_a e_i, -\mathrm{ad}(e_i)A\big), \qquad X_{F_i}(a, A) = \big(0, \mathrm{ad}(F)e_i\big) \tag{11.1.13}$$

(Exercise 11.1.3) and read off the Poisson brackets from (9.2.10),

$$\{L_i, L_j\} = \varepsilon_{ij}{}^k L_k, \qquad \{F_i, F_j\} = 0, \qquad \{L_i, F_j\} = \varepsilon_{ij}{}^k F_k. \tag{11.1.14}$$

Remark 11.1.6 Due to (11.1.14), the subspace of $C^\infty(\mathrm{SO}(3) \times \mathfrak{so}(3))$ spanned by the functions L_i and F_i forms a Lie subalgebra. This Lie subalgebra is isomorphic to the Lie algebra of the Euclidean group $\mathrm{E}(3) = \mathrm{SO}(3) \ltimes \mathbb{R}^3$, see Exercise 8.4.2. One can check that the Casimir functions[7] $\|F\|^2$ and $\langle F, L\rangle$ Poisson-commute with all the generators L_i and F_i and that they label the coadjoint orbits of E(3) (Exercise 11.1.4).

Now, we turn to the discussion of integrability. Since the system is autonomous, the Hamiltonian

$$I_1 := H$$

is a constant of motion. Besides that, the projection

$$I_2 := \mathbf{L} \cdot \boldsymbol{\varepsilon}, \tag{11.1.15}$$

of angular momentum to the direction $\boldsymbol{\varepsilon} := \|\mathbf{F}\|^{-1}\mathbf{F}$ of the external field $\mathbf{F}$ in the body frame should be a constant of motion, because it corresponds to the symmetry of this model under rotations about the axis defined by $\mathbf{F}$. Indeed, since $\mathbf{N} \cdot \mathbf{F} = 0$, the angular momentum balance equation (11.1.12) yields

$$0 = \mathbf{N} \cdot \boldsymbol{\varepsilon} = \dot{\mathbf{L}} \cdot \boldsymbol{\varepsilon} + (\boldsymbol{\omega} \times \mathbf{L}) \cdot \boldsymbol{\varepsilon} = \dot{\mathbf{L}} \cdot \boldsymbol{\varepsilon} + \mathbf{L} \cdot \dot{\boldsymbol{\varepsilon}} = \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{L} \cdot \boldsymbol{\varepsilon}).$$

A third constant of motion which is in involution with I_2 has been found in the following special cases. For simplicity, we assume that the body frame is chosen in such a way that the tensor of inertia $\boldsymbol{\Theta}$ is diagonal,

$$\boldsymbol{\Theta} = \begin{bmatrix} \Theta_1 & 0 & 0 \\ 0 & \Theta_2 & 0 \\ 0 & 0 & \Theta_3 \end{bmatrix}.$$

This is only relevant in the cases (b) and (c).

[7] In Lie algebra theoretic terms, the quantities $\|F\|^2$ and $\langle F, L\rangle$ are the quadratic Casimir operators of the Lie algebra of $\mathrm{E}(3) = \mathrm{SO}(3) \ltimes \mathbb{R}^3$ in the representation defined by L_i, F_i.

(a) The Euler top: this is the situation discussed in Example 10.6.5. In this case, either $\mathbf{S} = 0$, that is, the fixed point coincides with the centre of mass, or $\mathbf{F} = 0$. By scalar multiplication of Eq. (10.6.36) with $\mathbf{L}$ we obtain the additional constant of motion

$$I_3 = \mathbf{L}^2.$$

According to (11.1.14), I_2 and I_3 are in involution. Note that in the case $\mathbf{F} = 0$, for $\boldsymbol{\varepsilon}$ one can choose an arbitrary unit vector in $\mathbb{R}^3$,

(b) The Lagrange top: here one assumes $\Theta_1 = \Theta_2$ and $\mathbf{S} = (0, 0, S_3)$. By the axial symmetry, the additional constant of motion in involution is given by the projection of angular momentum to the symmetry axis of the inertia tensor:

$$I_3 = \mathbf{L} \cdot \mathbf{e}_3.$$

(c) Kovalevskaya top: here one puts $\Theta_1 = \Theta_2 = 2\Theta_3$ and $\mathbf{S} = (S_1, S_2, 0)$. Without loss of generality, we may assume that the centre of mass vector is parallel to $\mathbf{e}_1$. Then, the additional constant of motion in involution is given by

$$I_3 = \left(L_1^2 - L_2^2 + 2S\Theta_1 F_1\right)^2 + 4(L_1 L_2 + S\Theta_1 F_2)^2,$$

where S is the distance from the fixed point to the centre of mass.

We leave it to the reader to check that in each case, I_3 is a constant of motion and that the differentials of the functions in involution are linearly independent on a dense subset (Exercise 11.1.5). Let us add that in a completely analogous way, one can discuss the top in an ideal fluid. Here, a number of integrable situations exists as well, see Sect. 2.2. in [237].

Remark 11.1.7 Since the model is symmetric under SO(2)-rotations about the axis given by the external field $\mathbf{F}$, the configuration space reduces to the two-sphere S^2 and since $I_2 = \mathbf{L} \cdot \boldsymbol{\varepsilon}$ is a constant of motion, the dynamics in phase space reduces to the 4-dimensional level sets

$$\Sigma_l = \left\{(\boldsymbol{\varepsilon}, \mathbf{L}) \in \mathbb{R}^6 : \|\boldsymbol{\varepsilon}\| = 1, \mathbf{L} \cdot \boldsymbol{\varepsilon} = l\right\}. \tag{11.1.16}$$

Via the mapping

$$\Phi : \Sigma_l \to \mathrm{TS}^2, \quad (\boldsymbol{\varepsilon}, \mathbf{L}) \mapsto (\boldsymbol{\varepsilon}, \mathbf{L} - l\boldsymbol{\varepsilon}),$$

each of the level sets Σ_l is diffeomorphic to TS^2. This way, these level sets become symplectic manifolds. As a consequence of the fact that the Casimir functions $\|F\|^2$ and $\langle L, F\rangle$ label the coadjoint orbits of the Euclidean group E(3), cf. Remark 11.1.6, Σ_l is symplectomorphic to such a coadjoint orbit. Moreover, one can show that Σ_l corresponds to the reduced phase space at momentum level l obtained by symplectic reduction of the Hamiltonian SO(2)-manifold which arises from that of the Euler top, discussed in Example 10.6.5, by restricting the group action to the subgroup SO(2) of rotations about the axis defined by $\mathbf{F}$.

Example 11.1.8 (Spherical pendulum) Recall from Example 10.6.7 that the spherical pendulum has the phase space

$$\mathrm{TS}^2 = \left\{(\mathbf{x}, \mathbf{y}) \in \mathrm{T}\mathbb{R}^3 : \|\mathbf{x}\| = 1, \mathbf{x} \cdot \mathbf{y} = 0\right\}$$

and that the Hamiltonian function is

$$H(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\|\mathbf{y}\|^2 + \mathbf{x} \cdot \mathbf{e}_3 .$$

A further constant of motion is given by the momentum mapping of the SO(2)-symmetry of this system:

$$J(\mathbf{x}, \mathbf{y}) = x_1 y_2 - x_2 y_1 .$$

The subset of TS^2 where $\mathrm{d}H$ and $\mathrm{d}J$ are linearly independent coincides with the complement of the set of critical points of the energy momentum mapping. In Example 10.8.10, we have seen that this set is dense. Thus, the spherical pendulum is integrable.

Example 11.1.9 (Toda Lattice) The non-periodic Toda lattice is the Hamiltonian system on $\mathrm{T}^*\mathbb{R}^n$ defined by

$$H = \frac{1}{2}\sum_{k=1}^{n} p_k^2 + \sum_{k=1}^{n-1} e^{2(q^k - q^{k+1})}, \tag{11.1.17}$$

where q^i, p_i denote the standard bundle coordinates. It describes a linear molecule consisting of n atoms with exponential nearest neighbour interaction. The Hamilton equations are

$$\begin{aligned} &\dot{q}^k = p_k, && k = \quad 1, \dots, n, \\ &\dot{p}_k = 2\mathrm{e}^{2(q^{k-1} - q^k)} - 2\mathrm{e}^{2(q^k - q^{k+1})}, && k = 2, \dots, n-1, \\ &\dot{p}_1 = -2\mathrm{e}^{2(q^1 - q^2)}, \qquad \dot{p}_n = 2\mathrm{e}^{2(q^{n-1} - q^n)}. && \end{aligned} \tag{11.1.18}$$

Since $\sum_{k=1}^{n} \dot{p}_k = 0$, we can separate out the motion of the centre of mass by passing to the rest frame. Then, $\sum_i p_i = 0$ and the relative motion can be described in terms of the new variables

$$a_k = \mathrm{e}^{q^k - q^{k+1}}, \qquad b_k = p_k$$

fulfilling $\sum_{k=1}^{n} b_k = 0$. Define

$$L = \begin{bmatrix} b_1 & a_1 & 0 & \cdots & 0 \\ a_1 & b_2 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & b_{n-1} & a_{n-1} \\ 0 & \cdots & 0 & a_{n-1} & b_n \end{bmatrix},$$

$$M = \begin{bmatrix} 0 & a_1 & 0 & \cdots & 0 \\ -a_1 & 0 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & a_{n-1} \\ 0 & \cdots & 0 & -a_{n-1} & 0 \end{bmatrix}.$$

We show that the Hamilton equations (11.1.18) are equivalent to the Lax equation

$$\dot{L} = [L, M]. \tag{11.1.19}$$

For that purpose, we calculate

$$[L, M] = \begin{bmatrix} -2a_1^2 & a_1(b_1 - b_2) & 0 & \cdots & 0 \\ a_1(b_1 - b_2) & 2a_1^2 - 2a_2^2 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 2a_{n-2}^2 - 2a_{n-1}^2 & a_{n-1}(b_{n-1} - b_n) \\ 0 & \cdots & 0 & a_{n-1}(b_{n-1} - b_n) & +2a_{n-1}^2 \end{bmatrix}$$

and read off that (11.1.19) corresponds to the following system of equations:

$$\begin{aligned} \dot{a}_k &= a_k(b_k - b_{k+1}), && k = 1, \ldots, n-1, \\ \dot{b}_k &= 2a_{k-1}^2 - 2a_k^2, && k = 2, \ldots, n-2, \\ \dot{b}_1 &= -2a_1^2, \qquad \dot{b}_n = 2a_{n-1}^2. \end{aligned} \tag{11.1.20}$$

By direct inspection, one can check that this system is equivalent to (11.1.18). Due to

$$\frac{\mathrm{d}}{\mathrm{d}t} \operatorname{tr}(L^k) = k \operatorname{tr}(\dot{L} L^{k-1}) = k \operatorname{tr}([L, M] L^{k-1}) = \operatorname{tr}([M, L^k]) = 0,$$

the quantities

$$I_k = \frac{1}{k} \operatorname{tr}(L^k) \tag{11.1.21}$$

are constants of motion for every k. Since the coefficients of the characteristic polynomial of L can be expressed in terms of the traces $\operatorname{tr}(L^k)$, the eigenvalues of L are constants of motion, too. In the next section, we will show that the I_k are in involution. Moreover, one can check that the differentials $\mathrm{d}I_2, \ldots, \mathrm{d}I_n$ are linearly independent on a dense subset of $\mathrm{T}^*\mathbb{R}^n$ (Exercise 11.1.6). We conclude that the Toda lattice is integrable. Finally, let us calculate I_1 and I_2:

$$I_1 = \sum_{k=1}^{n} b_k = \sum_{k=1}^{n} p_k = 0 \quad \text{(total momentum)},$$

$$I_2 = \frac{1}{2}\sum_{k=1}^{n} b_k^2 + \sum_{k=1}^{n-1} a_k^2 = H \quad \text{(energy)}.$$

Exercises

11.1.1 Show that the six constants of motion of the two-body problem given by (11.1.3) are in involution and that the differentials of these functions are linearly independent on a dense subset.

11.1.2 Prove the statements (a) and (b) in Example 11.1.4.

11.1.3 Prove Formulae (11.1.13) and (11.1.14).

11.1.4 Verify that the Casimir functions $\|F\|^2$ and $\langle F, S\rangle$ of Example 11.1.5 commute with all generators of the Euclidean group E(3) and prove that they label the coadjoint orbits of E(3).

11.1.5 Show that both for the Lagrange and for the Kovalevskaya top, the quantity I_3 is a constant of motion in involution with I_2 and check the linear independence of the differentials of H, I_2 and I_3.

11.1.6 Determine the subset of $\mathrm{T}^*\mathbb{R}^n$ where the differentials of the constants of motion $I_2, \ldots, I_n$ defined by (11.1.21) are linearly independent.

11.2 Lax Pairs and Coadjoint Orbits

As we have seen, sometimes the Hamilton equations can be written in the form of a Lax equation

$$\dot{L} = [L, M] \tag{11.2.1}$$

with L and M being matrices whose entries are functions of positions and momenta. As noted above, the quantities I_k defined by (11.1.21) are constants of motion, and so are the eigenvalues of L. One says that the motion is isospectral. Thus, if one can find a sufficiently large number of invariants of this type which are in involution and functionally independent, the system is integrable.

Remark 11.2.1 Assume that $t \mapsto L(t)$ is a solution of the Lax equation (11.2.1) with initial value $L(0)$. Then, there exists a smooth matrix-valued function $t \mapsto g(t)$ such that

$$L(t) = g(t)L(0)g(t)^{-1}, \qquad M(t) = -\dot{g}(t)g(t)^{-1}, \qquad g(0) = \mathbb{1}. \qquad (11.2.2)$$

Indeed, the function g is found by solving the initial value problem for the second equation.[8] Then, by direct inspection, one checks that the matrix-valued function $g(t)L(0)g(t)^{-1}$ solves the Lax equation for the initial value $L(0)$. Now, by uniqueness, the first equation in (11.2.2) follows. Thus, one can always seek the solution of the Lax equation by making the ansatz $L(t) = g(t)L(0)g(t)^{-1}$. This procedure is called the method of isospectral deformation. Finally, we note that (11.2.2) immediately implies

$$I_k = \frac{1}{k}\,\mathrm{tr}\big(L(0)^k\big).$$

This is an alternative argument showing that the I_k are constants of motion.

In this section, we discuss aspects of integrability for Hamiltonian systems on coadjoint orbits in the dual space $\mathfrak{g}^*$ of a Lie algebra $\mathfrak{g}$. Recall from Sect. 8.4 that the coadjoint orbits in $\mathfrak{g}^*$ are the symplectic leaves of the Lie-Poisson structure (8.2.18) on $\mathfrak{g}^*$, and that their symplectic form is given by the (positive) Kirillov form (8.4.2). In this context, there exists a variety of constructive methods [237]. Here, we discuss one of these methods and apply it to the study of the Toda lattice. The setup is as follows. Let $\tilde{G}$ be a Lie group with Lie algebra $\tilde{\mathfrak{g}}$, let $\mathfrak{g}$ be a Lie subalgebra and let G be the connected Lie subgroup of $\tilde{G}$ associated with $\mathfrak{g}$. Assume that there exists a Lie subalgebra $\mathfrak{k}$ of $\tilde{\mathfrak{g}}$ such that

$$\tilde{\mathfrak{g}} = \mathfrak{g} \oplus \mathfrak{k}$$

(direct sum of vector spaces). Then, the dual spaces $\mathfrak{g}^*$ and $\mathfrak{k}^*$ can be naturally identified with $\mathfrak{k}^0$ and $\mathfrak{g}^0$, respectively, and thus give rise to the induced decomposition

$$\tilde{\mathfrak{g}}^* = \mathfrak{g}^* \oplus \mathfrak{k}^*.$$

Denote the natural projections corresponding to these decompositions by $\pi_{\mathfrak{g}}$, $\pi_{\mathfrak{k}}$ and $\pi_{\mathfrak{g}^*}$, $\pi_{\mathfrak{k}^*}$. Below, for the adjoint and coadjoint representations of G and $\mathfrak{g}$, we will use the conventional notation Ad etc., whereas for the adjoint and coadjoint representations of $\tilde{G}$ and $\tilde{\mathfrak{g}}$, we will occasionally use the notation $\widetilde{\mathrm{Ad}}$ etc. Finally, recall from the discussion of Formula (8.2.18), that for a smooth function f on $\tilde{\mathfrak{g}}$, the exterior differential $\mathrm{d}f$ can be viewed as a smooth mapping $\mathrm{d}f : \tilde{\mathfrak{g}}^* \to \tilde{\mathfrak{g}}$.

The following result belongs to Kostant [171] and Symes [284].[9]

---

[8] Note that $M(t)$ depends on $L(t)$.

[9] The statement can be formulated entirely on the level of Lie algebras. To see this, reformulate the requirement of $\widetilde{\mathrm{Ad}}^*$-invariance in terms of $\widetilde{\mathrm{ad}}^*$, drop $\tilde{G}$ and replace G by the Lie subgroup of $\mathrm{GL}(\mathfrak{g})$ generated by the automorphisms $\exp(\mathrm{ad}(A))$, $A \in \mathfrak{g}$, defined by the exponential series.

Proposition 11.2.2 *Under the assumptions made, the following holds.*

1. *If f and g are $\widetilde{\mathrm{Ad}}^*$-invariant smooth functions on $\tilde{\mathfrak{g}}^*$, their restrictions to $\mathfrak{g}^*$ are in involution with respect to the natural Poisson structure on $\mathfrak{g}^*$.*
2. *For a Hamiltonian function H on $\mathfrak{g}^*$ which is obtained by restriction of an $\widetilde{\mathrm{Ad}}^*$-invariant function $\tilde{H}$ on $\tilde{\mathfrak{g}}^*$, the Hamilton equations have the form*

$$\dot{\nu} = \widetilde{\mathrm{ad}}^*\big(\pi_{\mathfrak{k}}\big(\mathrm{d}\tilde{H}(\nu)\big)\big)\nu. \tag{11.2.3}$$

According to Example 9.2.3, every coadjoint orbit of G is invariant under the flow defined by H and (11.2.3) is the Hamilton equation on this orbit.

Proof 1. Let $\tilde{f}, \tilde{g} \in C^\infty(\tilde{\mathfrak{g}}^*)$ and let $f := \tilde{f}_{\restriction \mathfrak{g}^*}$ and $g := \tilde{g}_{\restriction \mathfrak{g}^*}$ be the restrictions. By (8.2.18),

$$\{\tilde{f}, \tilde{g}\}(\nu) := \big\langle \nu, [\mathrm{d}\tilde{f}, \mathrm{d}\tilde{g}]\big\rangle,$$

where $\mathrm{d}\tilde{f}$ and $\mathrm{d}\tilde{g}$ are viewed as mappings $\tilde{\mathfrak{g}}^* \to \tilde{\mathfrak{g}}$. Writing

$$\mathrm{d}_1\tilde{f} = \pi_{\mathfrak{g}} \circ \mathrm{d}\tilde{f} : \tilde{\mathfrak{g}}^* \to \mathfrak{g}, \qquad \mathrm{d}_2\tilde{f} = \pi_{\mathfrak{k}} \circ \mathrm{d}\tilde{f} : \tilde{\mathfrak{g}} \to \mathfrak{k}$$

and using that $\widetilde{\mathrm{Ad}}^*$-invariance implies

$$\big\langle \nu, [\mathrm{d}\tilde{f}, X]\big\rangle = \big\langle \nu, [\mathrm{d}\tilde{g}, X]\big\rangle = 0 \quad \text{for all } X \in \tilde{\mathfrak{g}}, \tag{11.2.4}$$

we find

$$\{f, g\}_{\mathfrak{g}^*}(\nu) = \big\langle \nu, \big[\mathrm{d}f(\nu), \mathrm{d}g(\nu)\big]\big\rangle = \big\langle \nu, \big[\mathrm{d}_1\tilde{f}(\nu), \mathrm{d}_1\tilde{g}(\nu)\big]\big\rangle = \big\langle \nu, \big[\mathrm{d}_2\tilde{f}(\nu), \mathrm{d}_2\tilde{g}(\nu)\big]\big\rangle.$$

Since $\mathfrak{k}$ is a Lie subalgebra, $[\mathrm{d}_2 f(\nu), \mathrm{d}_2 g(\nu)] \in \mathfrak{k}$. Since $\nu \in \mathfrak{g}^* = \mathfrak{k}^0$, the right hand side vanishes.

2. A smooth curve $t \mapsto \nu(t)$ in $\mathfrak{g}^*$ is an integral curve of the Hamiltonian vector field X_H generated by H iff

$$\big\langle X_H\big(\nu(t)\big), \mathrm{d}f\big(\nu(t)\big)\big\rangle = \big\langle \dot{\nu}(t), \mathrm{d}f\big(\nu(t)\big)\big\rangle$$

for all $f \in C^\infty(\mathfrak{g}^*)$ and all t. While, originally, the pairing is that of tangent vectors with covectors, by the usual identifications, we can interpret it as the pairing of elements of $\mathfrak{g}^*$ with elements of $\mathfrak{g}$ and rewrite the left hand side as follows, denoting $\mu = \nu(t)$:

$$\big\langle X_H(\mu), \mathrm{d}f(\mu)\big\rangle = \{H, f\}(\mu) = \big\langle \mu, \big[\mathrm{d}H(\mu), \mathrm{d}f(\mu)\big]\big\rangle = \big\langle \mu, \big[\mathrm{d}_1\tilde{H}(\mu), \mathrm{d}f(\mu)\big]\big\rangle.$$

In the last expression, $\mathrm{d}f(\mu)$ is viewed as an element of $\tilde{\mathfrak{g}}$ and the bracket is that of $\tilde{\mathfrak{g}}$. Using (11.2.4), we can rewrite the last expression as

$$-\big\langle \mu, \big[\mathrm{d}_2\tilde{H}(\mu), \mathrm{d}f(\mu)\big]\big\rangle = \big\langle \widetilde{\mathrm{ad}}^*\big(\mathrm{d}_2\tilde{H}(\mu)\big)\mu, \mathrm{d}f(\mu)\big\rangle.$$

This yields the assertion. □

According to Remark 5.4.11/2, in case $\tilde{\mathfrak{g}}$ admits a non-degenerate symmetric Ad-invariant bilinear form k, one has a natural isomorphism $F : \tilde{\mathfrak{g}} \to \tilde{\mathfrak{g}}^*$ of representations. Then, F^{-1} maps $\mathfrak{g}^*$ and $\mathfrak{k}^*$ to their mutual k-orthogonal complements in $\tilde{\mathfrak{g}}$ and the Hamilton equation (11.2.3) takes the form of the Lax equation (11.2.1) with

$$L = \nu, \qquad M = -\pi_{\mathfrak{k}}\big(\mathrm{d}\tilde{H}(\nu)\big). \tag{11.2.5}$$

Example 11.2.3 (Toda Lattice) Using Proposition 11.2.2, we analyze the Toda lattice. On the way, we use a number of facts about the Lie algebra $\mathfrak{sl}(n,\mathbb{R})$ of traceless real $(n \times n)$-matrices. We leave it to the reader to check them (Exercise 11.2.1).

Let $\tilde{G} = \mathrm{SL}(n,\mathbb{R})$, the group of real $(n \times n)$-matrices of unit determinant, let $\mathfrak{g} \subset \mathfrak{sl}(n,\mathbb{R})$ be the Lie subalgebra of traceless upper triangular $(n \times n)$-matrices, and let $\mathfrak{k} = \mathfrak{so}(n) \subset \mathfrak{sl}(n,\mathbb{R})$ be the Lie subalgebra of real skew-symmetric $(n \times n)$-matrices. As a vector space,

$$\tilde{\mathfrak{g}} = \mathfrak{g} \oplus \mathfrak{k}.$$

To write down the projections $\pi_{\mathfrak{g}}$ and $\pi_{\mathfrak{k}}$, observe that every $A \in \tilde{\mathfrak{g}}$ can be decomposed uniquely as $A = A_l + A_d + A_u$, where A_d is diagonal and A_l and A_u are lower and upper triangular with zero diagonal, respectively. We have

$$\pi_{\mathfrak{g}}(A) = A_d + A_u + A_l^{\mathrm{T}}, \qquad \pi_{\mathfrak{k}}(A) = A_l - A_l^{\mathrm{T}}. \tag{11.2.6}$$

The trace form $(A,B) \mapsto \mathsf{k}(A,B) := \mathrm{tr}(AB)$ on $\tilde{\mathfrak{g}}$ is non-degenerate, symmetric and Ad-invariant. The corresponding isomorphism of representations $F : \tilde{\mathfrak{g}} \to \tilde{\mathfrak{g}}^*$ identifies $\tilde{\mathfrak{g}}^*$ with $\tilde{\mathfrak{g}}$, $\mathfrak{g}^*$ with the subspace $\mathrm{S}_n(\mathbb{R})$ of $\tilde{\mathfrak{g}}$ of real symmetric matrices and $\mathfrak{k}^*$ with the subspace of upper triangular matrices with zero diagonal. The corresponding projections $\pi_{\mathfrak{g}^*}$ and $\pi_{\mathfrak{k}^*}$ are given by

$$\pi_{\mathfrak{g}^*}(\nu) = \nu_d + \nu_l + \nu_l^{\mathrm{T}}, \qquad \pi_{\mathfrak{k}^*}(\nu) = \nu_u - \nu_l^{\mathrm{T}}. \tag{11.2.7}$$

The Lie subgroup G of $\tilde{G}$ generated by $\mathfrak{g}$ is the identity connected component of the subgroup of upper triangular matrices with unit determinant. Since the diagonal entries of the elements of this subgroup must be nonzero, G consists of those elements whose diagonal entries are positive. Under the identifications made, the coadjoint action of G on $\mathfrak{g}^* = \mathrm{S}_n(\mathbb{R})$ is given by

$$\mathrm{Ad}^*(g)\nu = \pi_{\mathfrak{g}^*}\big(g\nu g^{-1}\big), \quad g \in G, \quad \nu \in \mathfrak{g}^*. \tag{11.2.8}$$

As the Hamiltonian H on $\mathfrak{g}^* = \mathrm{S}_n(\mathbb{R})$ we take

$$H = \frac{1}{2}\,\mathrm{tr}\big(\nu^2\big).$$

By letting ν range through the whole of $\tilde{\mathfrak{g}}$, we can extend H to a function $\tilde{H}$ on $\tilde{\mathfrak{g}}$. A brief computation shows $\mathrm{d}\tilde{H}(\nu) = \nu$, where ν is viewed as an element of $\tilde{\mathfrak{g}}$.

Hence, (11.2.5) and (11.2.6) yield that the Hamilton equation is a Lax equation with $L = \nu$ and

$$M = \nu_l^{\mathrm{T}} - \nu_l. \tag{11.2.9}$$

Now let us restrict this equation to the coadjoint orbit of

$$\nu_0 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 1 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \in \mathfrak{g}^*.$$

Using (11.2.7) and (11.2.8), we find that points on this orbit have the form

$$\nu = \pi_{\mathfrak{g}^*}\big(\mathrm{Ad}^*(g)\nu_0\big) = \begin{bmatrix} b_1 & a_1 & 0 & \cdots & 0 \\ a_1 & b_2 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & b_{n-1} & a_{n-1} \\ 0 & \cdots & 0 & a_{n-1} & b_n \end{bmatrix}, \quad g \in G, \tag{11.2.10}$$

where

$$a_k = \frac{g_{k+1k+1}}{g_{kk}}, \quad k = 1, \ldots, n-1,$$

and

$$b_1 = \frac{g_{12}}{g_{11}}, \qquad b_k = \frac{g_{kk+1}}{g_{kk}} - \frac{g_{k-1k}}{g_{k-1k-1}}, \qquad b_n = -\frac{g_{n-1n}}{g_{n-1n-1}},$$

with $k = 2, \ldots, n-1$. The real numbers a_k and b_k can take arbitrary values such that $a_i > 0$ and $\sum_{k=1}^n b_k = 0$. In particular, the orbit has dimension $2(n-1)$. Thus, according to (11.2.9), the Lax pair (L, M) representing the Hamilton equation on the coadjoint orbit of ν_0 is given by

$$L = \begin{bmatrix} b_1 & a_1 & 0 & \cdots & 0 \\ a_1 & b_2 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & b_{n-1} & a_{n-1} \\ 0 & \cdots & 0 & a_{n-1} & b_n \end{bmatrix},$$

$$M = \begin{bmatrix} 0 & a_1 & 0 & \cdots & 0 \\ -a_1 & 0 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & a_{n-1} \\ 0 & \cdots & 0 & -a_{n-1} & 0 \end{bmatrix}.$$

We conclude that the Lax pair constructed this way coincides with the Lax pair obtained in Example 11.1.9. Hence, the substitutions

$$a_k = \mathrm{e}^{(q^k - q^{k+1})}, \qquad b_k = p_k$$

yield the non-periodic Toda lattice model with the motion of the centre of mass separated out. Now, Proposition 11.2.2/1 immediately implies that the invariants I_k given by (11.1.21) are in involution. This proves that the non-periodic Toda lattice is integrable.

Remark 11.2.4

1. We have shown that the phase space of the Toda lattice is symplectomorphic to a certain coadjoint orbit of the group of upper triangular matrices with unit determinant. This observation goes back to Kostant [171] and Adler [5]. The dual space of the Lie algebra of this group can be realized in different ways. Above we used the so-called symmetric Lax representation. Another representation is provided by the subspace of lower triangular matrices with zero trace. This yields the so-called non-symmetric Lax representation. For this and a number of generalizations of the above construction to the case of coadjoint orbits of parabolic subgroups of simple Lie groups which yield generalized Toda systems, we refer to [237].
2. The periodic Toda lattice is defined by the Hamiltonian

$$H = \frac{1}{2}\sum_{k=1}^{n} p_k^2 + \sum_{k=1}^{n} e^{2(q^k - q^{k+1})} \quad \text{with } q_{n+1} = q_1.$$

   This system is integrable, too. For both the periodic and the non-periodic Toda lattice, the integration of the equations of motion can be carried out explicitly. It turns out that the dynamics of the non-periodic model is asymptotically free. Using a method proposed by Moser [223], the solution can be constructed recursively, see [140] and [237] for details. The solution for the periodic model is much more complicated, see [155] and [173].

While solving the equations of motion of an integrable system explicitly can be quite involved, these systems have an interesting geometric structure. A careful analysis of this structure shows that there exists a special sort of adapted local Darboux coordinates, called action and angle variables. Once such coordinates have been found, the integration of the equations of motion becomes trivial. This will be the topic of Sect. 11.4.

Exercises

11.2.1 Complete the discussion of the Toda model in Example 11.2.3 by showing the following.
 (a) As a vector space, $\mathfrak{sl}(n, \mathbb{R})$ is the direct sum of the subspace of upper triangular matrices with zero trace and the subspace $\mathfrak{so}(n)$ of skew-symmetric matrices.
 (b) The trace form $\mathsf{k}(A, B) := \mathrm{tr}(AB)$ is a non-degenerate Ad-invariant symmetric bilinear form on $\mathfrak{sl}(n, \mathbb{R})$.
 (c) The k-orthogonal complement of $\mathfrak{so}(n)$ in $\mathfrak{sl}(n, \mathbb{R})$ is given by the subspace $\mathrm{S}_n(\mathbb{R})$ of symmetric matrices.
 (d) Points on the coadjoint orbit of ν_0 are given by (11.2.10).

11.2.2 In the notation of Example 11.2.3, show that in the coordinates a_k, b_k the Poisson structure on the coadjoint orbit of ν_0 is given by

$$\{a_i, b_i\} = -a_i, \qquad \{a_i, b_{i+1}\} = a_i,$$

whereas all other Poisson brackets vanish. Using these relations, confirm that in these coordinates the Hamilton equations read

$$\dot{a}_i = \{H, a_i\}, \qquad \dot{b}_i = \{H, b_i\}.$$

11.3 The Arnold Theorem

In this section, we start to discuss the foliation of an integrable system $(M, \omega, \mathscr{H})$ defined by the level set components of the mapping $\mathscr{H}$. We confine our attention to the restriction $\mathscr{H}_r$ of $\mathscr{H}$ to the subset $M^{\mathscr{H}}$ of regular points of $\mathscr{H}$, cf. (11.1.1). Recall from Remark 11.1.2 that the level set components of $\mathscr{H}_r$ coincide with the integral manifolds of the regular distribution $D^{\mathscr{H}_r}$ on $M^{\mathscr{H}}$ spanned by the Hamiltonian vector fields X_{H_i}.

Belonging to the same level set component of $\mathscr{H}_r$ is an equivalence relation in $M^{\mathscr{H}}$. Let $\widetilde{V}^{\mathscr{H}}$ denote the set of equivalence classes[10] (the space of leaves of $D^{\mathscr{H}_r}$), endowed with the quotient topology, and let

$$\widetilde{\mathscr{H}_r} : M^{\mathscr{H}} \to \widetilde{V}^{\mathscr{H}} \tag{11.3.1}$$

be the natural projection, assigning to $m \in M^{\mathscr{H}}$ the level set component of $\mathscr{H}_r$ containing m. Since a submersion is open, the values of $\mathscr{H}_r$ form an open subset $V^{\mathscr{H}}$ of $\mathbb{R}^n$ and the mapping $\mathscr{H}_r : M^{\mathscr{H}} \to V^{\mathscr{H}}$ decomposes into $\mathscr{H}_r = \hat{\mathscr{H}_r} \circ \widetilde{\mathscr{H}_r}$, where

$$\hat{\mathscr{H}_r} : \widetilde{V}^{\mathscr{H}} \to V^{\mathscr{H}} \tag{11.3.2}$$

[10]Sometimes referred to as the Reeb graph of the mapping $\mathscr{H}_r$.

assigns to a level set component of $\mathscr{H}_r$ the corresponding value. In this section we prove the Arnold Theorem, which states that each level set component of $\mathscr{H}_r$ on which the Hamiltonian vector fields X_{H_i} are complete is diffeomorphic to $\mathrm{T}^k \times \mathbb{R}^{n-k}$ for some $0 \leq k \leq n$. Here, T^k denotes the k-dimensional torus. In the next section, we will show that the restriction of the foliation (11.3.1) to the subset $\widetilde{V}_c^{\mathscr{H}} \subset \widetilde{V}^{\mathscr{H}}$ of compact level set components is a locally trivial fibre bundle with typical fibre T^n. This will be a consequence of the existence of certain adapted coordinates, called action and angle variables. In Sect. 11.7, we will discuss topological aspects of this bundle.

Let us start with the following classical result.

Theorem 11.3.1 (Liouville) *Let $(M, \omega, H_1, \dots, H_n)$ be an integrable system and let $m \in M$ be a regular point of $\mathscr{H} = (H_1, \dots, H_n)$. There exists an open neighbourhood U of m and smooth functions $G^1, \dots, G^n$ on U complementing $H_1, \dots, H_n$ to Darboux coordinates. In these coordinates, the flow Φ^i of the Hamiltonian vector field X_{H_i} is given by*

$$\Phi_t^i(G, H) = \big(G^1, \dots, G^i + t, \dots, G^n, H_1, \dots, H_n\big). \tag{11.3.3}$$

Proof Since the subset of regular points $M^{\mathscr{H}}$ is open, we may search for U inside $M^{\mathscr{H}}$. Since $D^{\mathscr{H}_r}$ is a regular distribution of rank n and since the X_{H_i} commute, Remark 3.5.11/2 yields a local chart adapted to $D^{\mathscr{H}_r}$ on some open neighbourhood U of m, whose first n coordinate functions $G^1, \dots, G^n$ satisfy $X_{H_i} = \partial_{G^i}$. Obviously, we can replace the last n coordinate functions of this chart by $H_1, \dots, H_n$. By construction, the local representative of the flow Φ^i is given by (11.3.3) and the Poisson brackets are

$$\{H_i, H_j\} = 0, \qquad \big\{H_i, G^j\big\} = X_{H_i}\big(G^j\big) = \delta_i^j.$$

A brief computation shows that the latter implies

$$\omega = \mathrm{d}H_i \wedge \mathrm{d}G^i + h^{ij}\mathrm{d}H_i \wedge \mathrm{d}H_j$$

with smooth functions h^{ij}, uniquely determined by the condition $h^{ij} = -h^{ji}$. By (11.3.3), the 2-forms $\mathrm{d}H_i \wedge \mathrm{d}G^i$ and $\mathrm{d}H_i \wedge \mathrm{d}H_j$ are invariant under each of the flows Φ^l. Since ω is invariant under these flows, too, each of the functions h^{ij} must be invariant. Hence, the second term can be written as

$$h^{ij}\mathrm{d}H_i \wedge \mathrm{d}H_j = \mathscr{H}^*\big(k_{ij}\mathrm{d}x^i \wedge \mathrm{d}x^j\big)$$

where x^i are the standard coordinates on $\mathbb{R}^n$ and k_{ij} are smooth functions on the open subset $\mathscr{H}(U) \subset \mathbb{R}^n$. Since $\mathrm{d}(h^{ij}\mathrm{d}H_i \wedge \mathrm{d}H_j) = \mathrm{d}(\omega - \mathrm{d}H_i \wedge \mathrm{d}G^i) = 0$, we have

$$\mathscr{H}^*\mathrm{d}\big(k_{ij}\mathrm{d}x^i \wedge \mathrm{d}x^j\big) = 0.$$

Since $\mathscr{H}$ is a submersion, this implies that $k_{ij}\mathrm{d}x^i \wedge \mathrm{d}x^j$ is closed. Hence, by the Poincaré Lemma,[11] there exists a 1-form $\alpha_i \mathrm{d}x^i$ on $\mathscr{H}(U)$ such that

$$k_{ij}\mathrm{d}x^i \wedge \mathrm{d}x^j = \mathrm{d}\big(\alpha_i \mathrm{d}x^i\big).$$

We leave it to the reader to check that after replacing the coordinate functions G^i by $G^i - \alpha^i \circ \mathscr{H}$ we have $\omega = \mathrm{d}H_i \wedge \mathrm{d}G^i$, whereas (11.3.3) remains unchanged. □

Remark 11.3.2

1. The functions G^i provide coordinates on the level set components of $\mathscr{H}_r$. Since G^i, H_i are Darboux coordinates, their Poisson brackets are

   $$\{H_i, H_j\} = \big\{G^i, G^j\big\} = 0, \qquad \big\{H_i, G^j\big\} = \delta_i{}^j.$$

   Since the Hamiltonian is given by the coordinate function H_1, the Hamilton equations take the following simple form:

   $$\dot{H}_i = 0, \quad i = 1, \dots, n, \qquad \dot{G}^1 = 1, \qquad \dot{G}^i = 0, \quad i = 2, \dots, n.$$

   Hence, the integral curves are given by (11.3.3) with $i = 1$.
2. Since in the proof of Theorem 11.3.1 we can replace α by $\alpha + d\lambda$, with λ being an arbitrary smooth function on $\mathbb{R}^n$, there is the following freedom in the choice of the functions G^i:

   $$G^i \mapsto G^i + \frac{\partial \lambda(H)}{\partial H_i}.$$

   Now we discuss the structure of the level set components of $\mathscr{H}_r$.

Theorem 11.3.3 (Arnold) *Let $(M, \omega, \mathscr{H})$ be an integrable system and let Σ be a level set component of $\mathscr{H}_r$. If Σ is compact, it is diffeomorphic to T^n. If Σ is not compact but the restrictions of the Hamiltonian vector fields X_{H_i} to Σ are complete, it is diffeomorphic to $\mathrm{T}^k \times \mathbb{R}^{n-k}$ for some $0 \leq k < n$.*

Proof It suffices to prove that if the restrictions of the Hamiltonian vector fields X_{H_i} to Σ are complete, the latter is diffeomorphic to $\mathrm{T}^k \times \mathbb{R}^{n-k}$ for some $0 \leq k \leq n$. By assumption, the flows Φ^i of the restrictions of the vector fields X_{H_i} to Σ are complete and commute with one another. Thus, they define an action of $\mathbb{R}^n$ on Σ by

$$\Psi : \mathbb{R}^n \times \Sigma \to \Sigma, \qquad \Psi(\mathbf{t}, m) := \big(\Phi^1_{t_1} \circ \cdots \circ \Phi^n_{t_n}\big)(m), \tag{11.3.4}$$

cf. Example 6.1.2/4. We show that this action is transitive. First, we observe that Ψ'_m is invertible for every $m \in \Sigma$, because the vector fields X_{H_i} are linearly independent

[11] We can assume $\mathscr{H}(U)$ to be contractible.

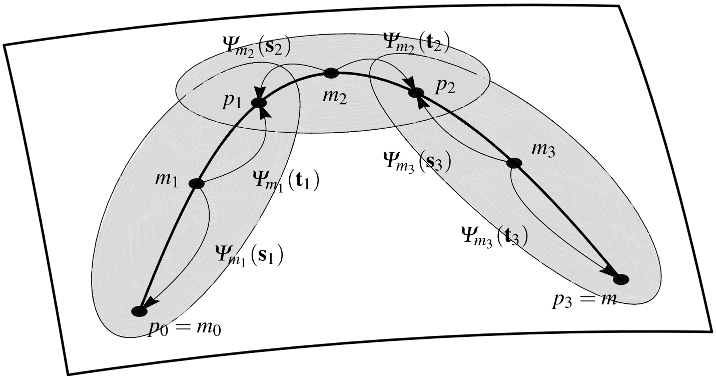


Fig. 11.1 Transitivity of the action Ψ in the proof of the Arnold Theorem

on Σ. Hence, the orbit mapping $\Psi_m : \mathbb{R}^n \to \Sigma$ defined by m restricts to a diffeomorphism from some open neighbourhood V_m of the origin of $\mathbb{R}^n$ onto an open neighbourhood U_m of m in Σ. Now, let m_0 and m be arbitrary points in Σ. Since Σ is connected, we find a curve $\gamma : [0, 1] \to \Sigma$ from m_0 to m. Since $[0, 1]$ is compact, we find points $m_1, \ldots, m_r$ on γ such that the subsets $U_{m_1}, \ldots, U_{m_r}$ form a covering of γ. By an appropriate reordering we may assume that $U_{m_i} \cap U_{m_{i+1}}$ is nonempty and hence contains a point p_i for all $i = 1, \ldots, r-1$. Put $p_0 \equiv m_0$ and $p_r \equiv m$. For $i = 1, \ldots, r$, there exist unique $\mathbf{s}_i, \mathbf{t}_i \in V_{m_i}$ such that $\Psi_{m_i}(\mathbf{s}_i) = p_{i-1}$ and $\Psi_{m_i}(\mathbf{t}_i) = p_i$, see Fig. 11.1. Then, $\Psi_{\mathbf{t}_i - \mathbf{s}_i}(p_{i-1}) = \Psi_{\mathbf{t}_i - \mathbf{s}_i} \circ \Psi_{\mathbf{s}_i}(m_i) = \Psi_{\mathbf{t}_i}(m_i) = p_i$ and hence

$$\Psi_{\mathbf{t}_1 - \mathbf{s}_1 + \cdots + \mathbf{t}_r - \mathbf{s}_r}(m_0) = m.$$

Thus, the action Ψ is transitive, indeed. Now, the Orbit Theorem 6.2.8 yields that Σ is diffeomorphic to the homogeneous space given by the quotient of $\mathbb{R}^n$ by the common[12] stabilizer of the points of Σ. Since $(\Psi_m)_{\restriction V_m}$ is injective, the stabilizer is a discrete subgroup of $\mathbb{R}^n$ and hence isomorphic to the integer lattice generated by k linearly independent elements of $\mathbb{R}^n$, where $0 \le k \le n$. By complementing the generators to a basis of $\mathbb{R}^n$, we finally obtain $\Sigma \cong \mathrm{T}^k \times \mathbb{R}^{n-k}$. □

Motivated by the proof of the Arnold Theorem, let us discuss the special case where the Hamiltonian vector fields X_{H_i} are complete on the whole of M. In this case, there is a deep relation between integrability and symplectic reduction. Since the vector fields X_{H_i} and hence their flows Φ^i commute, the mapping

$$\Psi : \mathbb{R}^n \times M \to M, \qquad \Psi(\mathbf{t}, m) := \left(\Phi^1_{t_1} \circ \cdots \circ \Phi^n_{t_n}\right)(m) \tag{11.3.5}$$

[12] Since $\mathbb{R}^n$ is Abelian, all stabilizers along an orbit coincide.

is an action of the additive Lie group $\mathbb{R}^n$ on M. In the local coordinates G^i, H_i provided by Theorem 11.3.1, this action reads

$$\Psi(\mathbf{t}, m) = \big(G^1(m) + t_1, \ldots, G^n(m) + t_n, H_1(m), \ldots, H_n(m)\big). \tag{11.3.6}$$

Since each of the flows Φ^i leaves ω invariant, this action is symplectic. We show that $\mathscr{H}$ can be viewed as a momentum mapping. For that purpose, recall from Example 5.2.10 that the Lie algebra of $G = \mathbb{R}^n$ is given by the vector space $\mathfrak{g} = \mathbb{R}^n$ endowed with the trivial Lie bracket. Thus, if we identify $\mathfrak{g}^*$ with $\mathbb{R}^n$ via the canonical scalar product, we can view $\mathscr{H}$ as a mapping

$$J : M \to \mathfrak{g}^*, \qquad J(m) := \mathscr{H}(m).$$

Next, we calculate the Killing vector field of $\mathbf{x} \in \mathfrak{g}$. On the one hand, by Example 5.3.15, we have $\exp(s\mathbf{x}) = s\mathbf{x}$ for all $s \in \mathbb{R}$ and hence

$$\mathbf{x}_*(m) = \frac{\mathrm{d}}{\mathrm{d}s}\restriction_0 \Psi_{s\mathbf{x}}(m) = \sum_{i=1}^{n} x_i X_{H_i}(m) = X_{\sum_{i=1}^n x_i H_i}(m).$$

On the other hand,

$$J_{\mathbf{x}}(m) = \big\langle J(m), \mathbf{x}\big\rangle = \sum\nolimits_{i=1}^{n} x_i H_i(m).$$

Thus, $\mathbf{x}_* = X_{J_{\mathbf{x}}}$, and $J = \mathscr{H}$ is a momentum mapping for Ψ, indeed. It is trivially equivariant, because Proposition 9.1.10 implies that J is constant on the orbits of Ψ. Let us add that the fact that $\mathscr{H}$ is a momentum mapping for Ψ implies, in particular, that the distribution $D^{\mathfrak{g}}$ spanned by the Killing vector fields of Ψ coincides with the distribution $D^{\mathscr{H}}$ generated by Hamiltonian vector fields X_{H_i}. Hence, the Orbit Theorem 6.2.8 implies that the orbits of Ψ in $M^{\mathscr{H}}$ coincide with the level set components of $\mathscr{H}$ in $M^{\mathscr{H}}$. This yields an alternative proof of the Arnold Theorem in this special situation.

Now, consider the invariant open submanifold $M^{\mathscr{H}}$ of the regular points of $\mathscr{H}$. Since J is a submersion on $M^{\mathscr{H}}$, every value is regular. Since the coadjoint action is trivial, the stabilizers $G_{\mathbf{h}}$ of the values $\mathbf{h} \in \mathfrak{g}^* \equiv \mathbb{R}^n$ of J coincide with $G = \mathbb{R}^n$. While we cannot directly apply the theory of regular symplectic reduction as discussed in Sect. 10.3, because the action Ψ need not be proper, we can nevertheless form the topological quotient $J^{-1}(\mathbf{h})/G_{\mathbf{h}}$. Since the $G_{\mathbf{h}}$-orbits in $J^{-1}(\mathbf{h})$ coincide with the connected components of $J^{-1}(\mathbf{h})$, this quotient is a discrete space consisting of at most countably many isolated points. Hence, it is trivially a symplectic manifold and can be interpreted as the reduced phase space at momentum $\mathbf{h}$. Let us summarize.

Proposition 11.3.4 *Let $(M, \omega, H_1, \ldots, H_n)$ be an integrable system. If the Hamiltonian vector fields X_{H_i} are complete, the $\mathbb{R}^n$-action (11.3.5) endows (M, ω) with the structure of a Hamiltonian G-manifold with equivariant momentum mapping given by $\mathscr{H} = (H_1, \ldots, H_n)$. For the invariant open submanifold $M^{\mathscr{H}}$, the reduced phase spaces are discrete and at most countable.*

Remark 11.3.5 Let Σ be a level set component of $\mathscr{H}_r$ and assume that the restriction of the Hamiltonian vector fields X_{H_i} to Σ is complete. Then, the action (11.3.4) of $\mathbb{R}^n$ on Σ endows the latter with an affine connection,[13] that is, with a globally defined prescription for the parallel transport of tangent vectors from a point m_0 to another point m_1 along a curve connecting them. This affine connection turns out to be flat, meaning that the result of the parallel transport does not depend on the connecting curve. By parallely transporting every tangent vector to the tangent space $\mathrm{T}_{m_0}\Sigma$ at some chosen point m_0, one can represent vector fields on Σ by smooth mappings $\Sigma \to \mathrm{T}_{m_0}\Sigma$. Vector fields represented by constant mappings are called constant vector fields. They form a vector space which is isomorphic to $\mathrm{T}_{m_0}\Sigma$. By virtue of the basis $\{X_{H_1}(m_0), \ldots, X_{H_n}(m_0)\}$ in $\mathrm{T}_{m_0}\Sigma$, we can identify this vector space with $\mathbb{R}^n$ in a natural way. As noted in the proof of the Arnold Theorem, all the stabilizers of the action of $\mathbb{R}^n$ on Σ coincide with a certain integer lattice generated by k linearly independent elements of $\mathbb{R}^n$. We refer to this lattice as the period lattice of the mapping $\mathscr{H}$ at the value $\mathbf{h} \in \mathbb{R}^n$ and denote it by $P_{\mathscr{H}}(\Sigma)$. Under the above identification, $P_{\mathscr{H}}(\Sigma)$ corresponds to a subgroup of the additive group of constant vector fields.

11.4 Action and Angle Variables

By the Arnold Theorem, compact level set components of $\mathscr{H}_r$ are tori. There arises the question whether it is possible to pass to new constants of motion in involution which are adapted to these tori in the sense that their flow is periodic. This leads to the notion of action and angle variables.

Definition 11.4.1 (Action and angle variables) A system $\{I_1, \ldots, I_n\}$ of smooth functions on an open subset W of $M^{\mathscr{H}}$ is called a system of action variables if

1. the Hamiltonian vector fields X_{I_i} span the distribution $D^{\mathscr{H}}$ over W,
2. $\{I_i, I_j\} = 0$ for all i, j,
3. the flows of the Hamiltonian vector fields X_{I_i} are complete and 2π-periodic.

Any system of functions $\vartheta^1, \ldots, \vartheta^n$ on W such that (ϑ, I) yield Darboux coordinates on a neighbourhood of every point of W is called a system of angle variables.

Note that, up to the choice of the Lagrangian submanifold $\vartheta^i = 0$ transversal to the foliation, the functions ϑ^i are given by the flow parameters of the vector fields X_{I_i}. Hence, they are necessarily multi-valued mod 2π.

The historical origin of action and angle variables can be traced back to Jacobi and Liouville. The first modern formulation belongs to Arnold and Avez [23]. Before proving existence, let us collect the basic properties of action and angle variables, following immediately from the definition. Obviously, both the X_{I_i} and the

[13] To be defined in Sect. 11.7.

X_{H_i} form a local frame in $D^{\mathscr{H}}$ over W. These two frames are related by

$$X_{I_i} = b_i{}^j X_{H_j} \tag{11.4.1}$$

with a unique smooth mapping $b : W \to \mathrm{GL}(n, \mathbb{R})$. By taking rows of b we obtain mappings

$$b_i : W \to \mathbb{R}^n, \qquad b_i(m) := \big(b_i{}^1(m), \ldots, b_i^n(m)\big). \tag{11.4.2}$$

Proposition 11.4.2 *Let I be a system of action variables on $W \subset M^{\mathscr{H}}$.*

1. *One has $\{I_i, H_j\} = 0$ for all i, j and $\{H_i, b_j{}^k\} = 0$ for all i, j, k. In particular, the functions I_i and $b_i{}^j$ are constant on the level set components of $\mathscr{H}_r$.*
2. *With Φ^{H_i} and Φ^{I_i} denoting the flows of X_{H_i} and X_{I_i}, respectively, for $m \in W$ we have*

$$\Phi_s^{I_i}(m) = \Phi^{H_1}_{sb_i{}^1(m)} \circ \cdots \circ \Phi^{H_n}_{sb_i{}^n(m)}(m), \tag{11.4.3}$$

$$\Phi_s^{H_i}(m) = \Phi^{I_1}_{s(b^{-1})_i{}^1(m)} \circ \cdots \circ \Phi^{I_n}_{s(b^{-1})_i{}^n(m)}. \tag{11.4.4}$$

This implies the following.
(a) *The vector fields X_{H_i} are complete on W.*
(b) *The $\mathbb{R}^n$-actions Ψ^I and Ψ defined by the flows $\Phi^{I_1}, \ldots, \Phi^{I_n}$ and $\Phi^{H_1}, \ldots, \Phi^{H_n}$ on W, respectively, are related by*

$$\Psi^I_{\mathbf{t}}(m) = \Psi_{\sum_{i=1}^n t_i b_i(m)}(m). \tag{11.4.5}$$

In particular, they have the same orbits and these orbits coincide with compact level set components of $\mathscr{H}_r$.
(c) *For every $m \in W$, the vectors $2\pi b_1(m), \ldots, 2\pi b_n(m)$ form a set of generators for the stabilizer of m under Ψ.*

3. *If $\tilde{I}_1, \ldots, \tilde{I}_n$ is another system of action variables on $\tilde{W} \subset M^{\mathscr{H}}$, then, on $W \cap \tilde{W}$,*

$$I_i = A_i{}^j \tilde{I}_j + C_i, \qquad b_i = A_i{}^j \tilde{b}_j$$

with smooth mappings $A : W \cap \tilde{W} \to \mathrm{GL}(n, \mathbb{Z})$ and $C : W \cap \tilde{W} \to \mathbb{R}^n$ fulfilling $\mathrm{d}C = 0$.

Proof 1. By (11.4.1), we have

$$\{I_i, H_j\} = X_{I_i}(H_j) = b_i{}^k X_{H_k}(H_j) = b_i{}^k \{H_k, H_j\} = 0.$$

In turn, this implies

$$0 = [X_{H_i}, X_{I_j}] = \big[X_{H_i}, b_j{}^k X_{H_k}\big] = X_{H_i}\big(b_j{}^k\big) X_{H_k} = \big\{H_i, b_j{}^k\big\} X_{H_k},$$

which proves $\{H_i, b_j{}^k\} = 0$, because the X_{H_i} are pointwise linearly independent on $M^{\mathscr{H}}$. As a consequence, the functions I_i and $b_i{}^j$ are constant on the maximal integral manifolds of the distribution $D^{\mathscr{H}_r}$. As noted in Remark 11.1.2/3, these are the level set components of $\mathscr{H}_r$.

2. To prove (11.4.3), denote the level set component of $\mathscr{H}_r$ containing m by Σ. Since Σ is invariant under all the flows involved, it suffices to show that the curves on the two sides of (11.4.3) are integral curves through m of the same vector field on Σ. Indeed, by (11.4.1), for all $\tilde{m} \in \Sigma$,

$$X_{I_i}(\tilde{m}) = b(\tilde{m})_i{}^k X_{H_k}(\tilde{m})$$

and by point 1, $b(\tilde{m}) = b(m)$. The statements (a)–(c) now follow. In particular, the Arnold Theorem 11.3.3 yields that the orbits of Ψ are level set components of $\mathscr{H}_r$ and the periodicity of the flows Φ^{I_i} implies that they are compact.

3. Since $\{X_{I_i}\}$ and $\{X_{\tilde{I}_i}\}$ are frames in the same distribution on $W \cap \tilde{W}$, there exists a smooth mapping $A : W \cap \tilde{W} \to \mathrm{GL}(n, \mathbb{R})$ such that

$$X_{I_i}(m) = A(m)_i{}^j X_{\tilde{I}_j}(m). \tag{11.4.6}$$

Then, (11.4.1) implies $b_i(m) = A(m)_i{}^j \tilde{b}_j$. By point 2(c), both $\{b_i(m)\}$ and $\{\tilde{b}_i(m)\}$ are systems of generators of the stabilizer of m under the action Ψ. Hence, $A(m) \in \mathrm{GL}(n, \mathbb{Z})$. Then, $\mathrm{d}A_i{}^j = 0$ and (11.4.6) implies $I_i = A_i{}^j \tilde{I}_j + C_i$ with $\mathrm{d}C_i = 0$. □

Now, we show that action and angle variables exist. Given the great importance of these coordinates, especially in perturbation theory, various constructions can be found in the literature. Our proof will follow the approach presented in Libermann and Marle [181], which in our opinion is particularly transparent. It is based on Theorem 8.6.4 and the fact that every Lagrangian submanifold admits local generating functions. Thereafter, we will show that action and angle variables can also be represented in terms of line integrals. This yields a relation to the more conventional approaches in [1], [23], [18]. Other approaches can be found in [189], [71], [80] [34] and [116]. Below, we will comment on the relation to the latter.

Theorem 11.4.3 (Existence) *Action and angle variables exist in a neighbourhood of every compact level set component of* $\mathscr{H}_r$.

Proof Let $\Sigma_{\mathbf{h}^0}$ be a compact level set component of $\mathscr{H}_r$ with value $\mathbf{h}^0 \in \mathbb{R}^n$. By Remark 11.1.2/3, $\Sigma_{\mathbf{h}^0}$ is a Lagrangian submanifold of M. Thus, Theorem 8.6.4 implies that there exists an open neighbourhood W of $\Sigma_{\mathbf{h}^0}$ in $M^{\mathscr{H}}$ and a symplectomorphism $\Phi : W \to V$ onto an open neighbourhood V of the zero section of the cotangent bundle $\mathrm{T}^*\Sigma_{\mathbf{h}^0}$. Since $\mathscr{H}_r$ is a submersion, $U := \mathscr{H}(W)$ is an open neighbourhood of $\mathbf{h}^0$ in $\mathbb{R}^n$. Let π denote the canonical projection in $\mathrm{T}^*\Sigma_{\mathbf{h}^0}$. We can shrink W so that the mapping

$$\chi = (\pi \circ \Phi, \mathscr{H}) : W \to \Sigma_{\mathbf{h}^0} \times U \tag{11.4.7}$$

becomes a diffeomorphism: since π and $\mathscr{H}_{\restriction W}$ are submersions, by counting dimensions, we find that χ'_m is bijective for all $m \in W$. Therefore, for every $m \in \Sigma_{\mathbf{h}^0}$ we can find open neighbourhoods W_m of m in W and U_m of $\mathbf{h}^0$ in U such that χ restricts to a diffeomorphism from W_m onto $(W_m \cap \Sigma_{\mathbf{h}^0}) \times U_m$. Since $\Sigma_{\mathbf{h}^0}$ is compact, the open covering of $\Sigma_{\mathbf{h}^0}$ by the subsets W_m contains a finite subcovering, labelled by $m_1, \ldots, m_r$. If we replace W by $\mathscr{H}^{-1}(\bigcap_{i=1}^r U_{m_i}) \cap (\bigcup_{i=1}^r W_{m_i})$, the mapping (11.4.7) becomes a diffeomorphism, indeed. As a consequence, the level sets of $\mathscr{H}$ in W, given by

$$\Sigma_{\mathbf{h}} := \chi^{-1}\big(\Sigma_{\mathbf{h}^0} \times \{\mathbf{h}\}\big)$$

with $\mathbf{h} \in U$, are diffeomorphic to $\Sigma_{\mathbf{h}^0}$ and hence, in particular, connected.

Let $\{\mathbf{e}_i\}$ denote the standard basis of $\mathbb{R}^n$ and let $\mathbf{b}_i$ denote the generators of the stabilizer $P_{\mathscr{H}}(\Sigma_{\mathbf{h}^0})$ under the action Ψ of $\mathbb{R}^n$ on $\Sigma_{\mathbf{h}^0}$ given by (11.3.4). For a chosen point $m_0 \in \Sigma_{\mathbf{h}^0}$, we define

$$\rho : \mathbb{R}^n \to \Sigma_{\mathbf{h}^0}, \qquad \rho(\mathbf{q}) := \Psi_{m_0}\left(\frac{1}{2\pi}\sum_{i=1}^n q_i \mathbf{b}_i\right).$$

Since $\rho(\mathbf{q} + 2\pi\mathbf{e}_i) = \rho(\mathbf{q})$ for all i, the mapping ρ defines global angle coordinates[14] on the torus $\Sigma_{\mathbf{h}^0}$, denoted by q^i. Denoting the corresponding fibre coordinates in $\mathrm{T}^*\Sigma_{\mathbf{h}^0}$ by p_i, via the symplectomorphism Φ we obtain Darboux coordinates on W, which we also denote by q^i, p_i. The idea of the proof consists in constructing a canonical transformation from q^i, p_i to the desired action and angle coordinates ϑ^i, I_i in terms of a generating function $S(\mathbf{q}, \mathbf{I})$ of the second kind.

In the first step, we construct S as a function of the variables q^i and h_i. Since $\Sigma_{\mathbf{h}}$ is Lagrange for every $\mathbf{h} \in U$, the image $\Phi(\Sigma_{\mathbf{h}})$ is a Lagrangian submanifold of $\mathrm{T}^*\Sigma_{\mathbf{h}^0}$ and Proposition 8.3.10 implies that it coincides with the image of a closed 1-form $\beta_{\mathbf{h}}$ on $\Sigma_{\mathbf{h}^0}$. Since

$$\beta_{\mathbf{h}}(m) = \Phi \circ \chi^{-1}(m, \mathbf{h}) \tag{11.4.8}$$

for all $m \in \Sigma_{\mathbf{h}^0}$ and $\mathbf{h} \in U$, this family of 1-forms is smooth. Accordingly, $\rho^*\beta_{\mathbf{h}}$ is a smooth family of closed 1-forms on $\mathbb{R}^n$. By the Poincaré Lemma, there exists a smooth function $S\colon \mathbb{R}^n \times U \to \mathbb{R}$ such that

$$\rho^*\beta_{\mathbf{h}} = \mathrm{d}S_{\mathbf{h}}, \tag{11.4.9}$$

with $S_{\mathbf{h}}(\mathbf{q}) = S(\mathbf{q}, \mathbf{h})$. Then,

$$\rho^*\beta_{\mathbf{h}}(\mathbf{q}) = \frac{\partial S}{\partial q^i}(\mathbf{q}, \mathbf{h})\mathrm{d}q^i, \tag{11.4.10}$$

where $\frac{\partial S}{\partial q^i}(\mathbf{q} + 2\pi\mathbf{e}_j, \mathbf{h}) = \frac{\partial S}{\partial q^i}(\mathbf{q}, \mathbf{h})$ for all i, j. It follows that

$$S(\mathbf{q}, \mathbf{h}) = \tilde{S}(\mathbf{q}, \mathbf{h}) + I(\mathbf{h}) \cdot \mathbf{q} \tag{11.4.11}$$

[14] That is, composition with the natural projection $\mathbb{R}^n \to \mathbb{R}^n/2\pi\mathbb{Z}^n \equiv \mathrm{T}^n$ yields a diffeomorphism.

with a smooth function $\tilde{S} : \mathbb{R}^n \times U \to \mathbb{R}^n$ satisfying $\tilde{S}(\mathbf{q} + 2\pi \mathbf{e}_i, \mathbf{h}) = \tilde{S}(\mathbf{q}, \mathbf{h})$ for all i and a smooth mapping $I : U \to \mathbb{R}^n$. For the components of the latter, we read off

$$I_i(\mathbf{h}) = \frac{1}{2\pi}\big(S(\mathbf{q} + 2\pi \mathbf{e}_i, \mathbf{h}) - S(\mathbf{q}, \mathbf{h})\big). \tag{11.4.12}$$

Since $\beta_{\mathbf{h}^0}$ coincides with the zero section in $\mathrm{T}^* \Sigma_{\mathbf{h}^0}$, we have $\frac{\partial S}{\partial q^i}(\mathbf{q}, \mathbf{h}^0) = 0$ and hence

$$I_i\big(\mathbf{h}^0\big) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\partial S}{\partial q^i}\big(\mathbf{q}, \mathbf{h}^0\big) \mathrm{d}q^i = 0. \tag{11.4.13}$$

Obviously, the functions $I_i \circ \mathscr{H}$ on W are in involution. We take them as candidates for the desired action variables and S as the generating function for the desired canonical transformation. First, we show that the variables I_i on U are in one-to-one relation with the variables h_i.

Lemma 11.4.4

1. *For all* $m \in W$, $p_i(m) = \frac{\partial S}{\partial q^i}(q(m), \mathscr{H}(m))$.
2. *For all* $\mathbf{h} \in U$, $\det(\frac{\partial I_i}{\partial h_j}(\mathbf{h})) \neq 0$.

Proof of the Lemma 1. Denote $\mathbf{h} := \mathscr{H}(m)$. By the definition of $\beta_{\mathbf{h}}$, we have $\Phi(m) = \beta_{\mathbf{h}}(\pi \circ \Phi(m))$. By the definition of the coordinates q^i on W,

$$\rho \circ q(m) = \rho \circ q\big(\pi \circ \Phi(m)\big) = \pi \circ \Phi(m).$$

Thus, we obtain

$$p_i(m) \equiv p_i \circ \Phi(m) = (\beta_{\mathbf{h}})_i\big(\pi \circ \Phi(m)\big) = (\beta_{\mathbf{h}})_i\big(\rho \circ q(m)\big) = \big(\rho^* \beta_{\mathbf{h}}\big)_i\big(q(m)\big),$$

where $(\beta_{\mathbf{h}})_i$ and $(\rho^* \beta_{\mathbf{h}})_i$ denote the coefficient functions of the 1-forms $\beta_{\mathbf{h}}$ and $\rho^* \beta_{\mathbf{h}}$ with respect to the global frames $\{\mathrm{d}q^i\}$ in $\mathrm{T}^* \Sigma_{\mathbf{h}^0}$ and $\{\rho^*(\mathrm{d}q^i) \equiv \mathrm{d}q^i\}$ in $\mathrm{T}^* \mathbb{R}^n$, respectively. Hence, the assertion follows from (11.4.10).

2. According to point 1, in the coordinates q^i, p_i on W and q^i, h_i on $\Sigma_{\mathbf{h}^0} \times U$, the inverse of χ has the form

$$(\mathbf{q}, \mathbf{h}) \mapsto \left(\mathbf{q}, \mathbf{p} = \frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}, \mathbf{h})\right).$$

Since χ is a diffeomorphism, for every $(\mathbf{q}, \mathbf{h})$,

$$\det\left(\frac{\partial^2 S}{\partial q^i \partial h_j}(\mathbf{q}, \mathbf{h})\right) \neq 0. \tag{11.4.14}$$

Assume that the assertion of point 2 does not hold. Then,

$$\sum_{j=1}^n a_j \frac{\partial I_i}{\partial h_j}(\tilde{\mathbf{h}}) = 0$$

for some $\tilde{\mathbf{h}} \in U$ and some nonzero $\mathbf{a} \in \mathbb{R}^n$, and (11.4.11) implies

$$\sum_{j=1}^n a_j \frac{\partial S}{\partial h_j}(\mathbf{q}, \tilde{\mathbf{h}}) = \sum_{j=1}^n a_j \frac{\partial \tilde{S}}{\partial h_j}(\mathbf{q}, \tilde{\mathbf{h}}) + \sum_{k=1}^n q^k \sum_{i=1}^n a_j \frac{\partial I_k}{\partial h_j}(\tilde{\mathbf{h}}) = \sum_{j=1}^n a_j \frac{\partial \tilde{S}}{\partial h_j}(\mathbf{q}, \tilde{\mathbf{h}}).$$

Since the function $\mathbf{q} \mapsto \sum_{j=1}^n a_j \frac{\partial \tilde{S}}{\partial h_j}(\mathbf{q}, \tilde{\mathbf{h}})$ is periodic, it has at least one maximum, say at $\tilde{\mathbf{q}}$. There, we have

$$\sum_{j=1}^n a_j \frac{\partial^2 S}{\partial q^i \partial h_j}(\tilde{\mathbf{q}}, \tilde{\mathbf{h}}) = 0,$$

which contradicts (11.4.14). This proves the lemma.

Proof of Theorem 11.4.3 (continued) By Lemma 11.4.4/2, U and hence W can be shrunk so that the mapping

$$I : U \to I(U) \subset \mathbb{R}^n, \qquad \mathbf{h} \mapsto I(\mathbf{h}),$$

becomes a diffeomorphism onto an open neighbourhood of the origin in $\mathbb{R}^n$ and the functions I_i provide coordinates on U. In particular, the Hamiltonian vector fields of the functions $I_i \circ \mathscr{H}$ span the distribution $D^{\mathscr{H}}$ over W. That their flows are complete and 2π-periodic will be obvious after the construction of angle coordinates. To obtain the latter, we view S as a function of $\mathbf{q}$ and $\mathbf{I}$ and define functions

$$\vartheta^j : \mathbb{R}^n \times I(U) \to \mathbb{R}, \qquad \vartheta^j(\mathbf{q}, \mathbf{I}) := \frac{\partial S}{\partial I_j}(\mathbf{q}, \mathbf{I}). \tag{11.4.15}$$

Equation (11.4.11) yields

$$\vartheta^i(\mathbf{q} + 2\pi \mathbf{e}_j, \mathbf{I}) - \vartheta^i(\mathbf{q}, \mathbf{I}) = 2\pi \delta^i{}_j \tag{11.4.16}$$

for all i, j. Hence, for every $\mathbf{I}$, the composition of $\vartheta_{\mathbf{I}} : \mathbb{R}^n \to \mathbb{R}^n$ with the natural projection $\mathrm{pr}_{\mathrm{T}^n} : \mathbb{R}^n \to \mathbb{R}^n / 2\pi \mathbb{Z}^n = \mathrm{T}^n$ induces a smooth mapping

$$\varphi : \Sigma_{\mathbf{h}^0} \to \mathbb{R}^n / 2\pi \mathbb{Z}^n = \mathrm{T}^n, \qquad \varphi\big(\rho(\mathbf{q})\big) := \mathrm{pr}_{\mathrm{T}^n}\big(\vartheta_{\mathbf{I}}(\mathbf{q})\big).$$

Since by Lemma 11.4.4/2 and Eq. (11.4.14), we have

$$\det\left(\frac{\partial^2 S}{\partial q^i \partial I_j}(\mathbf{q}, \mathbf{I})\right) \neq 0, \tag{11.4.17}$$

this mapping is a local diffeomorphism. We show that it is a global diffeomorphism, which implies that the functions ϑ^i_I induce global angle coordinates on $\Sigma_{\mathbf{h}^0}$, denoted by the same symbol. First, the image of φ coincides with T^n, because it is open and by compactness of $\Sigma_{\mathbf{h}^0}$ it is also closed. To prove that φ is injective it suffices to show that it is homotopic to a diffeomorphism. Indeed, by (11.4.12) we have

$$\vartheta^i_{\mathbf{I}}(\mathbf{q}) = q^i + \frac{\partial \tilde{S}}{\partial I_i}(\mathbf{q}, \mathbf{I}),$$

so that

$$F\colon [0,1] \times \Sigma_{\mathbf{h}^0} \to \mathrm{T}^n, \qquad F\big(s, \rho(\mathbf{q})\big) := \mathrm{pr}_{\mathrm{T}^n}\left(\mathbf{q} + s\frac{\partial \tilde{S}}{\partial \mathbf{I}}(\mathbf{q}, \mathbf{I})\right)$$

is a smooth homotopy between φ and the mapping

$$\Sigma_{\mathbf{h}^0} \to \mathrm{T}^n, \qquad \rho(\mathbf{q}) \mapsto \mathrm{pr}_{\mathrm{T}^n}(\mathbf{q}),$$

which is a diffeomorphism, because the q^i are global angle coordinates. Now, the mappings

$$W \to \mathbb{R}^n, \quad m \mapsto \vartheta\big(\pi \circ \Phi(m), I \circ \mathscr{H}(m)\big), \qquad W \to I(U), \quad m \mapsto I \circ \mathscr{H}(m),$$

define coordinates on W, for which we keep the notation ϑ^i, I_i, respectively. By construction, the functions I_i are constant on the level set components $\Sigma_{\mathbf{h}}$ and the functions ϑ^i are global angle coordinates on $\Sigma_{\mathbf{h}}$. Finally, by Lemma 11.4.4/1 and by the definition of the angle coordinates ϑ^i, on W the coordinates q^i, p_i and ϑ^i, I_i satisfy the relations

$$p_i = \frac{\partial S(q,I)}{\partial q^i}, \qquad \vartheta^i = \frac{\partial S(q,I)}{\partial I_i}.$$

We conclude that the coordinate transformation $(q^i, p_i) \mapsto (\vartheta^i, I_i)$ is canonical and, therefore, the coordinates ϑ^i, I_i are Darboux. Then, we have $X_{I_i} = \partial_{\vartheta^i}$, so that the flow of X_{I_i} is complete and 2π-periodic and hence the I_i are action variables, indeed. □

Now, we restrict the foliation (11.3.1) to the subset $\widetilde{V}^{\mathscr{H}}_c \subset \widetilde{V}^{\mathscr{H}}$ consisting of the compact level set components of $\mathscr{H}_r$. Thus, let $M^{\mathscr{H}}_c$ denote the subset of $M^{\mathscr{H}}$ of points whose level set component of $\mathscr{H}_r$ is compact. By Theorem 11.4.3, $M^{\mathscr{H}}_c$ can be covered by Darboux charts on M built from action and angle variables. By point (b) of Proposition 11.4.2/2, the domains of action and angle variables are necessarily contained in $M^{\mathscr{H}}_c$. It follows that $M^{\mathscr{H}}_c$ is an open submanifold of M.[15] Hence, $\widetilde{V}^{\mathscr{H}}_c$ is an open subset of $\widetilde{V}^{\mathscr{H}}$ and the projection (11.3.1) restricts to a projection

$$\widetilde{\mathscr{H}}_{rc} : M^{\mathscr{H}}_c \to \widetilde{V}^{\mathscr{H}}_c. \tag{11.4.18}$$

[15]In fact, this is immediate after one has shown that the mapping χ defined by (11.4.7) can be made into a diffeomorphism.

Corollary 11.4.5 *The foliation* (11.4.18) *is a locally trivial fibre bundle with typical fibre* T^n.

Proof By Theorem 11.4.3, $M_c^{\mathscr{H}}$ can be covered by Darboux charts $(\vartheta^{(\alpha)}, I^{(\alpha)})$ on $W^{(\alpha)}$ built from action and angle variables. Since $W^{(\alpha)}$ is a union of level set components of $\mathscr{H}_r$ which are labelled by $I^{(\alpha)}$, the subset

$$\widetilde{U}^{(\alpha)} := \widetilde{\mathscr{H}}_{rc}\big(W^{(\alpha)}\big)$$

of $\widetilde{V}_c^{\mathscr{H}}$ is open and $I^{(\alpha)}$ induces a local chart $\widetilde{I}^{(\alpha)} : \widetilde{U}^{(\alpha)} \to \mathbb{R}^n$. Using Proposition 11.4.2/3, it is easy to check that these charts form an atlas on $\widetilde{V}_c^{\mathscr{H}}$, thus endowing $\widetilde{V}_c^{\mathscr{H}}$ with the structure of a smooth manifold. To prove local triviality, we observe that the local charts $(\vartheta^{(\alpha)}, I^{(\alpha)})$ define diffeomorphisms

$$W^{(\alpha)} \to \mathrm{T}^n \times I^{(\alpha)}\big(W^{(\alpha)}\big). \tag{11.4.19}$$

Composing these with the diffeomorphisms $(\widetilde{I}^{(\alpha)})^{-1} : I^{(\alpha)}(W^{(\alpha)}) \to \widetilde{U}^{(\alpha)}$, we obtain diffeomorphisms

$$\chi^{(\alpha)} : W^{(\alpha)} = \widetilde{\mathscr{H}}_{rc}^{-1}\big(\widetilde{U}^{(\alpha)}\big) \to \widetilde{U}^{(\alpha)} \times \mathrm{T}^n, \qquad \chi^{(\alpha)}(m) := \big(\widetilde{\mathscr{H}}_{rc}(m), \vartheta^{(\alpha)}(m)\big),$$

which are obviously local trivializations of the projection (11.4.18). □

Remark 11.4.6 Below we list different approaches to action and angle variables. The first three are close in spirit. They are based on the original ideas of Arnold.

1. In the approach of Arnold and Avez [23], see also [1], one assumes that one has a local torus fibration and defines action and angle variables via a canonical transformation given by a certain generating function. The action variables and the generating function are defined by line integrals over a symplectic potential. Here, point 2 of Lemma 11.4.4 is usually imposed as an assumption.
2. In the approach of Nekhoroshev [228], the Tubular Neighbourhood Theorem for embedded submanifolds is used to prove that a neighbourhood of any torus has the structure of a local torus fibration. On such a neighbourhood, the action variables are defined via line integrals over a symplectic potential. Angle variables are constructed using the Liouville Theorem.[16]
3. In the approach of Libermann and Marle [181], the Weinstein Theorem 8.6.4 and the fact that every Lagrangian submanifold admits a local generating function is used to find a generating function S of a canonical transformation providing action and angle variables. Here, the action variables are obtained directly from the periodicity properties of S.

[16] In view of the classical Carathéodory-Jacobi-Lie Theorem, this construction generalizes to noncommutative integrable systems, cf. Theorem 11.8.3.

4. The approach of Markus and Meyer [189], see also Duistermaat [80], Bates and Śniatycki [34] and Guillemin and Sternberg [116], is different. Here, the action variables are derived from a local frame in the period bundle,[17] made up by the stabilizers $P_{\mathscr{H}}(\Sigma)$ (the period lattices) of the level set components of $\mathscr{H}_r$ under the action of $\mathbb{R}^n$ defined by the flows of the Hamiltonian vector fields X_{H_i}, cf. Remark 11.3.5. Angle variables are constructed via the Liouville Theorem.

To build the bridge to the original approach of Arnold and Avez, described in point 1 of Remark 11.4.6, we derive explicit formulae for the action and angle variables constructed in the proof of Theorem 11.4.3 in terms of line integrals. We use the notation of that proof. Without loss of generality, as the potential $S(\mathbf{q},\mathbf{h})$ of the smooth family of closed 1-form $\rho^*\beta_{\mathbf{h}}$ we may choose

$$S(\mathbf{q},\mathbf{h}) = \int_{\tau_{\mathbf{q}}} \rho^*\beta_{\mathbf{h}} \quad \text{with } \tau_{\mathbf{q}} : [0,1] \to \mathbb{R}^n, \qquad \tau_{\mathbf{q}}(t) := t\mathbf{q}. \tag{11.4.20}$$

This can be rewritten as an integral over $\Sigma_{\mathbf{h}}$ as follows. Let θ denote the canonical 1-form in $\mathrm{T}^*\Sigma_{\mathbf{h}^0}$. By (8.3.2),

$$S(\mathbf{q},\mathbf{h}) = \int_{\tau_{\mathbf{q}}} \rho^*\beta_{\mathbf{h}} = \int_{\rho\circ\tau_{\mathbf{q}}} \beta_{\mathbf{h}} = \int_{\beta_{\mathbf{h}}\circ\rho\circ\tau_{\mathbf{q}}} \theta.$$

Here, $\beta_{\mathbf{h}}\circ\rho\circ\tau_{\mathbf{q}}$ is a curve in $\mathrm{T}^*\Sigma_{\mathbf{h}^0}$ which is contained in the image of $\Sigma_{\mathbf{h}}$ under Φ. Hence, there is a unique curve $\gamma_{\mathbf{q}}^{\mathbf{h}}$ in $\Sigma_{\mathbf{h}}$ such that

$$\beta_{\mathbf{h}}\circ\rho\circ\tau_{\mathbf{q}} = \Phi\circ\gamma_{\mathbf{q}}^{\mathbf{h}}.$$

Thus,

$$S(\mathbf{q},\mathbf{h}) = \int_{\gamma_{\mathbf{q}}^{\mathbf{h}}} \Phi^*\theta. \tag{11.4.21}$$

To obtain from (11.4.21) an integral representation for the action variables I_i, denote $\gamma_i^{\mathbf{h}} \equiv \gamma_{2\pi\mathbf{e}_i}^{\mathbf{h}}$. By (11.4.12) and (11.4.21),

$$I_i(\mathbf{h}) = \frac{1}{2\pi}\int_{\gamma_i^{\mathbf{h}}} \Phi^*\theta. \tag{11.4.22}$$

Here, $\Phi^*\theta$ is a potential for the symplectic form ω and the curve $\gamma_i^{\mathbf{h}}$ encloses the i-th factor[18] of $\Sigma_{\mathbf{h}}$ and is contractible with respect to the remaining factors. Thus, $\{\gamma_i^{\mathbf{h}}\}$ is a system of fundamental cycles[19] for $\Sigma_{\mathbf{h}}$. Now, since the mapping $\mathbf{h}\mapsto I(\mathbf{h})$

---

[17] To be introduced in Sect. 11.7.

[18] As defined by the diffeomorphism $\Sigma_{\mathbf{h}^0} \cong \mathrm{T}^n$ induced by the mapping ρ.

[19] A system of closed curves whose homotopy classes yield a minimal set of generators for the fundamental group of $\Sigma_{\mathbf{h}}$.

satisfies the condition in Lemma 11.4.4/2, one can express $\mathbf{h}$ in (11.4.21) in terms of $\mathbf{I}$. Thus, we obtain the generating function S as

$$S(\mathbf{q}, \mathbf{I}) = \int_{\gamma_{\mathbf{q}}^{\mathbf{I}}} \Phi^* \theta, \tag{11.4.23}$$

with $\gamma_{\mathbf{q}}^{\mathbf{I}} = \gamma_{\mathbf{q}}^{h(\mathbf{I})}$, and the angle variables ϑ^i are given by (11.4.15). That the latter satisfy (11.4.16) can also be read off from the integral representation (11.4.23), because integration over the curve $\gamma_{\mathbf{q}+2\pi\mathbf{e}_i}^{\mathbf{h}}$ yields

$$S(\mathbf{q} + 2\pi \mathbf{e}_i, \mathbf{I}) = \int_{\gamma_{\mathbf{q}+2\pi\mathbf{e}_i}^{\mathbf{I}}} \Phi^* \theta = \int_{\gamma_{\mathbf{q}}^{\mathbf{I}}} \Phi^* \theta + \int_{\gamma_i^{\mathbf{I}}} \Phi^* \theta = S(\mathbf{q}, \mathbf{I}) + 2\pi I_i .$$

Now, we show that Formulae (11.4.22) and (11.4.23), with $\Phi^*\theta$ replaced by an arbitrary symplectic potential for ω, $\{\gamma_i^{\mathbf{h}}\}$ replaced by an arbitrary system of fundamental cycles and $\gamma_{\mathbf{q}}^{\mathbf{h}}$ replaced by some homotopic curve allow for locally defining action and angle variables.

Theorem 11.4.7 (Representation by line integrals) *Let $m_0 \in M$, let $\mathbf{h}^0 = \mathscr{H}(m_0)$ be a regular value of $\mathscr{H}$ and assume that the connected component $\Sigma_{\mathbf{h}^0}$ of m_0 of the level set $\mathscr{H}^{-1}(\mathbf{h}^0)$ is compact. Let the following data be given*:

1. *a connected open neighbourhood W of $\Sigma_{\mathbf{h}^0}$ in $M^{\mathscr{H}}$ such that*
$$\Sigma_{\mathbf{h}} := W \cap \mathscr{H}_r^{-1}(\mathbf{h})$$
is a compact level set component of $\mathscr{H}_r$ for all $\mathbf{h} \in U = \mathscr{H}(W)$,
2. *a potential τ for ω on W,*
3. *smooth families $\{\gamma_1^{\mathbf{h}}\}, \dots, \{\gamma_n^{\mathbf{h}}\}$ of curves in W such that $\{\gamma_1^{\mathbf{h}}, \dots, \gamma_n^{\mathbf{h}}\}$ is a system of fundamental cycles in $\Sigma_{\mathbf{h}}$ for all $\mathbf{h} \in U$,*
4. *Darboux coordinates q^i, p_i on some open neighbourhood V of m_0 in W such that $\tau = p_i \mathrm{d}q^i$ and such that, for all $\mathbf{h} \in U$, the subset $V \cap \Sigma_{\mathbf{h}}$ is simply connected and can be coordinatized by the q^i,*
5. *a smooth mapping $s : U \to V$ such that $\mathscr{H} \circ s = \mathrm{id}_U$.*

Define

$$I_i(\mathbf{h}) := \frac{1}{2\pi} \int_{\gamma_i^{\mathbf{h}}} \tau, \qquad S(\mathbf{q}, \mathbf{h}) := \int_{\gamma_{\mathbf{q}}^{\mathbf{h}}} \tau \tag{11.4.24}$$

where $\gamma_{\mathbf{q}}^{\mathbf{h}}$ is some curve in $V \cap \Sigma_{\mathbf{h}}$ from $s(\mathbf{h})$ to the point with coordinate value $\mathbf{q}$. Then, V and W can be shrunk so that

1. *The first equation in (11.4.24) can be resolved for $\mathbf{h}$ and hence S can be written as a function of $\mathbf{q}$ and $\mathbf{I}$,*
2. *the functions ϑ^i on V defined by*
$$\vartheta^i(m) = \frac{\partial S}{\partial I_i}\big(q(m), I(m)\big), \tag{11.4.25}$$

where $I \equiv I \circ \mathscr{H}_{\restriction W}$, *complement the functions* I_i *to Darboux coordinates on* V *which have a unique extension to action and angle variables on* W.

For the proof, we need

Lemma 11.4.8 *Let* $\mathbf{h} \in U$.

1. *If* γ_1 *and* γ_2 *are curves in* $\Sigma_{\mathbf{h}}$ *which are homotopic with fixed endpoints or if they are closed curves in* $\Sigma_{\mathbf{h}}$ *which are homotopic in* $\Sigma_{\mathbf{h}}$, *then* $\int_{\gamma_1} \tau = \int_{\gamma_2} \tau$.
2. *If* τ_1 *and* τ_2 *are potentials for* ω *on* W, *then* $\int_{\gamma_i^{\mathbf{h}}} \tau_2 - \int_{\gamma_i^{\mathbf{h}}} \tau_1$ *does not depend on* $\mathbf{h}$.

Proof of Lemma 11.4.8 1. Let $F : [0, 1] \times [0, 1] \to \Sigma_{\mathbf{h}}$ be a smooth homotopy from γ_1 to γ_2, viewed for simplicity as a surface in $\Sigma_{\mathbf{h}}$. By Stokes' Theorem and the fact that $\Sigma_{\mathbf{h}}$ is Lagrange,

$$\int_{\gamma_1} \tau - \int_{\gamma_2} \tau = \int_{\partial F} \tau = \int_F \omega = 0.$$

2. Since W is connected, so is U. Hence, for every $\mathbf{h} \in U$ we find a smooth curve $t \mapsto \mathbf{h}(t)$ such that $\mathbf{h}(0) = \mathbf{h}^0$ and $\mathbf{h}(1) = \mathbf{h}$. Then, the mapping $F_i : [0, 1] \times [0, 1] \to W$ defined by $F_i(t, r) := \gamma_i^{\mathbf{h}(t)}(r)$ represents a smooth surface in W and Stokes' Theorem yields

$$\int_{\gamma_i^{\mathbf{h}}} (\tau_2 - \tau_1) - \int_{\gamma_i^{\mathbf{h}^0}} (\tau_2 - \tau_1) = \int_{F_i} \mathrm{d}(\tau_2 - \tau_1) = 0.$$

□

Proof of Theorem 11.4.7 The idea of the proof is to relate the functions I_i and ϑ^i defined in the proposition with the action and angle variables constructed in the proof of Theorem 11.4.3 in a neighbourhood of $\Sigma_{\mathbf{h}^0}$ and with the point m_0 chosen to define the mapping ρ used there. Here, the objects from that proof will be denoted by their original symbols endowed with a hat. For simplicity, without loss of generality we may assume $\hat{W} = W$ and $\hat{U} = U$.

In the first step, we relate I with $\hat{I}$. Let $\mathbf{h} \in U$ be given. Since both $\{\gamma_i^{\mathbf{h}}\}$ and $\{\hat{\gamma}_i^{\mathbf{h}}\}$ are systems of fundamental cycles for $\Sigma_{\mathbf{h}}$, there are unique integer-valued $(n \times n)$-matrices A and B such that $\gamma_i^{\mathbf{h}}$ is homotopic to $A_i{}^j \hat{\gamma}_j^{\mathbf{h}}$ and $\hat{\gamma}_i^{\mathbf{h}}$ is homotopic to $B_i{}^j \gamma_j^{\mathbf{h}}$. Here, by the sum of two curves we mean the composite curve. Up to homotopy, the order of composition is not relevant. Obviously, A and B are inverse to one another. Using Lemma 11.4.8 and (11.4.22), we obtain

$$I_i(\mathbf{h}) = \frac{1}{2\pi} \int_{\gamma_i^{\mathbf{h}}} \Phi^*\theta + C_i = \frac{1}{2\pi} \int_{A_i{}^j \hat{\gamma}_j^{\mathbf{h}}} \Phi^*\theta + C_i = A_i{}^j \hat{I}_j(\mathbf{h}) + C_i,$$

where $(C_1, \dots, C_n)$ is a vector in $\mathbb{R}^n$ which is independent of $\mathbf{h}$. Since the families $\{\gamma_i^{\mathbf{h}}\}$ depend smoothly on $\mathbf{h} \in U$, so do the coefficients $A_i{}^j$ and $B_i{}^j$. Since U is connected, these coefficients are therefore constant in $\mathbf{h}$. To summarize, the functions I_i and $\hat{I}_i$ on W are related by an invertible affine transformation with constant coefficients:

$$I_i = A_i{}^j \hat{I}_j + C_i, \qquad \hat{I}_i = B_i{}^j (I_j - C_j). \tag{11.4.26}$$

In particular, W and hence U and V can be shrunk so that the first equation in (11.4.24) can be resolved for $\mathbf{h}$, because this is true for the corresponding equation for $\hat{I}$.

In the second step, we show that the functions ϑ^i are well defined and complement the functions I_i to Darboux coordinates on a possibly smaller V. Consider the integral in the definition of S. According to point 1 of Lemma 11.4.8, the value of this integral does not depend on the homotopy class with fixed endpoints of the curve $\gamma_{\mathbf{q}}^{\mathbf{h}}$. Since $V \cap \Sigma_{\mathbf{h}}$ is simply connected, it therefore depends on $\mathbf{q}$ and $\mathbf{h}$ only, and the dependence is smooth. Thus, S is a smooth function in the variables $\mathbf{q}$ and $\mathbf{h}$, and by expressing $\mathbf{h}$ in terms of $\mathbf{I}$ we find that the ϑ^i are well-defined smooth functions of $\mathbf{q}$ and $\mathbf{I}$. By plugging in for $\mathbf{q}$ and $\mathbf{I}$ the mappings q and $I \equiv I \circ \mathscr{H}_{\restriction W}$, ϑ^i and S become functions on V. Let $\iota_{\mathbf{h}} : V \cap \Sigma_{\mathbf{h}} \to V$ be the natural inclusion mapping. By construction,

$$\iota_{\mathbf{h}}^* \mathrm{d}S = \iota_{\mathbf{h}}^* \tau. \tag{11.4.27}$$

Since the functions q^i provide coordinates on $V \cap \Sigma_{\mathbf{h}}$ for all $\mathbf{h} \in U$, the functions q^i and I_i yield coordinates on V. In these coordinates, the left hand side reads

$$\iota_{\mathbf{h}}^* \mathrm{d}S = \iota_{\mathbf{h}}^* \big((\partial_{q^i} S) \mathrm{d}q^i + (\partial_{I_i} S) \mathrm{d}I_i \big) = \big((\partial_{q^i} S) \circ \iota_{\mathbf{h}} \big) \mathrm{d}\big(q^i \circ \iota_{\mathbf{h}} \big),$$

because $I_i \circ \iota_{\mathbf{h}}$ is constant. On the other hand, in the coordinates q^i and p_i, the right hand side reads

$$\iota_{\mathbf{h}}^* \tau = (p_i \circ \iota_{\mathbf{h}}) \mathrm{d}\big(q^i \circ \iota_{\mathbf{h}} \big).$$

Since the $\mathrm{d}(q^i \circ \iota_{\mathbf{h}})$ are pointwise linearly independent, and since (11.4.27) holds for all $\mathbf{h} \in U$, we conclude that

$$\partial_{q^i} S = p_i \tag{11.4.28}$$

and hence

$$\mathrm{d}S = p_i \mathrm{d}q^i + \vartheta^i \mathrm{d}I_i \tag{11.4.29}$$

on all of V. By taking the exterior differential of (11.4.29), we obtain $\omega = \mathrm{d}I_i \wedge \mathrm{d}\vartheta^i$. First, this implies that the mapping $(\vartheta, I) : V \to \mathbb{R}^{2n}$ is a local diffeomorphism. Hence, we may shrink V so that ϑ^i and I_i define coordinates. Second, this implies that these coordinates are Darboux.

In the third step, we show that ϑ^i and I_i extend to unique action and angle coordinates on W, where W is shrunk appropriately according to V. Since the I_i are already defined on W, it suffices to consider the ϑ^i. By (11.4.26), on V we have

$$\mathrm{d}I_i \wedge \mathrm{d}\vartheta^i = \omega = \mathrm{d}\hat{I}_i \wedge \mathrm{d}\hat{\vartheta}^i = \mathrm{d}\big(B_i{}^j I_j \big) \wedge \mathrm{d}\hat{\vartheta}^i = \mathrm{d}I_j \wedge \mathrm{d}\big(B_i{}^j \hat{\vartheta}^i \big).$$

Thus, by Remark 8.1.6/3, ${B_i}^j \hat{\vartheta}^i = \vartheta^j + \alpha^j \circ I$ and hence

$$\vartheta^i = {B_j}^i \hat{\vartheta}^j - \alpha^i \circ I, \tag{11.4.30}$$

where α^i are the coefficient functions of a closed 1-form on $I(V) \subset \mathbb{R}^n$. The functions on the right hand side are defined on W. Hence, this equation defines extensions of ϑ^i to functions $\tilde{\vartheta}^i$ on W. These functions arise from the $\hat{\vartheta}^i$ by an invertible linear transformation with integer coefficients, followed by a shift which depends on m only through I, that is, which is constant on each $\Sigma_{\mathbf{h}}$. Since the $\hat{\vartheta}^i$ provide global angle coordinates on each $\Sigma_{\mathbf{h}}$ in W, so do the $\tilde{\vartheta}^i$. Thus, $\tilde{\vartheta}^i$ and I_i are action and angle coordinates on W. That the $\tilde{\vartheta}^i$ are uniquely determined by their restrictions ϑ^i to V follows once more from Remark 8.1.6/3. □

Remark 11.4.9

1. According to (11.4.25) and (11.4.28), the function S defined by (11.4.24) is a generating function of the second kind for the canonical coordinate transformation $(q, p) \mapsto (\vartheta, I)$ on V. In particular,

$$\det\left(\frac{\partial^2 S}{\partial q^i \partial I_j}\right) \neq 0.$$

This generating function is related to the generating function used in the proof of Theorem 11.4.3, here denoted by $\hat{S}$, as follows. In view of (11.4.30), from (11.4.29) and the corresponding equation

$$\mathrm{d}\hat{S} = p_i \mathrm{d}q^i + \hat{\vartheta}^i \mathrm{d}\hat{I}_i$$

we read off that

$$\mathrm{d}(\hat{S} - S) = \hat{\vartheta}^j \mathrm{d}\hat{I}_j - \vartheta^j \mathrm{d}I_j = I^* \alpha.$$

Since the 1-form α on $I(U)$ is closed, on every contractible subset of V, the generating functions $\hat{S}$ and S differ by a function of the action variables I_i.
2. For later reference we note the following representation of (11.4.24) in terms of local Darboux coordinates:

$$I_i(\mathbf{h}) = \frac{1}{2\pi} \int_{\gamma_i^{\mathbf{h}}} p_i \mathrm{d}q^i, \qquad S(\mathbf{q}, \mathbf{h}) = \int_{\gamma_{\mathbf{q}}^{\mathbf{h}}} p_i \mathrm{d}q^i. \tag{11.4.31}$$

To conclude this section, we discuss dynamics in terms of action and angle variables. Since the Hamiltonian H is constant on level set components of $\mathscr{H}$, in action and angle variables ϑ^i and I_i, it depends on the I_i only. Hence, with the notation

$$\omega^j := \frac{\partial H}{\partial I_j},$$

the Hamilton equations take the following simple form:

$$\dot{\vartheta}^j(t) = \omega^j\big(\mathbf{I}(t)\big), \qquad \dot{I}_j(t) = -\frac{\partial H}{\partial \vartheta^j} = 0. \tag{11.4.32}$$

The corresponding integral curves are given by

$$\vartheta^j(t) = \omega^j(\mathbf{I})t + \vartheta^j(0), \qquad I_j(t) = I_j(0). \tag{11.4.33}$$

The functions $\omega^j(\mathbf{I})$ are called the characteristic frequencies associated with the action variables I_j. To summarize, the action variables I_i are functionally independent constants of motion in involution. For each value $\mathbf{I}$, the dynamics of the system reduces to a motion on the torus $\Sigma_{\mathbf{I}}$ with constant angular velocities $\omega^i(\mathbf{I})$. At this point, the reader should recall the discussion of Sect. 9.6, where we found this structure in the neighbourhood of a critical point of the Hamiltonian function of an arbitrary Hamiltonian system. As noted there, there are two qualitatively distinct cases: if the frequencies $\omega^i(\mathbf{I})$ are rationally independent, that is, if for all nonzero $\mathbf{k} \in \mathbb{Z}^n$ one has

$$\sum\nolimits_{i=1}^{n} k_i \omega^i(\mathbf{I}) \neq 0,$$

each integral curve is dense in $\Sigma_{\mathbf{I}}$. In this case, the torus $\Sigma_{\mathbf{I}}$ is said to be non-resonant and the motion on it is said to be quasiperiodic. If, in contrast, the frequencies $\omega^i(\mathbf{I})$ are rationally dependent, $\Sigma_{\mathbf{I}}$ is said to be resonant. In this case, we have

Proposition 11.4.10 *If the frequencies $\omega_0^i = \omega^i(\mathbf{I}_0)$ are rationally dependent, there exists a canonical transformation $(\vartheta, I) \mapsto (\bar{\vartheta}, \bar{I})$ to new action and angle variables such that a certain number $l < n$ of the new frequencies is rationally independent on $\Sigma_{\mathbf{I}_0}$, whereas the remaining frequencies vanish.*

Thus, the torus $\Sigma_{\mathbf{I}_0}$ is decomposed into l-dimensional invariant subtori and the motion on these subtori is quasiperiodic.

Proof The frequencies ω_0^i define a module

$$\left\{ \sum_{i=1}^{n} \omega_0^i k_i : k_i \in \mathbb{Z} \right\}$$

over the integers. Let $\{b^1, \ldots, b^l\}$ be a basis in this module. Since, by assumption, the real numbers ω_0^i are rationally dependent, $l < n$ and one has $n - l$ rationally independent equations

$$B^i{}_j \omega_0^j = 0, \quad i = l+1, \ldots, n, \; j = 1, \ldots, n,$$

with $B^i{}_j$ being an integer-valued matrix whose rows $(B^i{}_1, \ldots, B^i{}_n)$, $i = l+1, \ldots, n$, are rationally independent. The basis elements b^i possess a (not necessarily unique) decomposition

$$b^i = D^i{}_j \omega_0^j, \quad i = 1, \ldots, l, \ j = 1, \ldots, n.$$

We combine the matrices B and D to an $(n \times n)$-matrix $A = [D^{\mathrm{T}} B^{\mathrm{T}}]$ with integer entries and define

$$\bar{\vartheta} := A^{\mathrm{T}} \vartheta, \qquad \bar{I} := A^{-1} I. \tag{11.4.34}$$

This transformation is canonical, because $I_j \mathrm{d}\vartheta^j = \bar{I}_j \mathrm{d}\bar{\vartheta}^j$. Moreover, since A is an integer matrix, $\bar{\vartheta}^i$ and $\bar{I}_i$ are action and angle variables again. In the new variables, the Hamilton equations take the form

$$\dot{\bar{\vartheta}}^i = \frac{\partial H}{\partial \bar{I}_i}(\bar{\mathbf{I}}) \equiv \bar{\omega}^i(\bar{\mathbf{I}}), \qquad \dot{\bar{I}}_i = -\frac{\partial H}{\partial \bar{\vartheta}^i} = 0,$$

where the Hamiltonian is given by $H(A\bar{I})$. For the new frequencies $\bar{\omega}_0^i$ corresponding to $\bar{I}_0 = A^{-1} I_0$ we obtain

$$\bar{\omega}_0^i = \frac{\partial H}{\partial \bar{I}_i}(\bar{I}_0) = \frac{\partial I_j}{\partial \bar{I}_i} \frac{\partial H}{\partial I_j}(I_0) = A^i{}_j \omega_0^j = \begin{cases} b^i & i \leq l, \\ 0 & i > l. \end{cases}$$

This proves the proposition. □

Finally, let us recall the important notion of degeneracy. If the frequencies do not depend on the action variables, the system is called isochronous.Otherwise, it is called anisochronous. In the latter case, it is said to be non-degenerate if

$$\det\left(\frac{\partial^2 H}{\partial I_i \partial I_j}\right) = \det\left(\frac{\partial \omega^i}{\partial I_j}\right) \neq 0. \tag{11.4.35}$$

Then, the mapping $\mathbf{I} \mapsto \boldsymbol{\omega}(\mathbf{I})$ is a local diffeomorphism and one can use the frequencies as (non-canonical) coordinates labelling the tori. In this case, almost all tori are quasiperiodic. Nonetheless, there is a dense set of resonant tori as well. If (11.4.35) does not hold, the system is said to be degenerate. Often the occurrence of degeneracies is related to the fact that the system possesses more than n functionally independent constants of motion.[20] This situation will be discussed in Sect. 11.8.

11.5 Examples

In this section we present the construction of action and angle variables for the harmonic oscillator, the Kepler problem and the symmetric Euler top. For the first two of these examples, we will use Theorem 11.4.7.

[20] Which, of course, cannot all be in involution.

Example 11.5.1 (Harmonic Oscillator) We consider the Hamiltonian system $(\mathbb{R}^2, \mathrm{d}p \wedge \mathrm{d}q, H)$ with

$$H(q, p) = \frac{p^2}{2m} + \frac{kq^2}{2}.$$

Being autonomous and having one degree of freedom, this system is integrable, where $\mathscr{H} = H$. Thus, the level sets of $\mathscr{H}$ coincide with the energy surfaces

$$\Sigma_E = \left\{(q, p) \in \mathbb{R}^2 : \frac{p^2}{2m} + \frac{kq^2}{2} = E\right\},$$

which at the same time coincide with the orbits of the system, that is, with the images of the maximal integral curves of the Hamiltonian vector field X_H. For $E > 0$, Σ_E is an ellipse, whereas for $E = 0$, it consists of the origin, which is the only critical point of X_H. In either case, Σ_E is connected and compact. Thus, we can apply Theorem 11.4.7. Let $W = \mathbb{R}^2 \setminus \{0\}$ and $\tau = p\mathrm{d}q$. Since the fundamental cycle γ^E is given by Σ_E itself, and since on Σ_E we have

$$p(q, E) = \sqrt{2m\left(E - \frac{k}{2}q^2\right)},$$

where q runs between $-\sqrt{\frac{2E}{k}}$ and $\sqrt{\frac{2E}{k}}$, for the action variable we obtain

$$I(E) = \frac{1}{2\pi}\int_{\Sigma_E} \tau = \frac{1}{\pi}\int_{-\sqrt{\frac{2E}{k}}}^{\sqrt{\frac{2E}{k}}} p(q, E)\mathrm{d}q = \frac{E}{\omega} \tag{11.5.1}$$

with $\omega = \sqrt{\frac{k}{m}}$. To calculate the generating function and the angle variable, we choose, for example, $V = \{(q, p) : p > 0\}$ and obtain

$$S(q, I) = I \arcsin\left(\sqrt{\frac{m\omega}{2I}}\, q\right) + \frac{q}{2}\sqrt{2m\omega I - m^2\omega^2 q^2}, \tag{11.5.2}$$

$$\vartheta = \arcsin\left(\sqrt{\frac{m\omega}{2I}}\, q\right) \tag{11.5.3}$$

(Exercise 11.5.1). By extending ϑ in the obvious way to an angle coordinate on W, we obtain the desired action and angle variables on $W = \mathbb{R}^2 \setminus \{0\}$. Obviously, these coordinates induce a symplectomorphism $(W, \mathrm{d}p \wedge \mathrm{d}q) \cong (\mathrm{S}^1 \times \mathbb{R}_+, \mathrm{d}I \wedge \mathrm{d}\vartheta)$. In the coordinates ϑ and I, the Hamiltonian is given by $H = I\omega$ and hence the frequency corresponding to I is

$$\omega(I) = \frac{\partial H}{\partial I} = \omega.$$

Thus, for $E > 0$, we have a periodic motion on the torus Σ_I with frequency ω:

$$\vartheta(t) = \omega \cdot t + \vartheta(0), \qquad I(t) = I(0).$$

Example 11.5.2 (Kepler Problem) Let us consider the Hamiltonian system $(\mathrm{T}^*(\mathbb{R}^3 \setminus \{0\}), \mathrm{d}\theta, H)$ with

$$H(\mathbf{q}, \mathbf{p}) = \frac{\mathbf{p}^2}{2} - \frac{1}{\|\mathbf{q}\|}.$$

From the discussion in Example 11.1.3 we know that the functions H, $\mathbf{L}^2$ and L_z are constants of motion in involution which are functionally independent on a dense subset of $\mathrm{T}^*(\mathbb{R}^3 \setminus \{0\})$. It is clear that the level sets of $\mathscr{H}$ are compact only in the case

$$-\frac{1}{2\mathbf{L}^2} < E < 0, \qquad \mathbf{L} \neq 0. \tag{11.5.4}$$

In spherical coordinates $r \equiv \|\mathbf{q}\|$, ϑ, ϕ and the corresponding fibre coordinates (conjugate momenta) p_r, p_ϑ, p_ϕ, the constants of motion read

$$H = \frac{1}{2}\left(p_r^2 + \frac{p_\vartheta^2}{r^2} + \frac{p_\phi^2}{r^2 \sin^2 \vartheta}\right) - \frac{1}{r}, \qquad \mathbf{L}^2 = p_\vartheta^2 + \frac{p_\phi^2}{r^2 \sin^2 \vartheta}, \qquad L_z = p_\phi \tag{11.5.5}$$

(Exercise 11.5.2). Denote the values of $\mathscr{H} = (H, \mathbf{L}^2, L_z)$ by $\mathbf{h} = (E, \Theta^2, \Phi)$. On the level set $\Sigma_\mathbf{h}$, the momenta satisfy the relations

$$p_r^2 = 2E + \frac{2}{r} - \frac{\Theta^2}{r^2}, \qquad p_\vartheta^2 = \Theta^2 - \frac{\Phi^2}{\sin^2 \vartheta}, \qquad p_\phi = \Phi. \tag{11.5.6}$$

Since $\Phi^2 \leq \frac{\Phi^2}{\sin^2 \vartheta} \leq \Theta^2$, the right hand side of the second equation in (11.5.6) has two zeros, $\vartheta^\mathbf{h}_{\min} \leq \vartheta^\mathbf{h}_{\max}$, and it is non-negative in the interval between $\vartheta^\mathbf{h}_{\min}$ and $\vartheta^\mathbf{h}_{\max}$. Thus, on this interval, we can define a function $p_\vartheta(\vartheta, \mathbf{h})$ by the positive square root of the right hand side. In the same way, since $E < 0$, the right hand side of the first equation in (11.5.6) has two zeros, $r^\mathbf{h}_{\min} \leq r^\mathbf{h}_{\max}$, and in the interval $r^\mathbf{h}_{\min} \leq r \leq r^\mathbf{h}_{\max}$, a function $p_r(r, \mathbf{h})$ can be defined by the positive square root. To construct fundamental cycles on the tori $\Sigma_\mathbf{h}$, define curves

$$\begin{aligned}
\gamma^\mathbf{h}_{r,\pm}(t) &:= \big(t, 0, 0; \pm p_r(t, \mathbf{h}), 0, \Phi\big), & t &\in \big[r^\mathbf{h}_{\min}, r^\mathbf{h}_{\max}\big],\\
\gamma^\mathbf{h}_{\vartheta,\pm}(t) &:= \big(r^\mathbf{h}_{\min}, t, 0; 0, \pm p_\vartheta(t, \mathbf{h}), \Phi\big), & t &\in \big[\vartheta^\mathbf{h}_{\min}, \vartheta^\mathbf{h}_{\max}\big],\\
\gamma^\mathbf{h}_{\phi}(t) &:= \big(r^\mathbf{h}_{\min}, \vartheta^\mathbf{h}_{\min}, t; 0, 0, \Phi\big), & t &\in [0, 2\pi].
\end{aligned}$$

Denote $\gamma^\mathbf{h}_\vartheta := \gamma^\mathbf{h}_{\vartheta,-} \circ \gamma^\mathbf{h}_{\vartheta,+}$ and $\gamma_r := \gamma^\mathbf{h}_{r,-} \circ \gamma^\mathbf{h}_{r,+}$. Then, $\{\gamma^\mathbf{h}_r\}$, $\{\gamma^\mathbf{h}_\vartheta\}$ and $\{\gamma^\mathbf{h}_\phi\}$ are smooth families of curves and, for every $\mathbf{h}$, $\{\gamma^\mathbf{h}_r, \gamma^\mathbf{h}_\vartheta, \gamma^\mathbf{h}_\phi\}$ is a system of fundamental cycles on $\Sigma_\mathbf{h}$. Using these cycles and choosing the canonical 1-form θ as a potential for the symplectic form, we calculate the action variables:

$$I_1(\mathbf{h}) = \frac{1}{2\pi}\int_{\gamma^\mathbf{h}_r} \theta = \frac{1}{\pi}\int_{r^\mathbf{h}_{\min}}^{r^\mathbf{h}_{\max}} p_r(t, \mathbf{h})\mathrm{d}t = \frac{1}{\sqrt{-2E}} - \Theta,$$

$$I_2(\mathbf{h}) = \frac{1}{2\pi}\int_{\gamma_\vartheta^{\mathbf{h}}} \theta = \frac{1}{\pi}\int_{\vartheta_{\min}^{\mathbf{h}}}^{\vartheta_{\max}^{\mathbf{h}}} p_\vartheta(t,\mathbf{h})\mathrm{d}t = \Theta - \Phi, \tag{11.5.7}$$

$$I_3(\mathbf{h}) = \frac{1}{2\pi}\int_{\gamma_\phi^{\mathbf{h}}} \theta = \frac{1}{2\pi}\int_0^{2\pi} \Phi\mathrm{d}t = \Phi.$$

For the inverse of the mapping $(E, \Theta, \Phi) \mapsto (I_1, I_2, I_3)$ we find

$$E = -\frac{1}{2(I_1+I_2+I_3)^2}, \qquad \Theta = I_2 + I_3, \qquad \Phi = I_3. \tag{11.5.8}$$

Finally, to compute S and the angle variables $\vartheta^1, \vartheta^2, \vartheta^3$, we choose the mapping $s : U \to V$ as

$$s(\mathbf{h}) = \left(r_{\min}^{\mathbf{h}}, \vartheta_{\min}^{\mathbf{h}}, 0; 0, 0, \Phi\right)$$

and express $\mathbf{h}$ in terms of $\mathbf{I}$ and θ in terms of the variables r, ϑ, ϕ and I_1, I_2, I_3. We find

$$S = S_r(r, I_1, I_2, I_3) + S_\vartheta(\vartheta, I_2, I_3) + S_\phi(\phi, I_3)$$

with

$$S_r = \int_{r_{\min}^{\mathbf{I}}}^{r} \sqrt{2\left(\frac{1}{t} - \frac{1}{2(I_1+I_2+I_3)^2}\right) - \frac{(I_2+I_3)^2}{t^2}}\mathrm{d}t,$$

$$S_\vartheta = \int_{\vartheta_{\min}^{\mathbf{I}}}^{\vartheta} \sqrt{(I_2+I_3)^2 - \frac{I_3^2}{\sin^2 t}}\mathrm{d}t,$$

$$S_\phi = I_3\phi$$

and

$$\vartheta^1 = \frac{\partial S_r}{\partial I_1}, \qquad \vartheta^2 = \frac{\partial S_r}{\partial I_2} + \frac{\partial S_\vartheta}{\partial I_2}, \qquad \vartheta^3 = \frac{\partial S_r}{\partial I_3} + \frac{\partial S_\vartheta}{\partial I_3} + \phi.$$

By (11.5.8), in the action and angle variables ϑ^i, I_i, the Hamiltonian reads

$$H(\mathbf{I}) = -\frac{1}{2(I_1+I_2+I_3)^2}$$

and the characteristic frequencies are

$$\omega^1(\mathbf{I}) = \omega^2(\mathbf{I}) = \omega^3(\mathbf{I}) = \frac{1}{(I_1+I_2+I_3)^3} =: \omega(\mathbf{I}). \tag{11.5.9}$$

Thus, the system is degenerate, and Lemma 11.4.10 yields that the dynamics reduces to one-dimensional subtori, on which a periodic motion with period

$$T = \frac{2\pi}{\omega} = \frac{2\pi}{\sqrt{(-2E)^3}} \tag{11.5.10}$$

takes place. This is Kepler's Third Law, because the semimajor axis is equal to $-\frac{1}{2E}$, cf. Example 10.6.3. Let us apply the procedure of Lemma 11.4.10 to determine new action and angle variables (φ, J) adapted to this reduction. The module generated over $\mathbb{Z}$ by the characteristic frequencies is $\mathbb{Z}\omega$. We choose $\omega^1 = \omega$ as a basis and $\omega^2 - \omega^1 = 0$ and $\omega^3 - \omega^2 = 0$ as the relations. This leads to the transformation matrix

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

Then, (11.4.34) yields

$$\varphi^1 = \vartheta^1, \qquad \varphi^2 = \vartheta^2 - \vartheta^1, \qquad \varphi^3 = \vartheta^3 - \vartheta^2 \tag{11.5.11}$$

and

$$J_1 = I_1 + I_2 + I_3 = \sqrt{-\frac{1}{2E}}, \qquad J_2 = I_2 + I_3 = \Theta, \qquad J_3 = I_3 = \Phi. \tag{11.5.12}$$

The new action and angle variables φ^i, J_i are referred to as the Delaunay elements. In order to avoid degeneracy, in addition to (11.5.4) we have to assume

$$L_z < |\mathbf{L}|. \tag{11.5.13}$$

In terms of the Delaunay elements, the Hamiltonian and the frequencies are given by

$$H(\mathbf{J}) = -\frac{1}{2J_1^2}, \qquad \bar{\omega}^1(\mathbf{J}) = \frac{1}{J_1^3}, \qquad \bar{\omega}^2(\mathbf{J}) = \bar{\omega}^3(\mathbf{J}) = 0. \tag{11.5.14}$$

To clarify the geometrical meaning of the Delaunay elements, from (11.5.12) we read off that

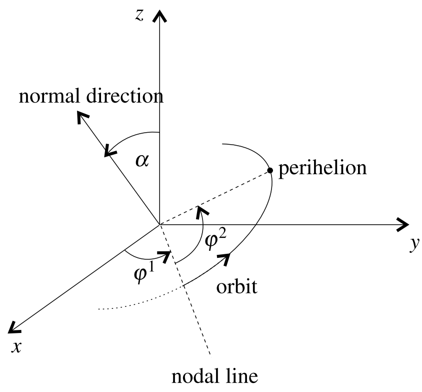
$$J_1^2 = -\frac{1}{2E}, \qquad J_2^2 = \mathbf{L}^2, \qquad J_3 = L_z. \tag{11.5.15}$$

Moreover, we use (10.6.31) to express the semimajor axis a, the semiminor axis b and the inclination angle α (the angle between the z-axis and the direction normal to the plane of motion) in terms of the J_i. This yields

$$a = \frac{p}{1-\varepsilon^2} = J_1^2, \qquad b = a\sqrt{1-\varepsilon^2} = J_1 J_2, \qquad \cos\alpha = \frac{L_z}{|\mathbf{L}|} = \frac{J_3}{J_2}.$$

The geometrical meaning of the angle coordinates φ^1 and φ^2 is illustrated in Fig. 11.2: φ^1 is the angle between the x-axis of the inertial frame and the nodal line, that is, the line of intersection of the plane of motion with the x–y-plane. Therefore, it is called the longitude of the ascending node. The angle coordinate φ^2 is the angle between the perihelion and the nodal line. The angle coordinate φ^3 coincides with

Fig. 11.2 Geometrical interpretation of the Delaunay elements



the so-called mean anomaly. For a detailed discussion of the angle coordinates φ^i, see [1] or [67].

To summarize, let us stress once again that the Delaunay elements are well defined under the conditions (11.5.4) and (11.5.13), that is, for elliptic orbits. Thus, circular orbits and orbits lying on the ecliptic plane are excluded.

Remark 11.5.3 For given $\mathbf{J}$, the invariant 3-torus $\Sigma_{\mathbf{J}}$ is defined by the equations

$$p_r^2 = 2\left(E + \frac{1}{r} - \frac{\mathbf{L}^2}{2r^2}\right) = -\frac{1}{J_1^2} + \frac{2}{r} - \frac{J_2^2}{r^2},$$

$$p_\vartheta^2 = \mathbf{L}^2 - \frac{L_z^2}{\sin^2\vartheta} = J_2^2 - \frac{J_3^2}{\sin^2\vartheta},$$

$$p_\phi = J_3.$$

On the other hand, we know that the motion reduces to the level sets of the energy-momentum mapping $\mathscr{E} = (H, L_x, L_y, L_z)$. According to (10.8.5), for negative energy E and $\mathbf{L} \neq 0$, the level sets are 2-tori. Obviously, these invariant 2-tori foliate the 3-tori $\Sigma_{\mathbf{J}}$ provided by the Arnold Theorem. This is in accordance with the Nekhoroshev Theorem, to be discussed in Sect. 11.8. Indeed, the invariant 2-tori coincide with the level sets of the $n + 1 = 4$ functionally independent constants of motion $H, \mathbf{L}^2, L_y$ and L_z which are in involution with the $n - 1 = 2$ constants of motion H and $\mathbf{L}^2$.

Example 11.5.4 (Symmetric Euler Top) Consider the Hamiltonian system $(\mathrm{T}^*\mathrm{SO}(3) \cong \mathrm{SO}(3) \times \mathfrak{so}(3), \mathrm{d}\theta, H)$ with H given by (11.1.9). The action and angle coordinates we are going to construct are the so-called Andoyer variables [7, 43]. Unlike the preceding examples, here we will proceed by first defining the coordinates geometrically and then proving that they possess the defining properties of action and angle variables. For clarity, we use the vector notation, that is, we identify $\mathfrak{so}(3)$ with $\mathbb{R}^3$ via the isomorphism (5.2.6).

Let us assume that the tensor of inertia Θ has been diagonalized and let us denote the principal moments of inertia by Θ_i, $i = 1, 2, 3$. As a local chart on the configuration space SO(3), we use the Euler angles[21] ϕ, ϑ and ψ, where $0 < \phi, \psi < 2\pi$ and $0 < \vartheta < \pi$. Let us denote the corresponding fibre coordinates in $T^*SO(3)$ (conjugate momenta) by p_ϕ, p_ϑ and p_ψ. For later use, we introduce the notation

$$A = \frac{p_\phi - p_\psi \cos\vartheta}{\sin\vartheta}, \qquad B = \frac{p_\psi - p_\phi \cos\vartheta}{\sin\vartheta}.$$

To define the Andoyer variables, we start with expressing the angular momentum **L** in terms of the Euler angles and their momenta. This can be done by rewriting the canonical 1-form on $T^*SO(3)$ in terms of the L_i, or by means of the Legendre transformation generated by the Lagrangian

$$T = \frac{1}{2}\boldsymbol{\omega} \cdot (\Theta\boldsymbol{\omega}) = \frac{1}{2}\big(\Theta_1\omega_1^2 + \Theta_2\omega_2^2 + \Theta_3\omega_3^2\big),$$

where $\boldsymbol{\omega} = \boldsymbol{\omega}(\dot\phi, \dot\vartheta, \dot\psi)$ denotes the angular velocity of the top. We shall follow the latter strategy. The Legendre transformation is given by

$$p_\phi = \frac{\partial T}{\partial\dot\phi}, \qquad p_\vartheta = \frac{\partial T}{\partial\dot\vartheta}, \qquad p_\psi = \frac{\partial T}{\partial\dot\psi}.$$

Using this and $\mathbf{L} = \Theta\boldsymbol{\omega} = (\Theta_1\omega_1, \Theta_2\omega_2, \Theta_3\omega_3)$, we obtain

$$L_1 = p_\vartheta \cos\psi + A\sin\psi, \qquad L_2 = p_\vartheta \sin\psi - A\cos\psi, \qquad L_3 = p_\psi \tag{11.5.16}$$

(Exercise 11.5.4). Now, we can construct the Andoyer variables j, g, l and J, G, L. For that purpose, recall that the elements of the inertial frame and of the body frame are denoted by $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$ and $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, respectively. Define

$$J := \mathbf{L} \cdot \mathbf{n}_3, \qquad G := \|\mathbf{L}\|, \qquad L := \mathbf{L} \cdot \mathbf{e}_3. \tag{11.5.17}$$

To define the remaining Andoyer variables j, g, l, we have to assume that **L** is neither parallel to $\mathbf{e}_3$ nor to $\mathbf{n}_3$, so that the vectors

$$\mathbf{n}' := \mathbf{n}_3 \times \mathbf{L}, \qquad \mathbf{n}'' := \mathbf{L} \times \mathbf{e}_3$$

do not vanish. Then, we can define j to be the angle between the $\mathbf{n}_1$-axis and the nodal line defined by the vector $\mathbf{n}'$, g to be the angle between the nodal lines defined by $\mathbf{n}'$ and $\mathbf{n}''$ and l to be the angle between the $\mathbf{e}_1$-axis and the nodal line defined by $\mathbf{n}''$. While G, L and l are defined intrinsically, J, g and j depend on the choice of the inertial frame. A brief calculation (Exercise 11.5.4) yields

$$J = p_\phi, \qquad G^2 = p_\vartheta^2 + p_\psi^2 + A^2, \qquad L = p_\psi \tag{11.5.18}$$

[21] Cf. Exercise 5.5.4.

and

$$\cos(j) = \frac{p_\vartheta \sin\phi - B\cos\phi}{\sqrt{p_\vartheta^2 + B^2}}, \tag{11.5.19}$$

$$\cos(g) = \frac{p_\phi p_\psi - \cos\vartheta\,(p_\vartheta^2 + p_\psi^2 + A^2)}{\sqrt{(p_\vartheta^2 + A^2)(p_\vartheta^2 + B^2)}}, \tag{11.5.20}$$

$$\cos(l) = \frac{p_\vartheta \sin\psi - A\cos\psi}{\sqrt{p_\vartheta^2 + A^2}}. \tag{11.5.21}$$

We note that the coordinates G, L, l and G, J, j parameterize the angular momentum $\mathbf{L}$ in the body frame and $a\mathbf{L}$ in the inertial frame, respectively:

$$\mathbf{L} = \big(\sqrt{G^2 - L^2}\sin(l), \sqrt{G^2 - L^2}\cos(l), L\big),$$
$$a\mathbf{L} = \big(\sqrt{G^2 - J^2}\sin(j), -\sqrt{G^2 - J^2}\cos(j), J\big).$$

By a lengthy but straightforward calculation (Exercise 11.5.4) one can prove that, in the Andoyer variables, the canonical 1-form θ on $\mathrm{T}^*\mathrm{SO}(3)$ reads

$$\theta = J\mathrm{d}j + G\mathrm{d}g + L\mathrm{d}l. \tag{11.5.22}$$

Thus, these variables provide local Darboux coordinates on $\mathrm{T}^*\mathrm{SO}(3)$. Now, consider the Hamiltonian (11.1.9). In Euler coordinates, it reads

$$H = \frac{\left(p_\vartheta \cos\psi + \frac{p_\phi - p_\psi\cos\vartheta}{\sin\vartheta}\sin\psi\right)^2}{2\Theta_1} + \frac{\left(p_\vartheta \sin\psi - \frac{p_\phi - p_\psi\cos\vartheta}{\sin\vartheta}\cos\psi\right)^2}{2\Theta_2} + \frac{p_\psi^2}{2\Theta_3} + V(\phi, \vartheta, \psi).$$

Rewriting this formula in terms of the Andoyer variables, we obtain

$$H = \left(\frac{\sin^2 l}{2\Theta_1} + \frac{\cos^2 l}{2\Theta_2}\right)\big(G^2 - L^2\big) + \frac{L^2}{2\Theta_3} + V(j, g, l) \tag{11.5.23}$$

(Exercise 11.5.4). As we know from the discussion in Example 11.1.5, in general this system is not integrable. Therefore, let us assume that $\Theta_1 = \Theta_2$ and $V = 0$, that is, let us consider a symmetric Euler top. Then, the Hamiltonian (11.5.23) boils down to

$$H = \frac{1}{2\Theta_1}\big(G^2 + \alpha L^2\big), \qquad \alpha = \frac{\Theta_1 - \Theta_3}{\Theta_3}. \tag{11.5.24}$$

Since it does not depend on the Andoyer angles j, g, l, the Andoyer variables J, G and L are constants of motion. By (11.5.22), they are functionally independent and in involution. Hence, the system is integrable. Moreover, since J, G and L are canonically conjugate to angle coordinates, the flows of their Hamiltonian vector fields are 2π-periodic. Thus, J, G and L are action variables and j, g and l can be taken as the corresponding angle variables. The characteristic frequencies are

$$\omega^j = \frac{\partial H}{\partial J} = 0, \qquad \omega^g = \frac{\partial H}{\partial G} = \frac{G}{\Theta_1}, \qquad \omega^l = \frac{\partial H}{\partial L} = \frac{\alpha L}{\Theta_1}. \tag{11.5.25}$$

Since $\omega^j = 0$, the system is degenerate.[22] As in the previous example, we have more than three constants of motion, and their level sets decompose the 3-torus into 2-tori. In the inertial frame, the constant of motion $\mathbf{L}$ is fixed in space and the figure axis performs a circular motion (regular precession) around the axis of $\mathbf{L}$ with frequency ω^g. The frequency ω^l characterizes the rotation of the top around its own symmetry axis, given by $\mathbf{e}_3$. Altogether, we find a quasiperiodic motion on the 2-torus, parameterized by the angles g and l. This can be further illustrated by the following observation: since $L_i = \Theta_i \omega_i$, at each moment of time, the vectors $\mathbf{L}$, $\mathbf{e}_3$ and $\boldsymbol{\omega}$ lie in a plane and we have

$$\boldsymbol{\omega} = \frac{1}{\Theta_1}\left(\alpha L \mathbf{e}_3 + G\frac{\mathbf{L}}{|\mathbf{L}|}\right). \tag{11.5.26}$$

The motions of rotation and precession define two cones: the first one, called the space cone, is defined by the rotation of $\boldsymbol{\omega}$ around $\mathbf{L}$. The second one, called the body cone, is defined by the rotation of $\boldsymbol{\omega}$ around the figure axis $\mathbf{e}_3$. The body cone rolls without slipping on the space cone, with the instantaneous tangent line coinciding with the axis defined by $\boldsymbol{\omega}$.

Remark 11.5.5

1. If $\mathbf{L}$ is parallel to $\mathbf{e}_3$, we have $\mathbf{L} \times \mathbf{e}_3 = 0$ and the Andoyer variables are not well defined. In this case, we have $\mathbf{L} = L_3 \mathbf{e}_3$ and hence $G^2 \equiv \mathbf{L}^2 = L^2$. Thus,

$$H = \frac{G^2}{2\Theta_3}. \tag{11.5.27}$$

   Obviously, in this case $\boldsymbol{\omega}$ and $\mathbf{e}_3$ coincide, that is, the body cone and the space cone degenerate to a ray defined by the axis of symmetry and the top rotates around this axis with frequency $\omega = \frac{\partial H}{\partial G} = G/\Theta_3$. We note that for $L = 0$ we get an analogous degeneracy. In this case we have

$$H = \frac{G^2}{2\Theta_1}. \tag{11.5.28}$$

   Thus, $\boldsymbol{\omega}$ and $\mathbf{L}$ coincide, that is, the space cone degenerates to a ray and the body cone opens to build a half space. The symmetry axis rotates in the plane

[22] For the special case $\Theta_1 = \Theta_3$ there is an additional degeneracy.

orthogonal to $\mathbf{L}$ with angular velocity ω^g and there is no rotation of the body with respect to its symmetry axis.[23] Finally, there is the trivial case defined by $G = 0$. Then, $\mathbf{L} = 0$ and the top remains at rest.

2. The case $G = L$ corresponds to the two critical points $\mathbf{L} = (0, 0, \pm l)$ of the reduced Hamiltonian with corresponding critical value given by (11.5.27). The case $L = 0$ corresponds to the critical circle defined by $\mathbf{L}(\alpha) = (l \cos\alpha, l \sin\alpha, 0)$ with critical value given by (11.5.28). These critical subsets can be viewed as intersection sets of the 2-sphere S^2_l with the ellipsoid defined by the kinetic energy. They belong to the bifurcation set of the energy momentum mapping, cf. Example 10.6.5 and Exercise 10.8.3.
3. We note that there are further coordinate singularities, defined by the condition $\mathbf{n}_3 \times \mathbf{L} = 0$. That the Andoyer variables are not defined globally suggests that the bundle (11.4.18) defined by $\mathscr{H} = (I_1, I_2, I_3)$, with the constants of motion of Example 11.1.5, is nontrivial. This will be explained in Example 11.8.10.

Let us summarize the above discussion: If one finds action and angle variables for an integrable system, then, locally, one has an explicit description of the Hamiltonian flow and of the invariant tori of this system in terms of a (quasi)periodic motion on tori. Action and angle variables usually cannot be extended to global coordinates on the subset of regular points of the phase space. Instead, there may exist topological obstructions. This will be discussed in Sect. 11.7.

Exercises

11.5.1 Prove the equations (11.5.2) and (11.5.3).

11.5.2 Verify the formulae in (11.5.5).

11.5.3 Find action and angle variables for the planar two-centre problem, cf. Example 11.1.4.
Hint. Use elliptic coordinates.

11.5.4 Prove the following formulae in Example 11.5.4: (11.5.16), (11.5.18)–(11.5.22) and (11.5.23).

11.6 Small Perturbations

In this section, we show that the description of an integrable system in terms of action and angle variables is well adapted to the study of small perturbations. Thus, let $(M, \omega, \mathscr{H})$ be a $2n$-dimensional integrable system with Hamiltonian function H_0, possessing some compact connected component of a level set of $\mathscr{H}$. By Theorem 11.4.3, in a neighbourhood of this connected component there exist action and angle variables ϑ^i, I_i and dynamics is given by

$$\dot{I}_i = 0, \qquad \dot{\vartheta}^i = \frac{\partial H_0}{\partial I_i}(\mathbf{I}) \equiv \omega^i(\mathbf{I}).$$

[23] ω^l tends to zero with L tending to zero.

Assume that the system is non-degenerate,

$$\det\left(\frac{\partial \omega^i}{\partial I_j}\right) = \det\left(\frac{\partial^2 H_0}{\partial I_i \partial I_j}\right) \neq 0. \tag{11.6.1}$$

Consider a small perturbation of this system, given by

$$H(\boldsymbol{\vartheta}, \mathbf{I}) = H_0(\mathbf{I}) + \varepsilon H_1(\boldsymbol{\vartheta}, \mathbf{I}), \tag{11.6.2}$$

where ε is a small parameter and H_1 is a Hamiltonian function which is 2π-periodic in the variables ϑ^i. The Hamilton equations for this system read

$$\dot{I}_i(t) = -\varepsilon \frac{\partial H_1}{\partial \vartheta^i}\big(\boldsymbol{\vartheta}(t), \mathbf{I}(t)\big), \qquad \dot{\vartheta}^i(t) = \omega^i\big(\mathbf{I}(t)\big) + \varepsilon \frac{\partial H_1}{\partial I_i}\big(\boldsymbol{\vartheta}(t), \mathbf{I}(t)\big). \tag{11.6.3}$$

We aim at finding an iterative canonical transformation $(\boldsymbol{\vartheta}, \mathbf{I}) \to (\bar{\boldsymbol{\vartheta}}, \bar{\mathbf{I}})$ which makes the full system integrable, order by order in ε. Let us discuss the first step of this procedure in detail: we seek the canonical transformation in terms of a generating function $S = S(\boldsymbol{\vartheta}, \bar{\mathbf{I}})$ of the second kind, that is,

$$I_i = \frac{\partial S}{\partial \vartheta^i}(\boldsymbol{\vartheta}, \bar{\mathbf{I}}), \qquad \bar{\vartheta}^i = \frac{\partial S}{\partial \bar{I}_i}(\boldsymbol{\vartheta}, \bar{\mathbf{I}}). \tag{11.6.4}$$

We make the ansatz $S(\boldsymbol{\vartheta}, \bar{\mathbf{I}}) = \vartheta^j \bar{I}_j + \varepsilon S_1(\boldsymbol{\vartheta}, \bar{\mathbf{I}})$ and require that the Hamiltonian function in the new variables, given by $\bar{H}(\bar{\boldsymbol{\vartheta}}, \bar{\mathbf{I}}) = H(\boldsymbol{\vartheta}, \mathbf{I})$, be integrable to first order in ε,

$$\bar{H}(\bar{\boldsymbol{\vartheta}}, \bar{\mathbf{I}}) = \bar{H}_0(\bar{\mathbf{I}}) + \varepsilon \bar{H}_1(\bar{\mathbf{I}}) + \varepsilon^2 \bar{H}_2(\bar{\boldsymbol{\vartheta}}, \bar{\mathbf{I}}). \tag{11.6.5}$$

By (11.6.4),

$$I_i = \bar{I}_i + \varepsilon \frac{\partial S_1}{\partial \vartheta^i}(\boldsymbol{\vartheta}, \bar{\mathbf{I}}), \qquad \vartheta^i = \bar{\vartheta}^i - \varepsilon \frac{\partial S_1}{\partial \bar{I}_i}(\boldsymbol{\vartheta}, \bar{\mathbf{I}}). \tag{11.6.6}$$

Plugging this in into the Hamiltonian (11.6.2) and expanding by powers of ε, we obtain[24]

$$\begin{aligned}
H(\boldsymbol{\vartheta}, \mathbf{I}) &= H\left(\boldsymbol{\vartheta}, \bar{\mathbf{I}} + \varepsilon \frac{\partial S_1}{\partial \boldsymbol{\vartheta}}(\boldsymbol{\vartheta}, \bar{\mathbf{I}})\right) \\
&= H_0\left(\bar{\mathbf{I}} + \varepsilon \frac{\partial S_1}{\partial \boldsymbol{\vartheta}}(\boldsymbol{\vartheta}, \bar{\mathbf{I}})\right) + \varepsilon H_1\left(\boldsymbol{\vartheta}, \bar{\mathbf{I}} + \varepsilon \frac{\partial S_1}{\partial \boldsymbol{\vartheta}}(\boldsymbol{\vartheta}, \bar{\mathbf{I}})\right) \\
&= H_0(\bar{\mathbf{I}}) + \varepsilon \frac{\partial H_0}{\partial \bar{I}_j}(\bar{\mathbf{I}}) \frac{\partial S_1}{\partial \vartheta^j}(\boldsymbol{\vartheta}, \bar{\mathbf{I}}) + \varepsilon^2 R_0(\boldsymbol{\vartheta}, \bar{\mathbf{I}}) + \varepsilon H_1(\boldsymbol{\vartheta}, \bar{\mathbf{I}}) + \varepsilon^2 R_1(\boldsymbol{\vartheta}, \bar{\mathbf{I}})
\end{aligned}$$

---

[24] Since S is a function of the variables $\boldsymbol{\vartheta}$ and $\bar{\mathbf{I}}$, we express H in these variables.

$$= H_0(\bar{\mathbf{I}}) + \varepsilon\left(\omega^j(\bar{\mathbf{I}})\frac{\partial S_1}{\partial\vartheta^j}(\boldsymbol{\vartheta},\bar{\mathbf{I}}) + H_1(\boldsymbol{\vartheta},\bar{\mathbf{I}})\right) + \varepsilon^2(R_0+R_1)(\boldsymbol{\vartheta},\bar{\mathbf{I}}) \tag{11.6.7}$$

with

$$R_0(\boldsymbol{\vartheta},\bar{\mathbf{I}}) = \frac{1}{\varepsilon^2}\left(H_0\left(\bar{\mathbf{I}}+\varepsilon\frac{\partial S_1}{\partial\boldsymbol{\vartheta}}(\boldsymbol{\vartheta},\bar{\mathbf{I}})\right) - H_0(\bar{\mathbf{I}}) - \varepsilon\frac{\partial H_0}{\partial\bar{I}_j}(\bar{\mathbf{I}})\frac{\partial S_1}{\partial\vartheta^j}(\boldsymbol{\vartheta},\bar{\mathbf{I}})\right),$$

$$R_1(\boldsymbol{\vartheta},\bar{\mathbf{I}}) = \frac{1}{\varepsilon}\left(H_1\left(\boldsymbol{\vartheta},\bar{\mathbf{I}}+\varepsilon\frac{\partial S_1}{\partial\boldsymbol{\vartheta}}(\boldsymbol{\vartheta},\bar{\mathbf{I}})\right) - H_1(\boldsymbol{\vartheta},\bar{\mathbf{I}})\right).$$

Now we choose $\bar{\mathbf{I}}_0$ such that $\omega(\bar{\mathbf{I}}_0)$ is non-resonant.[25] For $\|\bar{\mathbf{I}}-\bar{\mathbf{I}}_0\|$ being of order ε, we may replace $\omega(\bar{\mathbf{I}})$ by $\boldsymbol{\omega}_0 := \omega(\bar{\mathbf{I}}_0)$ in (11.6.7), thus producing a correction of order ε^2. Then, requiring that (11.6.7) equals (11.6.5) and comparing coefficients, we find

$$\begin{aligned} \bar{H}_0(\bar{\mathbf{I}}) &= H_0(\bar{\mathbf{I}}), \\ \bar{H}_1(\bar{\mathbf{I}}) &= H_1(\boldsymbol{\vartheta},\bar{\mathbf{I}}) + \omega_0^j\frac{\partial S_1}{\partial\vartheta^j}(\boldsymbol{\vartheta},\bar{\mathbf{I}}), \\ \bar{H}_2(\boldsymbol{\vartheta},\bar{\mathbf{I}}) &= (R_0+R_1)(\boldsymbol{\vartheta},\bar{\mathbf{I}}). \end{aligned} \tag{11.6.8}$$

To analyze the second equation in (11.6.8), we use the Fourier expansion of S_1 and H_1 with respect to the angle variables ϑ^i,

$$S_1(\boldsymbol{\vartheta},\bar{\mathbf{I}}) = \sum_{\mathbf{k}\in\mathbb{Z}^n}\tilde{S}_{\mathbf{k}}(\bar{\mathbf{I}})\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\boldsymbol{\vartheta}}, \qquad H_1(\boldsymbol{\vartheta},\bar{\mathbf{I}}) = \sum_{\mathbf{k}\in\mathbb{Z}^n}\widetilde{H}_{\mathbf{k}}(\bar{\mathbf{I}})\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\boldsymbol{\vartheta}},$$

with $\mathbf{k}\cdot\boldsymbol{\vartheta} = \sum_{i=1}^n k_i\vartheta^i$. Then, this equation reads

$$\bar{H}_1(\bar{\mathbf{I}}) = \sum\nolimits_{\mathbf{k}\in\mathbb{Z}^n}\big(\mathrm{i}(\boldsymbol{\omega}_0\cdot\mathbf{k})\tilde{S}_{\mathbf{k}}(\bar{\mathbf{I}}) + \widetilde{H}_{\mathbf{k}}(\bar{\mathbf{I}})\big)\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\boldsymbol{\vartheta}}.$$

This yields $\bar{H}_1(\mathbf{I}) = \widetilde{H}_0(\bar{\mathbf{I}})$ for $\mathbf{k}=0$ and $\mathrm{i}(\boldsymbol{\omega}_0\cdot\mathbf{k})\tilde{S}_{\mathbf{k}}(\bar{\mathbf{I}}) + \widetilde{H}_k(\bar{\mathbf{I}}) = 0$ for $\mathbf{k}\neq 0$. Thus, the solution reads

$$\bar{H}_1(\bar{\mathbf{I}}) = \widetilde{H}_0(\bar{\mathbf{I}}), \qquad S_1(\boldsymbol{\vartheta},\bar{\mathbf{I}}) = -\sum_{\mathbf{k}\in\mathbb{Z}^n\setminus\{0\}}\frac{\widetilde{H}_{\mathbf{k}}(\bar{\mathbf{I}})}{\mathrm{i}\boldsymbol{\omega}_0\cdot\mathbf{k}}\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\boldsymbol{\vartheta}}, \tag{11.6.9}$$

and the transformed Hamiltonian function takes the form

$$\bar{H}(\bar{\boldsymbol{\vartheta}},\bar{\mathbf{I}}) = H_0(\bar{\mathbf{I}}) + \varepsilon\widetilde{H}_0(\bar{\mathbf{I}}) + \varepsilon^2\bar{H}_2(\bar{\boldsymbol{\vartheta}},\bar{\mathbf{I}}). \tag{11.6.10}$$

[25] Since the system is assumed to be non-degenerate, the tori can be labelled by their frequencies, and hence such $\bar{\mathbf{I}}_0$ are dense.

If $\boldsymbol{\omega}_0 \cdot \mathbf{k}$ becomes very small, there is no hope that the Fourier series in (11.6.9) converges. This is the famous problem of small denominators, which we met already in Sect. 9.6 where we had shown that a symplectomorphism can be brought to normal form in the neighbourhood of an elliptic 4-elementary fixed point. The same was true for the Hamiltonian function in the neighbourhood of an elliptic 4-elementary critical point. We had seen that in a neighbourhood U of such a critical point the normal form part of the Hamiltonian yields a foliation of U into invariant tori, defined by a set I_j of constants of motion given by (9.6.11), and that, in this approximation, the system becomes integrable. In this context, the choice of symplectic polar coordinates on U yields action and angle variables. We had also noted there that in the cases under consideration the KAM theory applies, see Remark 9.6.8. This theory yields that non-resonant tori persist the perturbation caused by passing to the full system, provided they fulfil a strong non-resonance condition of Diophantine type. Here, we meet another situation where KAM theory is applicable. If $\boldsymbol{\omega}_0$ is strongly non-resonant, cf. (9.6.21), that is, if there exist constants $\tau > 0$ and $\gamma > 0$ such that

$$|\boldsymbol{\omega}_0 \cdot \mathbf{k}| \geq \gamma \|\mathbf{k}\|^{-\tau}, \tag{11.6.11}$$

for all $\mathbf{k} \in \mathbb{Z}^n \setminus \{0\}$, the Fourier series in (11.6.9) is convergent.[26] In this case, the generating function S exists and for tori fulfilling $\|\bar{\mathbf{I}} - \bar{\mathbf{I}}_0\| < \varepsilon$, at first order, the small perturbation yields a change of the frequency by a constant:

$$\bar{\omega}^j(\mathbf{I}) = \frac{\partial \bar{H}}{\partial \bar{I}_j} = \omega^j(\bar{\mathbf{I}}) + \varepsilon \frac{\partial \widetilde{H}_0}{\partial \bar{I}_j}(\bar{\mathbf{I}}_0).$$

Moreover, on the torus defined by $\bar{\mathbf{I}}$, at first order we have a quasiperiodic motion with frequencies $\bar{\omega}^j$:

$$\dot{\bar{I}}_j = -\frac{\partial \bar{H}}{\partial \bar{\vartheta}^j} = 0, \qquad \dot{\bar{\vartheta}}^j = \frac{\partial \bar{H}}{\partial \bar{I}_j} = \bar{\omega}^j.$$

We also note that, since $\dot{\mathbf{I}}$ is of order ε^2, the condition $\|\bar{\mathbf{I}} - \bar{\mathbf{I}}_0\| < \varepsilon$ remains valid for large times, that is, times of order ε^{-1}.

Now, the above described procedure must be iterated. In the second step, one starts with the Hamiltonian

$$\bar{H}(\bar{\boldsymbol{\vartheta}}, \bar{\mathbf{I}}) = \bar{H}_0(\bar{\mathbf{I}}) + \varepsilon^2 \bar{H}_2(\bar{\boldsymbol{\vartheta}}, \bar{\mathbf{I}}).$$

This step removes the dependence on ϑ up to order ε^4. After n steps, one arrives at order ε^{2^n}. This is the reason why the KAM procedure is said to be superconvergent. See [26] for further comments and historical remarks. Finally, one arrives at

[26] Under the assumption that H_1 is analytic on a certain domain the proof is not difficult, see e.g. [286, §3.5].

the following result.[27] If the unperturbed system is non-degenerate in the sense of (11.6.1), for sufficiently small $\varepsilon > 0$, most non-resonant invariant tori persist and are only slightly deformed, so that in the phase space of the perturbed system there exist invariant tori densely filled with quasiperiodic integral curves. These invariant tori form a majority in the sense that the Lebesgue measure of the complement of their union is small for small perturbations.

Remark 11.6.1

1. The KAM method works for tori satisfying the strong non-resonance condition (11.6.11), provided both H_0 and H_1 are of class C^l with $l > 2\tau + 2 > 2n$. This weakens the regularity assumptions made in the historical papers cited before considerably [221, 244, 259].
2. Let us make precise what it means that a majority of invariant tori persists: let $\Delta_{\gamma,\tau}$ be the set of frequencies fulfilling the infinitely many conditions contained in (11.6.11), with γ and τ kept fixed. It can be shown that $\Delta_{\gamma,\tau}$ are Cantor sets in $\mathbb{R}^n$ and that the union $\Delta_\tau = \bigcup_{\gamma>0} \Delta_{\gamma,\tau}$ has full Lebesgue measure for $\tau > n-1$, whereas $\Delta_\tau = \varnothing$ for $\tau < n-1$. That is, almost every $\boldsymbol{\omega}$ belongs to Δ_τ provided $\tau > n-1$. On the other hand, the parameter γ turns out to limit the magnitude of the perturbation parameter through the condition that $\varepsilon \ll \gamma^2$. This means that for a given perturbation one cannot vary γ arbitrarily but has to keep it large enough. Therefore, one takes a subset Ω_γ of $\Delta_{\gamma,\tau}$ of frequencies whose distance from the boundary of the set Ω of all frequencies is at least γ. These are Cantor sets whose complements in Ω have Lebesgue measure of order γ. For more details we refer to the papers of Pöschel [245, 246]. In [245], it is shown that the persisting tori form a family over the Cantor set Ω_γ which is smooth in the sense of Whitney. Thus, one may say that the perturbed system is integrable over this Cantor set.
3. The above results remain true if the assumption (11.6.1) of non-degeneracy is replaced by

$$\det \begin{bmatrix} \frac{\partial^2 H_0}{\partial I_i \partial I_j} & \frac{\partial H_0}{\partial I_j} \\ \frac{\partial H_0}{\partial I_i} & 0 \end{bmatrix} \neq 0. \tag{11.6.12}$$

A system fulfilling this condition is said to be isoenergetically non-degenerate. In this case, the persisting invariant tori form a majority, in the sense explained above, on each energy surface.

---

[27] See the list of references in Remark 9.6.8. Additionally, we draw the attention of the reader to Sect. 3.6 of the textbook [286] by Thirring, where the proof is provided for a special class of Hamiltonians. This way, the complicated KAM theory analysis becomes quite transparent.

11.7 Global Aspects. Monodromy

In the present section, we use action and angle variables to study the global structure of the submanifold $M_c^{\mathscr{H}}$ of points which belong to a compact level set component of $\mathscr{H}_r$. In particular, we derive topological obstructions to the existence of global action variables and to the existence of global action and angle variables on $M_c^{\mathscr{H}}$. The presentation below is in the spirit of Duistermaat [80].

First, we discuss the existence of global action variables. Recall from Corollary 11.4.5 that the natural projection

$$\widetilde{\mathscr{H}}_{rc} : M_c^{\mathscr{H}} \to \widetilde{V}_c^{\mathscr{H}} \tag{11.7.1}$$

has the structure of a locally trivial fibre bundle with typical fibre T^n. Since $M_c^{\mathscr{H}}$ is made up by compact level set components of $\mathscr{H}_r$, the Hamiltonian vector fields X_{H_i} restrict to complete vector fields on $M_c^{\mathscr{H}}$. Hence, according to Example 6.1.2/4, their flows define an action Ψ of $\mathbb{R}^n$ on $M_c^{\mathscr{H}}$. By the Arnold Theorem 11.3.3 and Corollary 11.4.5, $\widetilde{V}_c^{\mathscr{H}}$ is the orbit manifold of this action and $\widetilde{\mathscr{H}}_{rc}$ is the corresponding natural projection. Let I_i be action variables on $W \subset M_c^{\mathscr{H}}$ and let Ψ^I denote the action of $\mathbb{R}^n$ on W defined by the flows of the Hamiltonian vector fields X_{I_i}. By (11.4.5), the actions Ψ and Ψ^I on W are related by

$$\Psi^I_{\mathbf{t}}(m) = \Psi_{\sum_{i=1}^n t_i b_i(m)}(m), \tag{11.7.2}$$

where $b_i : W \to \mathbb{R}^n$ are smooth mappings uniquely defined by

$$X_{I_i} = b_i{}^j X_{H_j}.$$

In particular, Ψ and Ψ^I have the same orbits in W. Hence, we may view the bundle structure of $M_c^{\mathscr{H}}$ over $\widetilde{U} = \widetilde{\mathscr{H}}_{rc}(W)$ as being induced by the action Ψ^I. Since all points of W have stabilizer $2\pi\mathbb{Z}^n$ under Ψ^I, this action descends to a free action $\hat{\Psi}^I$ of $\mathrm{T}^n = \mathrm{U}(1)^n$ on W, given by

$$\hat{\Psi}^I_{(\mathrm{e}^{\mathrm{i}t_1},\ldots,\mathrm{e}^{\mathrm{i}t_n})}(m) = \Psi^I_{(t_1,\ldots,t_n)}(m).$$

Over $\widetilde{U}$, the action $\hat{\Psi}^I$ turns (11.7.1) into a principal bundle with structure group T^n. Thus, the existence of global action variables is related to the existence of a global reduction of the action Ψ to a free action of T^n. Via the relation (11.7.2), the latter is equivalent to the existence of a smooth assignment to $\Sigma \in \widetilde{V}_c^{\mathscr{H}}$ of a set of generators for the stabilizer of Ψ on Σ. Hence, we have to discuss the existence of such an assignment. For that purpose, we first construct the period bundle mentioned earlier. Consider the subset $P_{\mathscr{H}}$ of $\widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n$ made up by the pairs $(\Sigma, \mathbf{t})$ such that $\mathbf{t}$ belongs to the stabilizer of Ψ on Σ. According to Proposition 11.4.2/2, a set of action variables I_i on W defines a local frame in the vector bundle $\widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n$ by

$$\Sigma \mapsto \left\{b_1(\Sigma), \ldots, b_n(\Sigma)\right\}$$

and hence a local chart on $\widetilde{\mathscr{H}_{rc}}(W) \times \mathbb{R}^n \subset \widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n$ which maps $P_{\mathscr{H}}$ onto $\widetilde{\mathscr{H}_{rc}}(W) \times 2\pi\mathbb{Z}^n$. This shows that $P_{\mathscr{H}}$ is an embedded submanifold of $\widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n$ and, in the smooth structure induced, a locally trivial $\mathbb{Z}$-module bundle over $\widetilde{V}_c^{\mathscr{H}}$.

Definition 11.7.1 (Period bundle) The bundle $P_{\mathscr{H}}$ is called the period bundle of the integrable system $(M, \omega, \mathscr{H})$.

The bundle projection $\pi^{P_{\mathscr{H}}} : P_{\mathscr{H}} \to \widetilde{V}_c^{\mathscr{H}}$ is induced from the natural projection to the first factor in $\widetilde{V}^{\mathscr{H}} \times \mathbb{R}^n$ and local trivializations are given by the frames $\{b_i\}$ induced by action variables. Thus, a smooth global assignment to $\Sigma \in \widetilde{V}_c^{\mathscr{H}}$ of a set of generators for the stabilizer of Ψ on Σ corresponds to a global frame in $P_{\mathscr{H}}$, that is, it exists iff $P_{\mathscr{H}}$ is trivial. A criterion for the (non-)triviality of $P_{\mathscr{H}}$ is given by the so-called monodromy, which will be constructed now.

Since the fibres of $P_{\mathscr{H}}$ are discrete, the projection $\pi^{P_{\mathscr{H}}}$ is a local diffeomorphism. Hence, for every curve $\gamma : [0, 1] \to \widetilde{V}_c^{\mathscr{H}}$ and every $\mathbf{t}$ in the fibre of $P_{\mathscr{H}}$ over $\gamma(0)$, there exists a unique curve $\tilde{\gamma}$ with $\pi^{P_{\mathscr{H}}} \circ \tilde{\gamma} = \gamma$ and $\tilde{\gamma}(0) = \mathbf{t}$. This curve is called the lift of γ to $\mathbf{t}$. Via its lifts, γ induces a mapping $P_{\mathscr{H}}(\gamma(0)) \to P_{\mathscr{H}}(\gamma(1))$ by assigning to $\mathbf{t}$ the endpoint of the lift of γ to $\mathbf{t}$. This mapping is called the parallel transport along γ. The parallel transport along a composite curve $\gamma_2 \cdot \gamma_1$ is the composition of the parallel transport along γ_1 with the parallel transport along γ_2. Moreover, curves in $\widetilde{V}_c^{\mathscr{H}}$ which are homotopic with fixed endpoints generate the same parallel transport, because every smooth homotopy in $\widetilde{V}_c^{\mathscr{H}}$ can be lifted to $P_{\mathscr{H}}$ by means of a covering by local trivializations. Thus, by choosing $\Sigma_0 \in \widetilde{V}_c^{\mathscr{H}}$ and by assigning to a closed curve based at Σ_0 its parallel transport, we obtain a group homomorphism from the fundamental group $\pi_1(\widetilde{V}_c^{\mathscr{H}}, \Sigma_0)$ of $\widetilde{V}_c^{\mathscr{H}}$ based at Σ_0 to the group of transformations of the fibre of $P_{\mathscr{H}}$ over Σ_0.

To compute the parallel transport along the closed curve $s \mapsto \gamma(s)$ based at Σ_0, we cover γ by open subsets U_l which admit action variables $I_i^{(l)}$, $l = 0, \ldots, r$. For $\mathbf{t} \in P_{\mathscr{H}}(\Sigma_0)$, we decompose $\mathbf{t} = 2\pi k^i b_i^{(0)}(\Sigma_0)$ with $k^i \in \mathbb{Z}$ and $\{b_i^{(0)}\}$ denoting the local frame in $\widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n$ induced by $I^{(0)}$. As long as γ stays in U_0, the lift of γ based at $\mathbf{t} \in P_{\mathscr{H}}(\Sigma_0)$ is given by $\tilde{\gamma}(s) = 2\pi k^i b_i^{(0)}(\gamma(s))$. When entering U_1, we merely have to express each vector $b_i^{(0)}(\gamma(s))$ in terms of the frame $\{b_i^{(1)}(\gamma(s))\}$. Hence, according to Proposition 11.4.2/3, over U_1 the lift is given by

$$\tilde{\gamma}(s) = 2\pi k^i A_i{}^j b_j^{(1)}\big(\gamma(s)\big)$$

with a constant invertible integer matrix A. As a result, on the level of the coefficients k^i with respect to the basis $\{2\pi b_i^{(0)}(\Sigma_0)\}$ in $P_{\mathscr{H}}(\Sigma_0)$, the parallel transport along γ is given by the product of the integer $(n \times n)$-matrices appearing in the change of action variables along γ. In particular, it is an automorphism of the Abelian group $P_{\mathscr{H}}(\Sigma_0)$.

Definition 11.7.2 (Monodromy) The group homomorphism

$$\pi_1\big(\widetilde{V}_c^{\mathscr{H}}, \Sigma_0\big) \to \operatorname{Aut}\big(P_{\mathscr{H}}(\Sigma_0)\big)$$

which assigns to a closed curve its parallel transport is called the monodromy based at Σ_0 of the integrable system $(M, \omega, \mathscr{H})$.

As noted above, after having fixed a system of generators of $P_{\mathscr{H}}(\Sigma_0)$, the monodromy at Σ_0 may be viewed as a homomorphism from $\pi_1(\widetilde{V}_c^{\mathscr{H}}, \Sigma_0)$ to $\mathrm{GL}(n, \mathbb{Z})$. Accordingly, the image of a closed curve γ under this homomorphism is referred to as the monodromy matrix of this curve. Another choice of generators results in this homomorphism composed with an inner automorphism of $\mathrm{GL}(n, \mathbb{Z})$, as does another choice of the base point Σ_0 in the same connected component of $\widetilde{V}_c^{\mathscr{H}}$. Thus, up to conjugacy in $\mathrm{GL}(n, \mathbb{Z})$, the monodromy can depend on the connected component of $\widetilde{V}_c^{\mathscr{H}}$ only. For simplicity, in the rest of this section we adopt the assumption that $\widetilde{V}_c^{\mathscr{H}}$ or, equivalently, $M_c^{\mathscr{H}}$ is connected. In the general case, the results apply to each connected component of the bundle $\widetilde{\mathscr{H}}_{rc} : M_c^{\mathscr{H}} \to \widetilde{V}_c^{\mathscr{H}}$.

Theorem 11.7.3 (Global action variables) *For an integrable system $(M, \omega, \mathscr{H})$, the following statements are equivalent.*

1. *$M_c^{\mathscr{H}}$ admits global action variables.*
2. *The monodromy is trivial.*
3. *The period bundle is trivial.*
4. *The action Ψ on $M_c^{\mathscr{H}}$ induced by the flows of the Hamiltonian vector fields X_{H_i} can be globally reduced to a free action of T^n, that is, there exists a free T^n-action $\hat{\Psi}$ on $M_c^{\mathscr{H}}$ and a smooth mapping $\lambda : \widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n \to \mathrm{T}^n$ such that $\lambda(\Sigma, \cdot)$ is a group homomorphism for all $\Sigma \in \widetilde{V}_c^{\mathscr{H}}$ and $\Psi_{\mathbf{t}}(m) = \hat{\Psi}_{\lambda(\widetilde{\mathscr{H}}_{rc}(m), \mathbf{t})}(m)$ for all $m \in M_c^{\mathscr{H}}$.*

The action $\hat{\Psi}$ turns (11.7.1) into a principal bundle with structure group $\mathrm{T}^n = \mathrm{U}(1)^n$.

Proof 1 ⇒ 4: This has been discussed above.

4 ⇒ 3: The mapping λ induces a unique mapping $\hat{\lambda} : \widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n \to \mathbb{R}^n$ by

$$\lambda(\Sigma, \mathbf{t}) = \left(\mathrm{e}^{\mathrm{i}\hat{\lambda}^1(\Sigma, \mathbf{t})}, \ldots, \mathrm{e}^{\mathrm{i}\hat{\lambda}^n(\Sigma, \mathbf{t})}\right).$$

One can check that the mapping

$$\tilde{\lambda} : \widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n \to \widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n, \qquad \tilde{\lambda}(\Sigma, \mathbf{t}) = \left(\Sigma, \hat{\lambda}(\Sigma, \mathbf{t})\right)$$

is bijective and has a pointwise injective tangent mapping. Hence, it is a diffeomorphism. Composing the assignment of the standard basis $\{2\pi \mathbf{e}_i\}$ of $\mathbb{R}^n$ to every $\Sigma \in \widetilde{V}_c^{\mathscr{H}}$ with the inverse of this diffeomorphism, we obtain a global frame in $P_{\mathscr{H}}$.

3 ⇒ 2: This is obvious.

2 ⇒ 1: Since the parallel transport along an arbitrary closed curve in $\widetilde{V}_c^{\mathscr{H}}$ is given by the product of the transformation matrices for the changes of the action variables along that curve, vanishing of the monodromy implies that any such product yields the unit matrix. Thus, for a given covering of $M_c^{\mathscr{H}}$ by systems of local

action variables, by choosing one such system and redefining all others by successively transforming them with the inverse transformation matrices and subtracting the constant shifts, we obtain a system of global action variables, cf. point 3 of Proposition 11.4.2. □

Next, we discuss the existence of global action and angle variables. This discussion involves the Chern class of a principal bundle with structure group $\mathrm{T}^n = \mathrm{U}(1)^n$, for which we refer to the standard literature, see e.g. [166] or [279].[28] First, we investigate under which conditions the locally trivial fibre bundle $\widetilde{\mathscr{H}_{rc}} : M_c^{\mathscr{H}} \to \widetilde{V}_c^{\mathscr{H}}$ is globally trivial. Recall that $M_c^{\mathscr{H}}$ is assumed to be connected.

Proposition 11.7.4 *The bundle* $\widetilde{\mathscr{H}_{rc}} : M_c^{\mathscr{H}} \to \widetilde{V}_c^{\mathscr{H}}$ *is globally trivial iff both the monodromy is trivial and the first Chern class of the principal* T^n*-bundle so induced vanishes.*

Proof If the monodromy is trivial and the Chern class of the induced principal T^n-bundle vanishes, this bundle must be trivial, because principal T^n-bundles are classified up to vertical bundle isomorphisms by their first Chern class. Conversely, if (11.7.1) is trivial, it suffices to show that the monodromy is trivial, because then, this bundle is a principal T^n-bundle and since it is trivial, the first Chern class vanishes. Thus, let $\chi : M_c^{\mathscr{H}} \to \widetilde{V}_c^{\mathscr{H}} \times \mathrm{T}^n$ be a trivialization of (11.7.1). Realize T^n as the n-fold product of the complex unit circle and consider the local diffeomorphism

$$\psi : \widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n \to \widetilde{V}_c^{\mathscr{H}} \times \mathrm{T}^n, \qquad \psi(\Sigma, \mathbf{t}) := \chi \circ \Psi_{\mathbf{t}} \circ \chi^{-1}\big(\Sigma, (1, \ldots, 1)\big).$$

Define a mapping $\varphi : \widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n \to \widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n$ as follows. For $(\Sigma, \mathbf{t}) \in \widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n$, the curve $s \mapsto \psi(\Sigma, s\mathbf{t})$ in $\widetilde{V}_c^{\mathscr{H}} \times \mathrm{T}^n$ has a unique lift to $\widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n$ with respect to the covering

$$\rho : \widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n \to \widetilde{V}_c^{\mathscr{H}} \times \mathrm{T}^n, \qquad \rho(\Sigma, \mathbf{t}) := \big(\Sigma, \big(\mathrm{e}^{\mathrm{i}t_1}, \ldots, \mathrm{e}^{\mathrm{i}t_n}\big)\big).$$

Assign to $(\Sigma, \mathbf{t})$ the endpoint of the lifted curve. Then, $\psi = \rho \circ \varphi$, hence φ is a local diffeomorphism. Since it is bijective, it is a diffeomorphism. Hence, it is a vector bundle automorphism of the trivial vector bundle $\widetilde{V}_c^{\mathscr{H}} \times \mathbb{R}^n$. Since it maps the period bundle $P_{\mathscr{H}}$ onto the trivial subbundle $\widetilde{V}_c^{\mathscr{H}} \times 2\pi\mathbb{Z}^n$, Theorem 11.7.3/3 yields that the monodromy is trivial, indeed. □

Theorem 11.7.5 (Global action and angle variables) *Let* $(M, \omega, \mathscr{H})$ *be an integrable system. Then,* $M_c^{\mathscr{H}}$ *admits global action and angle variables iff the monodromy is trivial, the first Chern class of the principal* T^n*-bundle* $\widetilde{\mathscr{H}_{rc}} : M_c^{\mathscr{H}} \to \widetilde{V}_c^{\mathscr{H}}$ *so induced vanishes and* ω *is exact.*

[28] Characteristic classes will be discussed in detail in volume II of this book.

Proof If $M_c^{\mathscr{H}}$ admits global action and angle variables ϑ^i, I_i, the bundle $\widetilde{\mathscr{H}}_{rc}$: $M_c^{\mathscr{H}} \to \widetilde{V}_c^{\mathscr{H}}$ is globally trivial and the triviality of the monodromy and the Chern class follow from Proposition 11.7.4. Moreover, ω is exact with potential $I_i \mathrm{d}\vartheta^i$. Conversely, if the monodromy and the Chern class are trivial, Theorem 11.7.3 yields the existence of global action variables I_i and Proposition 11.7.4 implies that the bundle $\widetilde{\mathscr{H}}_{rc} : M_c^{\mathscr{H}} \to \widetilde{V}_c^{\mathscr{H}}$ admits a global section $s : \widetilde{V}_c^{\mathscr{H}} \to M_c^{\mathscr{H}}$. Since ω is exact, so is the 2-form $s^*\omega$ on $\widetilde{V}_c^{\mathscr{H}}$. Choose a potential and expand it with respect to the global frame $\{\mathrm{d}I_i\}$ (with the I_i viewed as functions on $\widetilde{V}_c^{\mathscr{H}}$). Let ϑ^i be the expansion coefficients. Define a new section $\tilde{s} : \widetilde{V}_c^{\mathscr{H}} \to M_c^{\mathscr{H}}$ by $\tilde{s}(\Sigma) := \Psi^I_{\vartheta(\Sigma)}(s(\Sigma))$. A brief calculation shows that $\tilde{s}$ is Lagrange. It follows that the functions ϑ^i on $M_c^{\mathscr{H}}$ defined by

$$\vartheta^i\left(\Psi^I_{\mathbf{x}}\left(\tilde{s}(\Sigma)\right)\right) = x^i$$

combine with the functions I_i to global action and angle variables on $M_c^{\mathscr{H}}$. □

Remark 11.7.6

1. According to the Hurewicz Theorem, see e.g. [55] or [199], if the first and second homotopy groups of $M_c^{\mathscr{H}}$ are trivial, so are the first and second de Rham cohomology groups. This implies that the monodromy and the first Chern class are trivial and that the 2-form ω is exact. Hence, Theorem 11.7.5 yields that triviality of the first and second homotopy groups of $M_c^{\mathscr{H}}$ is a sufficient condition for the existence of global action and angle variables, a result which belongs to Nekhoroshev [228].
2. Let $\{(W_\alpha, (\vartheta_\alpha, I_\alpha))\}$ be an atlas of Darboux charts on $M_c^{\mathscr{H}}$ built from action and angle variables. The equations $\vartheta_\alpha = 0$ define Lagrangian sections s_α over $U_\alpha = \widetilde{\mathscr{H}}_{rc}(W_\alpha)$ in the bundle (11.7.1). For each pair (α, β) with $W_\alpha \cap W_\beta \neq \varnothing$, there exists a closed 1-form $\lambda_{\alpha\beta}$ on $U_\alpha \cap U_\beta$ such that

$$s_\alpha(\Sigma) = \phi_{\lambda_{\alpha\beta}(\Sigma)} s_\beta(\Sigma),$$

   where ϕ denotes the natural fibrewise action[29] of $\mathrm{T}^*\widetilde{V}_c^{\mathscr{H}}$ on the fibres of the bundle (11.7.1), that is, on the level set components of $\widetilde{\mathscr{H}}_{rc}$. Note that $\lambda_{\alpha\beta}$ is not unique, for the following reason. By virtue of the global frame $\{\hat{\mathscr{H}}_r^* \mathrm{d}x^i\}$ in $\mathrm{T}^*\widetilde{V}_c^{\mathscr{H}}$, where x^i denote the standard coordinates on $\mathbb{R}^n$, one can identify the period bundle $P_{\mathscr{H}}$ with a $\mathbb{Z}$-module subbundle of $\mathrm{T}^*\widetilde{V}_c^{\mathscr{H}}$. It turns out that this subbundle is the stabilizer of ϕ. Thus, to make $\lambda_{\alpha\beta}$ unique, one has to view it as a section in the quotient bundle $\mathrm{T}^*\widetilde{V}_c^{\mathscr{H}}/P_{\mathscr{H}}$, which is a locally trivial T^n-bundle. Note that this bundle is locally isomorphic to (11.7.1). In more detail, the action variables I_α define an isomorphism

$$W_\alpha \cong \mathrm{T}^*(U_\alpha)/P_{\mathscr{H}}(U_\alpha).$$

[29] Defined by (8.6.15).

One can check that the system $\lambda := \{\lambda_{\alpha\beta}\}$ defines a 1-cocycle on $\widetilde{V}_c^{\mathscr{H}}$ with values in the bundle $\mathrm{T}^*\widetilde{V}_c^{\mathscr{H}}/P_{\mathscr{H}}$ in the sense of Čech, that is, an element of the first Čech cohomology,

$$[\lambda] \in H^1_{\check{c}}\big(\widetilde{V}_c^{\mathscr{H}}, Z\big(\mathrm{T}^*\widetilde{V}_c^{\mathscr{H}}/P_{\mathscr{H}}\big)\big).$$

This cohomology element is called the Lagrange class of the bundle (11.7.1). The Lagrange class is related to the Chern class mentioned above as follows. There exists a natural coboundary operator

$$\delta : H^1\big(\widetilde{V}_c^{\mathscr{H}}, Z\big(\mathrm{T}^*\widetilde{V}_c^{\mathscr{H}}/P_{\mathscr{H}}\big)\big) \to H^2\big(\widetilde{V}_c^{\mathscr{H}}, P_{\mathscr{H}}\big),$$

which maps the Lagrange class to a cohomology class $\delta([\lambda]) \in H^2(\widetilde{V}_c^{\mathscr{H}}, P_{\mathscr{H}})$, called the Chern class of the bundle (11.7.1). If the monodromy is trivial, then the Chern class so defined coincides with the ordinary Chern class of the corresponding principal T^n-bundle. One can show that the Lagrange class characterizes the bundle (11.7.1) up to bundle isomorphisms which are symplectomorphisms, that is, for a given Chern class, there exists a family of isomorphic bundles which are distinguished as symplectic manifolds by their Lagrange class. For details, we refer to Duistermaat [80], Dazord and Delzant [71], and Zung [317, 318].

Let us add that, on $M_c^{\mathscr{H}}$, the action ϕ of $\mathrm{T}^*\widetilde{V}_c^{\mathscr{H}}$ can be used to replace the action Ψ of $\mathbb{R}^n$ induced by the Hamiltonian vector fields X_{H_i}. This shifts the focus from the constants of motion in involution H_i, whose choice contains some arbitrariness, to the Lagrangian torus foliation they define, which is more geometric in nature. The description of integrable systems in terms of Lagrangian torus foliations generalizes to the situation where the foliation is only locally generated by constants of motion in involution. For an introduction to symplectic toric manifolds, we refer to [28].

3. Recently, there have been successful attempts to include the singular points of $\mathscr{H}$, with the ultimate goal of understanding the global topological structure of the full system, see Zung [317, 318].

Example 11.7.7 (Spherical pendulum) Consider the spherical pendulum, with the phase space TS^2, realized as the level set of the mapping

$$f = (f_1, f_2)\colon T\mathbb{R}^3 \to \mathbb{R}^2, \qquad f_1(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\|^2 - 1, \qquad f_2(\mathbf{x}, \mathbf{y}) = \mathbf{x}\cdot\mathbf{y},$$

and with the Hamiltonian

$$H(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\|\mathbf{y}\|^2 + x_3. \tag{11.7.3}$$

According to Example 11.1.8, as constants of motion in involution we can choose $H_1 = H$ and $H_2 = J$, where

$$J(\mathbf{x}, \mathbf{y}) = x_1 y_2 - x_2 y_1 \tag{11.7.4}$$

is the momentum mapping of the SO(2)-symmetry given by rotations about the x_3-axis. Thus, $\mathscr{H}$ coincides with the energy-momentum mapping $\mathscr{E}$ studied in Example 10.8.10. In what follows, we will keep the notation $\mathscr{E}$ and we will use results from that example without further notice. The critical points of $\mathscr{E}$ are $m_\pm = (\mathbf{x}_\pm, 0)$ with $\mathbf{x}_\pm = (0, 0, \pm 1)$. Correspondingly, the critical values are $(h, j) = (\pm 1, 0)$. The level set of $(-1, 0)$ consists of m_- alone, whereas the level set of $(1, 0)$ consists of m_+ and an open cylinder over S^1. All the other level sets are compact and connected. Hence,

$$M^{\mathscr{H}} = \mathrm{TS}^2 \setminus \{m_+, m_-\}, \qquad \widetilde{V}^{\mathscr{H}} = V^{\mathscr{H}} = \mathscr{E}\big(\mathrm{TS}^2\big) \setminus \big\{(-1, 0)\big\}$$

and

$$M_c^{\mathscr{H}} = \mathrm{TS}^2 \setminus \mathscr{E}^{-1}\big(\big\{(\pm 1, 0)\big\}\big), \qquad \widetilde{V}_c^{\mathscr{H}} = \mathscr{E}\big(\mathrm{TS}^2\big) \setminus \big\{(\pm 1, 0)\big\},$$

cf. Fig. 10.8. In particular, $\widetilde{V}_c^{\mathscr{H}}$ coincides with the set of regular values of $\mathscr{E}$. We will show that for the spherical pendulum, both the bundle (11.7.1) and the monodromy are non-trivial and that, therefore, neither global action and angle variables nor global action variables exist.

The argument for the first assertion requires knowledge about the topology of the level sets of H. Since H has the same critical points as $\mathscr{E}$, it is a Morse function on $\mathrm{T}^*\mathrm{S}^2$. Hence, we can apply the results of Sect. 8.9. To determine the Morse indices, we observe that

$$\mathrm{T}_{m_\pm}\big(\mathrm{TS}^2\big) = \ker \mathrm{d} f(m_\pm) = \ker \begin{pmatrix} 0 & 0 & \pm 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \pm 1 \end{pmatrix},$$

so that $\mathrm{T}_{m_\pm}(\mathrm{TS}^2)$ is spanned by the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4$ and $\mathbf{e}_5$ of $\mathbb{R}^6$, and we calculate[30] the Hessian of X_H at $m_\pm$ in that basis:

$$\mathrm{Hess}_{m_\pm}(X_H) = \begin{pmatrix} \mp 1 & 0 & 0 & 0 \\ 0 & \mp 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Thus, the Morse index is 2 for m_+ and 0 for m_-, that is, m_+ is a non-degenerate saddle point and m_- is a non-degenerate minimum.

Proposition 11.7.8 *The energy surfaces of the spherical pendulum have the following topological structure.*

1. *For* $h = -1$ *we have* $H^{-1}(h) = \{m_-\}$.
2. *For* $-1 < h < 1$, *the energy surface* $H^{-1}(h)$ *is diffeomorphic to the* 3*-sphere* S^3.
3. *For* $h = 1$, *the energy surface* $H^{-1}(h)$ *is homeomorphic to the unit tangent* 1*-sphere bundle* $\mathrm{T}_1\mathrm{S}^2$ *with the fibre over* $\mathbf{x}_+$ *contracted to a point.*
4. *For* $h > 1$, *the energy surface* $H^{-1}(h)$ *is diffeomorphic to* $\mathrm{T}_1\mathrm{S}^2 \cong \mathrm{SO}(3) \cong \mathbb{R}\mathrm{P}^3$.

[30] This was posed as Exercise 10.8.7.

Proof 1. This is obvious.

2. The Morse Lemma 8.9.4 provides local coordinates $\xi_1, \dots, \xi_4$ in a neighbourhood of the critical point m_- such that

$$H(\xi_1, \xi_2, \xi_3, \xi_4) = -1 + \frac{1}{2}\big(\xi_1^2 + \cdots + \xi_4^2\big).$$

Thus, for values h close to -1, the level set $H^{-1}(h)$ is diffeomorphic to the 3-sphere S^3. By the Morse-Isotopy Lemma 8.9.6, this remains true for all $h < 1$.

3. and 4. In case $h > 1$, for all $(\mathbf{x}, \mathbf{y}) \in \mathrm{TS}^2$, we have $x_3 < h$. Consequently, $\mathbf{y}$ belongs to the 1-sphere in $\mathrm{T}_{\mathbf{x}}\mathrm{S}^2$ defined by

$$\frac{1}{2}\|\mathbf{y}\|^2 = h - x_3.$$

This proves point 4. If $h \to 1$, the 1-spheres over the points $\mathbf{x} \neq \mathbf{x}_+$ persist, whereas the 1-sphere over $\mathbf{x}_+$ degenerates to a single point. □

Corollary 11.7.9 *For the spherical pendulum, the bundle* $\widetilde{\mathscr{H}}_{rc} : M_c^{\mathscr{H}} \to \widetilde{V}_c^{\mathscr{H}}$ *is nontrivial. In particular, global action and angle variables cannot exist.*

Proof We show[31] that there exists a closed curve γ in $\widetilde{V}_c^{\mathscr{H}}$ such that the subbundle $\mathscr{E}^{-1}(\gamma) \to \gamma$ is non-trivial. Choose γ to wind once around the critical value $(1, 0)$ and to touch the two boundary curves of $\widetilde{V}_c^{\mathscr{H}}$ in one point each, see Fig. 11.3(a). At the points where it touches the boundary, cut it into two closed pieces γ_0 and γ_1. Deform these pieces diffeomorphically, and with endpoints on the boundary, into horizontal line segments α_0 and α_1 running through all points (h_0, j) and (h_1, j) in $\widetilde{V}_c^{\mathscr{H}}$, respectively, where $h_0 < 1$ and $h_1 > 1$ are fixed energy values. If the bundle $\mathscr{E}^{-1}(\gamma) \to \gamma$ was trivial, that is, if it was isomorphic to the product bundle $\mathrm{T}^2 \times \gamma$, the manifolds $\mathscr{E}^{-1}(\gamma_0)$ and $\mathscr{E}^{-1}(\gamma_1)$, and hence the manifolds $\mathscr{E}^{-1}(\alpha_0)$ and $\mathscr{E}^{-1}(\alpha_1)$, would be homeomorphic. However, $\mathscr{E}^{-1}(\alpha_0) = H^{-1}(h_0) \cong \mathrm{S}^3$ and $\mathscr{E}^{-1}(\alpha_1) = H^{-1}(h_1) \cong \mathrm{SO}(3)$ are not homeomorphic. □

Next, we calculate the monodromy. The fundamental group $\pi_1(\widetilde{V}_c^{\mathscr{H}})$ is generated by any closed curve γ which winds once around the critical value $(1, 0)$. We have to find the parallel transport of a basis of the period lattice along γ. For that purpose, we have to construct action variables. In spherical coordinates ϕ and ϑ on the unit 2-sphere, given by

$$x_1 = \cos\phi \sin\vartheta, \qquad x_2 = \sin\phi \sin\vartheta, \qquad x_3 = \cos\vartheta,$$

and their canonical conjugate momenta p_ϕ and p_ϑ, the momentum mapping (11.7.4) and the Hamiltonian (11.7.3) take the form

$$J = p_\phi, \qquad H = \frac{p_\vartheta^2}{2} + \frac{p_\phi^2}{2\sin^2\vartheta} + \cos\vartheta.$$

[31] The argument belongs to Cushman.

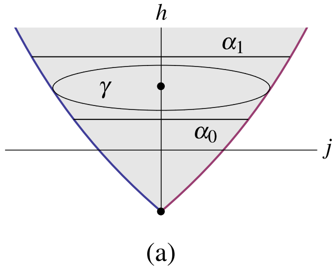


Fig. 11.3 The curves in $\widetilde{V}_c^{\mathscr{H}}$ used in the prof of Corollary 11.7.9 (**a**) and in the computation of the monodromy (**b**)

Thus, the Hamilton equations read

$$\dot{p}_\phi = 0, \qquad \dot{p}_\vartheta = \frac{p_\phi^2 \cos\vartheta}{\sin^3\vartheta} + \sin\vartheta, \qquad \dot{\phi} = \frac{p_\phi}{\sin^2\vartheta}, \qquad \dot{\vartheta} = p_\vartheta. \tag{11.7.5}$$

For a given regular value (h, j), we choose the fundamental cycles γ_1 and γ_2 as the integral curve of X_J and X_H, respectively, through a chosen point of $\mathscr{E}^{-1}(h, j)$. According to (10.8.14), γ_1 is given by $\vartheta = \text{const}$ and it is parameterized by the variable ϕ. Since H is a constant of motion, along γ_2, we have

$$p_\vartheta^2 = 2(h - \cos\vartheta) - \frac{j^2}{\sin^2\vartheta}.$$

Now, the action variables are given by

$$I_1 = \int_{\gamma_1} p_\phi d\phi = 2\pi j,$$

$$I_2 = \int_{\gamma_2} p_\vartheta d\vartheta = 2\int_{\vartheta_-}^{\vartheta_+} \sqrt{2(h - \cos\vartheta) - \frac{j^2}{\sin^2\vartheta}} d\vartheta,$$

with $\vartheta_\pm$ denoting the solutions of the equation $2(h - \cos\vartheta) - j^2 \sin^{-2}\vartheta = 0$. Note that the integral is of elliptic type and cannot be solved in terms of elementary functions. For the corresponding transformation of frames

$$X_{I_i} = b_i{}^k X_{H_k},$$

where $H_1 = H$ and $H_2 = J$, we find $b_i{}^k(h, j) = \frac{\partial I_i}{\partial H_k}(h, j)$, that is,

$$b_1{}^1(h, j) = 0, \qquad b_1{}^2(h, j) = 2\pi$$

and, by (11.7.5),

$$b_2{}^1(h, j) = 2\int_{\vartheta_-}^{\vartheta_+} \frac{\mathrm{d}\vartheta}{\sqrt{2(h - \cos\vartheta) - \frac{j^2}{\sin^2\vartheta}}} = 2\int_{\vartheta_-}^{\vartheta_+} \frac{\mathrm{d}\vartheta}{\dot{\vartheta}} = 2\int_{\vartheta_-}^{\vartheta_+} \mathrm{d}t$$

$$b_2{}^2(h,j) = -2j\int_{\vartheta_-}^{\vartheta_+} \frac{\mathrm{d}\vartheta}{\sin^2\vartheta\sqrt{2(h-\cos\vartheta)-\frac{j^2}{\sin^2\vartheta}}} = -2\int_{\vartheta_-}^{\vartheta_+}\frac{\dot\phi \mathrm{d}\vartheta}{\dot\vartheta}$$
$$= -2\int_{\vartheta_-}^{\vartheta_+} \mathrm{d}\phi.$$

Thus, $T := b_2{}^1(h,j)$ is the time needed for running through one period of the reduced dynamics and $\Delta\phi := b_2{}^2(h,j)$ is the corresponding increase of ϕ. T is called the time of return and $\Delta\phi$ is called the rotation number. A careful analysis, see [69], yields the following result:

(a) T is a uniquely defined real analytic function on $V_{\mathscr{H}}$.
(b) $\Delta\phi$ is a locally unique real analytic function on $V_{\mathscr{H}}$ fulfilling

$$\lim_{j\to 0}\Delta\phi = \begin{cases} -\pi & \text{for } -1 < h < 1 \\ -2\pi & \text{for } h > 1. \end{cases} \tag{11.7.6}$$

This can be understood heuristically by looking at the effective potential of the reduced dynamics.

Now, we choose γ to run along the edges of the square with corners

$$(h_0, -j_0), \qquad (h_1, -j_0), \qquad (h_1, j_0), \qquad (h_0, j_0),$$

where $h_0 < 1$, $h_1 > 1$ and $j_0 > 0$, see Fig. 11.3(b). Obviously, the parallel transport of $b_1(h_0, -j_0)$ along this curve is trivial and the component $b_2{}^1$ does not change as well. Let us calculate the change in the component $b_2{}^2(h_0, -j_0) = \Delta\phi$: we can choose j_0 arbitrarily small. In the limit $j_0 \to 0$, using (11.7.6) and the obvious relation $\Delta\phi(h,j) = -\Delta\phi(h,-j)$, for the first line segment we obtain

$$\lim_{j_0\to 0}\big(\Delta\phi(h_1,-j_0) - \Delta\phi(h_0,-j_0)\big) = 2\pi - \pi = \pi.$$

For the second and the fourth line segments there is no contribution and for the third segment we obtain again π. Thus, the parallel transport along γ yields the following transformation of the frame at $(h_0, -j_0)$:

$$b_1 \mapsto b_1, \qquad b_2 \mapsto b_2 + b_1.$$

Thus, the monodromy matrix is given by

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

In particular, the monodromy is non-trivial, so that the spherical pendulum does not admit global action variables either.

Remark 11.7.10 Another example with nontrivial monodromy is provided by the Lagrange top, see [69]. We stress that nontrivial monodromy leads to interesting quantum effects. These are studied for example in molecular physics, see [113] and [83] and the references therein.

11.8 Non-commutative Integrability

As already noted at the end of Sect. 11.4, the occurrence of degeneracies is often related to the fact that a system possesses more than n functionally independent[32] constants of motion.[33] This situation has been considered in an early paper by Nekhoroshev [228]. Below we give a proof of this result. Moreover, we prove the Mishchenko-Fomenko Theorem which applies when the constants of motion close under the Poisson bracket. For both of these results, one needs the following classical theorem, which generalizes the Liouville Theorem 11.3.1.

Theorem 11.8.1 (Carathéodory-Jacobi-Lie) *Let (M, ω) be a symplectic manifold of dimension $2n$ and let $f_1, \dots, f_l$, $l \leq n$, be independent smooth functions in involution. Then, for every $m \in M$, there exist $(2n - l)$ smooth functions $f_{l+1}, \dots, f_{2n}$ on an open neighbourhood of m such that $f_1, \dots, f_{2n}$ are Darboux coordinates,*

$$\omega = \mathrm{d} f_1 \wedge \mathrm{d} f_{n+1} + \cdots + \mathrm{d} f_n \wedge \mathrm{d} f_{2n}.$$

Proof We show that on some neighbourhood of m in M, the family $\{f_1, \dots, f_l\}$ may be extended to a family $\{f_1, \dots, f_n\}$ of independent functions in involution. Then, the assertion follows from the Liouville Theorem 11.3.1.

Since $\omega(X_{f_i}, X_{f_j}) = \{f_i, f_j\} = 0$ for all $i, j = 1, \dots, l$ and since the differentials $\mathrm{d} f_1, \dots, \mathrm{d} f_l$ are linearly independent, they generate an isotropic subbundle E of $\mathrm{T}^* M$. Let $\alpha = \mathrm{d} f_1 \wedge \cdots \wedge \mathrm{d} f_l$ and consider the closed $(2n - l)$-form

$$\beta = \alpha \wedge \omega^{(n-l)}.$$

Obviously, all the characteristic subspaces $\ker \beta_m$ have the same dimension. Hence, $\ker \beta$ coincides with the characteristic distribution[34] D^β of β which by Proposition 4.2.20 is integrable. Thus, the Frobenius Theorem implies that there exists an adapted chart (U, κ) such that D^β is spanned over U by the l vector fields $\partial_{2n-l+1}, \dots, \partial_{2n}$ and the annihilator $(D^\beta)^0$ is spanned over U by the 1-forms $\mathrm{d}\kappa^1, \dots, \mathrm{d}\kappa^{2n-l}$. It is not hard to see that $(D^\beta)^0$ is Π-orthogonal to E, where Π denotes the Poisson bivector defined by ω (Exercise 11.8.1). Thus, every function κ^i, $i = 1, \dots, 2n - l$, Poisson-commutes with every function f_j. Choosing one of them and denoting it by f_{l+1}, we end up with $(l + 1)$ independent functions in involution. This procedure can be iterated until we obtain n functions in involution. □

[32] For convenience, throughout this section we will assume the constants of motion under consideration to be functionally independent on the whole of M.

[33] One usually speaks of non-commutative (Liouville) integrability or of superintegrability.

[34] Cf. Definition 4.2.18.

Remark 11.8.2 The Carathéodory-Jacobi-Lie Theorem can be further generalized as follows. Let $f_1, \ldots, f_l$ and $h_1, \ldots, h_k$ be independent smooth functions fulfilling

$$\{f_i, f_j\} = 0, \qquad \{h_r, h_s\} = 0, \qquad \{f_i, h_r\} = \delta_{ir}.$$

Then, locally, there exist smooth functions $f_{l+1}, \ldots, f_n$ and $h_{k+1}, \ldots, h_n$ complementing the above functions to Darboux coordinates. This result is usually referred to as the Cartan-Lie Theorem, see [181, §III.13] for a proof.

Theorem 11.8.3 (Nekhoroshev) *Let there be given $(n+k)$ independent functions $H_1, \ldots, H_{n+k}$ on a $2n$-dimensional symplectic manifold (M, ω). Assume that all of these functions are in involution with the first $(n-k)$ functions. Let Σ be a compact level set component of the mapping $\mathscr{H} = (H_1, \ldots, H_{n+k}) : M \to \mathbb{R}^{n+k}$. Then, Σ is isotropic and diffeomorphic to an $(n-k)$-dimensional torus. Moreover, there exist coordinates $\vartheta^1, \ldots, \vartheta^{n-k}, q^1, \ldots, q^k, I_1, \ldots, I_{n-k}$ and $p_1, \ldots, p_k$ on an open neighbourhood U of Σ such that*

(a) *the symplectic form is given by $\omega = \sum_{i=1}^{n-k} dI_i \wedge d\vartheta^i + \sum_{j=1}^{k} dp_j \wedge dq^j$,*
(b) *the flows of the Hamiltonian vector fields X_{I_i} on U are complete and 2π-periodic,*
(c) *the coordinates I_i can be written as functions of $H_1, \ldots, H_{n-k}$ and the coordinates p_j and q^j can be written as functions of $H_1, \ldots, H_{n+k}$.*

Due to the periodicity of the flows of the Hamiltonian vector fields X_{I_i}, the coordinates I_i and ϑ^i are referred to as generalized action and angle variables. For $k = 0$, we obtain the Arnold Theorem 11.3.3 and the existence theorem 11.4.3 for action and angle variables as a special case. Thus, the proof of the Nekhoroshev Theorem, to be given below, yields an alternative existence proof for action and angle variables in the case of an ordinary integrable system, cf. Remark 11.4.6.

Proof Since

$$X_{H_i}(H_j) = \{H_i, H_j\} = 0, \qquad i = 1, \ldots, n-k, \quad j = 1, \ldots, n+k,$$

the Hamiltonian vector fields $X_{H_1}, \ldots, X_{H_{n-k}}$ are tangent to all level set components of $\mathscr{H}$. For dimensional reasons, they span the tangent spaces and thus, every level set component is isotropic. Since Σ is compact, the restrictions of $X_{H_1}, \ldots, X_{H_{n-k}}$ to Σ are complete. Hence, their flows define an action of $\mathbb{R}^{n-k}$ on Σ. Since they are linearly independent and since Σ has dimension $2n - (n+k) = n-k$, the orbit mapping of any point is a local diffeomorphism. By the same argument as in the proof of the Arnold Theorem we conclude that the action is transitive. Since Σ is compact, the common stabilizer of the action is an integer lattice in $\mathbb{R}^{n-k}$, generated by elements $\mathbf{b}_1, \ldots, \mathbf{b}_{n-k} \in \mathbb{R}^{n-k}$. This implies that Σ is diffeomorphic to T^{n-k} and that the vector fields

$$X_i := b_i{}^j X_{H_j} \tag{11.8.1}$$

on Σ have 2π-periodic flows. Moreover, since the constants of motion $(H_1, \dots, H_{n-k})$ commute, we have

$$\omega(X_{H_i}, X_{H_j}) = \{H_i, H_j\} = 0, \quad i, j = 1, \dots, n-k,$$

that is, Σ is isotropic. Using the Tubular Neighbourhood Theorem for the submanifold Σ and the Implicit Function Theorem, one can show that there exists a diffeomeorphism Φ from a neighbourhood W of Σ onto $B \times \mathrm{T}^{n-k}$, where $B \subset \mathbb{R}^{n+k}$ is some open ball, such that $\mathrm{pr}_B \circ \Phi = \mathscr{H}$. We conclude that all level set components in W are diffeomorphic to $\Sigma \cong \mathrm{T}^{n-k}$ and by the same arguments we obtain 2π-periodic vector fields X_i on each level set component, defined by (11.8.1). We choose the generators $\mathbf{b}_i$ so that, for every i, the integral curves of X_i on different level set components are homotopic in W.

Now, let γ_i be the image of the integral curve of X_i through $m_0 \in \Sigma$ and let $\gamma_i(m)$ be a closed curve through $m \in W$ which is homotopic to γ_i and which is contained in the torus $\Sigma(m)$ through m. The curves $\gamma_1(m), \dots, \gamma_{n-k}(m)$ form a system of fundamental cycles in this torus. Next, using the Poincaré Lemma, we want to show that on W there exists a potential form τ of ω. For that purpose, it is enough to note that ω vanishes on every 2-cycle of W. This is the case, indeed, because via the diffeomorphism Φ every 2-cycle on W is homotopic to a 2-cycle on Σ. Then, the isotropy of Σ implies the assertion. Thus, let us choose a potential τ and let us define

$$I_i(m) := \frac{1}{2\pi} \int_{\gamma_i(m)} \tau. \tag{11.8.2}$$

These functions are well-defined and smooth,[35] because all tori foliating the neighbourhood U are isotropic and thus, the I_i depend on the homotopy class of $\gamma_i(m)$ only.

Now, we will show that the functions I_i are in involution and that the Hamiltonian vector fields generated by them coincide with the 2π-periodic vector fields X_i defined by (11.8.1). Then, an application of Theorem 11.8.1 will yield the assertion of the theorem. For that purpose, let Φ^{H_j} be the flow of X_{H_j}. By the assumptions of the theorem, the foliation of W into tori is invariant with respect to Φ^{H_j}, that is, $\Phi_t^{H_j}(\Sigma(m)) = \Sigma(\Phi_t^{H_j}(m))$. We calculate

$$\{H_j, I_i\}(m) = \frac{\mathrm{d}}{\mathrm{d}t}{\restriction_0} I_i \circ \Phi_t^{H_j}(m) = \frac{1}{2\pi} \frac{\mathrm{d}}{\mathrm{d}t}{\restriction_0} \int_{\gamma_i(\Phi_t^{H_j}(m))} \tau = \frac{1}{2\pi} \frac{\mathrm{d}}{\mathrm{d}t}{\restriction_0} \int_{\Phi_t^{H_j} \circ \gamma_i(m)} \tau.$$

Let $A(t)$ be the 2-dimensional surface obtained by acting with $\Phi_t^{H_j}$ on γ_i. Then,

$$\frac{\mathrm{d}}{\mathrm{d}t}{\restriction_0} \int_{\Phi_t^{H_j} \circ \gamma_i(m)} \tau = \lim_{t \to 0} \frac{1}{t} \left(\int_{\Phi_t^{H_j} \circ \gamma_i(m)} \tau - \int_{\gamma_i(m)} \tau \right) = \lim_{t \to 0} \frac{1}{t} \int_{A(t)} \omega = 0,$$

[35] To see this, note that one can choose the family $\{\gamma_i(m)\}$ to be differentiable in m.

because $A(t)$ isotropic. Indeed, its tangent spaces are spanned by X_{H_j} and a linear combination of the vectors $(X_{H_1}, \dots, X_{H_{n-k}})$ commuting with X_{H_j}. We conclude that

$$\{H_j, I_i\} = 0, \quad j = 1, \dots, n+k,\ i = 1, \dots, n-k. \tag{11.8.3}$$

Since ω is non-degenerate, it follows that on W we can express the variables I_i as smooth functions of $(H_1, \dots, H_{n-k})$. Therefore, $X_{I_i} = \frac{\partial I_i}{\partial H_j} X_{H_j}$ and hence the functions I_i are in involution.

Let us calculate the Jacobi matrix $\frac{\partial I_i}{\partial H_j}$. Let $m \in \Sigma$, let $\beta_i \subset \Sigma$ be the integral curve through m of the 2π-periodic vector field X_i and let φ^i be the flow parameter on β_i. Consider the line segment

$$I_t = \big\{\big(H_1(m), \dots, H_{j-1}(m), H_j(m)+s, H_{j+1}(m), \dots, H_{n+k}(m)\big) : 0 \le s \le t\big\} \subset B$$

and take the 2-dimensional surface $S(t) = \Phi^{-1}(I_t \times \Phi(\beta_i)) \subset W$. Then, we have

$$\frac{\partial I_i}{\partial H_j}(m) = \lim_{t\to 0} \frac{1}{2\pi t} \int_{\partial S(t)} \tau = \lim_{t\to 0} \frac{1}{2\pi t} \int_{S(t)} \omega = \frac{1}{2\pi} \int_0^{2\pi} \omega(\partial_j, X_i)\mathrm{d}\varphi^i,$$

where ∂_j denotes the j-th partial derivative in the coordinate system defined by Φ. Using (11.8.1), we conclude

$$\frac{\partial I_i}{\partial H_j}(m) = b_i{}^j(m), \quad m \in \Sigma, \tag{11.8.4}$$

and thus $X_{I_i} = X_i$, $i = 1, \dots, n-k$. Thus, the vector fields X_{I_i} are 2π-periodic. Since the functions I_i are in involution, we can apply Theorem 11.8.1, which tells us that for any $m \in \Sigma$, there exists a neighbourhood U in M and smooth functions $\vartheta^1, \dots, \vartheta^{n-k}, q^1, \dots, q^k$ and $p_1, \dots, p_k$ on U such that

$$\omega_{\restriction U} = \mathrm{d}I_i \wedge \mathrm{d}\vartheta^i + \mathrm{d}p_j \wedge \mathrm{d}q^j.$$

Let Φ^{I_j} be the flow of X_{I_j} and let $\Psi^I_{\mathbf{t}} := \Phi^{I_1}_{t_1} \circ \dots \circ \Phi^{I_j}_{t_{n-k}}$ be the corresponding action of $\mathbb{R}^{n-k}$. This action is transitive on every level set component. Since the functions q^j and p_j are constant along the tori, via this mapping they trivially extend to functions on $W = \Psi^I_{\mathbb{R}^{n-k}}(U)$. To extend the functions ϑ^i, note that on U we have $X_{I_i} = \partial_{\vartheta^i}$ and thus

$$\vartheta^i\big(\Psi^I_{\mathbf{t}}(m)\big) = \vartheta^i(m) + t^i \tag{11.8.5}$$

for all $m \in U$ and $\mathbf{t} \in \mathbb{R}^{n-k}$ such that $\Psi^I_{\mathbf{t}}(m) \in U$. Thus, for a covering of W by open subsets $U_k := \Psi^I_{\mathbf{t}_k}(U)$ with appropriately chosen $\mathbf{t}_k \in \mathbb{R}^{n-k}$, $k \in \mathbb{Z}$, one can define the extension of ϑ^i to U_k by setting

$$\vartheta^i_{\restriction U_k}(m) := \vartheta^i\big(\Psi^I_{-\mathbf{t}_k}(m)\big) + t^i_k,$$

because (11.8.5) ensures that the functions so defined coincide on any nontrivial intersection of the U_k and thus combine to smooth functions (mod 2π) on W. In order to show that the functions q^j, p_j and ϑ^i complement the functions I_i to Darboux coordinates on W, it suffices to prove that the mapping $W \to \mathbb{R}^{2n}$ defined by ϑ, I, q and p is injective. The latter follows from the transitivity of Ψ^I and (11.8.5). $\square$

The Nekhoroshev Theorem immediately implies

Corollary 11.8.4 *Let (M, ω, H) be a $2n$-dimensional Hamiltonian system such that there exist $(n+k)$ constants of motion fulfilling the assumptions of Theorem 11.8.3. In the Darboux coordinates $\vartheta^i, q^j, I_i, p_j$ provided by this theorem, the Hamiltonian function H depends on I only and the Hamilton equations read*

$$\dot{I}_j = 0, \qquad \dot{\vartheta}^j = \frac{\partial H(\mathbf{I})}{\partial I_j} \equiv \omega^j(\mathbf{I}), \qquad \dot{p}_i = 0, \qquad \dot{q}^i = 0. \tag{11.8.6}$$

Thus, the integral curves of the system are located on $(n-k)$-dimensional tori and the motion is quasiperiodic with frequencies being functions of the constants of motion $I_1, \ldots, I_{n-k}$.

Another important class of non-commutatively integrable systems is the following, studied by Fomenko and Mishchenko [214]. Let (M, ω, H) be a $2n$-dimensional Hamiltonian system, let p be an integer between n and $2n$ and let $H_1, \ldots, H_p$ be functionally independent constants of motion. Let $\mathfrak{g}$ be the linear subspace of $C^\infty(M)$ spanned by the functions H_i. Assume that

(a) the Hamiltonian vector fields X_{H_i} are complete,
(b) $\mathfrak{g}$ is closed under the Poisson bracket, that is, $\{H_i, H_j\} = c^k_{ij} H_k$ with $c^k_{ij} \in \mathbb{R}$.

By assumption (b), $\mathfrak{g}$ is a Lie algebra of dimension p. It acts symplectically from the right on M by the homomorphism $\mathfrak{g} \to \mathfrak{X}(M)$, $A \mapsto X_A$. By assumption (a), this action of $\mathfrak{g}$ integrates to a symplectic action $\tilde{\Psi}$ of the corresponding simply connected Lie group[36] $\tilde{G}$ on M, given by

$$\tilde{\Psi}_{\exp_{\tilde{G}} A}(m) = \Phi^A_1(m)$$

for all $A \in \mathfrak{g}$ and $m \in M$, where Φ^A denotes the flow of X_A. One can show that the adjoint action of $\tilde{G}$ on $\mathfrak{g}$ is given by

$$\mathrm{Ad}(a^{-1})A = A \circ \tilde{\Psi}_a \tag{11.8.7}$$

for all $A \in \mathfrak{g}$ and $a \in \tilde{G}$ (Exercise 11.8.2). Note that in the special case where $p = n$ and $\mathfrak{g}$ is Abelian, $\tilde{G} \cong \mathbb{R}^n$ and we are in the situation of an ordinary integrable system whose Hamiltonian vector fields X_{H_i} are complete, cf. Sect. 11.3.

[36] Every finite-dimensional abstract Lie algebra is the Lie algebra of a simply connected Lie group $\tilde{G}$, which is unique up to isomorphy [302, Thm. 3.28].

Since, by definition, we have $A_* = X_A$ for all $A \in \mathfrak{g}$, the mapping $J : M \to \mathfrak{g}^*$ defined by

$$\big\langle J(m), A\big\rangle = A(m), \quad A \in \mathfrak{g}, \tag{11.8.8}$$

is a momentum mapping for the action $\tilde{\Psi}$. As a consequence of (11.8.7), it is equivariant (Exercise 11.8.2). If we identify $\mathfrak{g}^*$ with $\mathbb{R}^n$ by means of the basis dual to the basis $\{H_i\}$ in $\mathfrak{g}$, J coincides with the mapping $\mathscr{H} = (H_1, \ldots, H_p)$. Since the H_i are functionally independent, J is a submersion. This implies the following.

(a) The level set components of J are embedded submanifolds of dimension $2n - p$. They are the maximal integral manifolds of the regular distribution $\ker J'$, cf. Example 3.5.4/4.
(b) Corollary 10.2.2/2 implies that the stabilizer $\tilde{G}_m$ is discrete, that is, $\mathfrak{g}_m = 0$. Hence, the action of $\tilde{G}$ is locally free. Consequently, the kernel $\ker \tilde{\Psi} = \bigcap_{m \in M} G_m$ is discrete and the quotient group $G := \tilde{G}/\ker \tilde{\Psi}$ has Lie algebra $\mathfrak{g}$ and universal covering group $\tilde{G}$. The induced G-action on M is denoted by Ψ. It is obviously effective and locally free.
(c) The image of J is open and thus it intersects the principal stratum of the coadjoint action, that is, it contains an element μ such that the Lie algebra $\mathfrak{g}_\mu$ of the stabilizer G_μ has minimal dimension. This implies that $\mathfrak{g}_\mu$ is Abelian, see [77]. Hence, the identity connected component G^0_μ of G_μ is Abelian, too. Using Proposition 6.2.2/3 and Formula (6.2.3), we determine $\mathfrak{g}_\mu$ explicitly (Exercise 11.8.3):

$$\mathfrak{g}_\mu = \big\{A \in \mathfrak{g} : \{A, B\}(m) = 0 \text{ for all } B \in \mathfrak{g},\ m \in J^{-1}(\mu)\big\}. \tag{11.8.9}$$

Let $\mu \in \mathfrak{g}^*$ be a value of J. Since, by equivariance, $J^{-1}(\mu)$ is invariant under the action of the stabilizer G_μ of μ under the coadjoint representation, and since the induced action of G_μ on $J^{-1}(\mu)$ is locally free,

$$\dim \mathfrak{g}_\mu \le \dim J^{-1}(\mu) = \dim M - \dim \mathfrak{g}, \tag{11.8.10}$$

where the equality is due to the fact that J is a submersion. In the special case of an ordinary integrable system, $\mathfrak{g}$ is Abelian, and hence $\dim \mathfrak{g}_\mu = \dim \mathfrak{g} = n = \dim J^{-1}(\mu)$. A reasonable generalization to the present situation is to assume that $\dim \mathfrak{g}_\mu = \dim J^{-1}(\mu)$, which is equivalent to

$$\dim M = \dim \mathfrak{g} + \dim \mathfrak{g}_\mu. \tag{11.8.11}$$

Let us assume that there exist $\mu \in \mathfrak{g}^*$ which fulfil this equality and belong to the principal stratum[37] of $\mathfrak{g}^*$. Then, (11.8.10) implies that the image of J is completely contained in this stratum, that is, for every value μ of J condition (11.8.11) holds and $\mathfrak{g}_\mu$ is Abelian. Under this assumption, we have the following generalization of the Arnold Theorem.

[37] The subset of elements with minimal stabilizer under the coadjoint action.

Theorem 11.8.5 (Mishchenko-Fomenko) *Let $H_1, \dots, H_p$, $n \le p < 2n$, be independent functions on a $2n$-dimensional symplectic manifold (M, ω) whose Hamiltonian vector fields X_{H_i} are complete. Assume that this system closes under the Poisson bracket, thus spanning a p-dimensional Lie subalgebra $\mathfrak{g}$ of $C^\infty(M)$. Let G be the induced effective symmetry group and let Σ be a level set component of the associated momentum mapping J with value $J(\Sigma) = \mu$. Assume that $\mathfrak{g}_\mu$ belongs to the principal stratum of the coadjoint action of G and satisfies (11.8.11). Then, Σ is an orbit of the identity connected component G^0_μ of G_μ and hence diffeomorphic to $\mathrm{T}^l \times \mathbb{R}^{2n-p-l}$ for some $0 \le l \le 2n - p$.*

Proof Let $m_0 \in \Sigma$. Since $J^{-1}(\mu)$ is invariant under the action of G_μ, the level set component Σ is invariant under the action of the identity connected component G^0_μ. According to (11.8.11), the Killing vector fields of the action of G^0_μ on Σ span the tangent bundle of Σ. Hence, for every $m \in \Sigma$, the orbit mapping $G^0_\mu \to \Sigma$, $a \mapsto \Psi_a(m)$, is a local diffeomorphism. Using this, by the same argument as in the proof of the Arnold Theorem, one can show that G^0_μ acts transitively on Σ. Then, the Orbit Theorem 6.2.8 implies that Σ is diffeomorphic to the quotient of G^0_μ with respect to the stabilizer G_m of some[38] point m of Σ. Since $\mathfrak{g}_\mu$ is Abelian, so is G^0_μ. Hence, the quotient G^0_μ / G_m is an Abelian Lie group. Since G_m is discrete, G^0_μ / G_m has dimension $2n - p$ and is, therefore, isomorphic to $\mathrm{T}^l \times \mathbb{R}^{2n-p-l}$ for some l between 0 and $2n - p$.[39] □

Remark 11.8.6

1. In Sect. 11.3 we observed that in the case of commutative integrability, symplectic reduction of the level sets of J yields discrete reduced phase spaces. Here, under the assumption that Eq. (11.8.11) holds, we have a certain non-commutative analogue of this situation: by Theorem 11.8.5, the topological quotient

$$J^{-1}(\mu)/G_\mu \equiv \bigl(J^{-1}(\mu)/G^0_\mu\bigr)/\bigl(G_\mu/G^0_\mu\bigr)$$

   is discrete and hence trivially a symplectic manifold, which may be interpreted as the reduced phase space at μ.
2. Under the additional assumption that the G-action be free and proper, the theory of regular symplectic reduction applies: the level set $M_\mu = J^{-1}(\mu)$ is an embedded submanifold, G_μ acts freely and properly on M_μ and the discrete reduced phase space M_μ / G_μ results from the Regular Reduction Theorem 10.3.1. The level set $\mathscr{M}_\mu = J^{-1}(\mathscr{O}_\mu)$, with $\mathscr{O}_\mu \subset \mathfrak{g}^*$ denoting the coadjoint orbit through μ, is a submanifold of M diffeomorphic to $M_\mu \times_{G_\mu} G$, see (10.3.14). Since M_μ / G_μ is discrete, $\mathscr{M}_\mu$ is diffeomorphic to a direct sum of group manifolds G_μ. Since

[38] In fact, since G^0_μ is Abelian, all points have the same stabilizer.

[39] Every finite-dimensional connected Abelian Lie group is of this form, see Exercise 18 in Chap. 3 of [302].

G_μ acts freely on $\mathscr{M}_\mu$, the Killing vector fields generated by the elements of $\mathfrak{g}_\mu$ span an integrable distribution on $\mathscr{M}_\mu$, which yields a foliation of $\mathscr{M}_\mu$ into leaves diffeomorphic to G^0_μ.

Under the additional assumption that the G-action is free and proper, we have the following non-commutative analogue of the existence theorem 11.4.3 for action and angle variables.

Theorem 11.8.7 *Let (M, ω) be a $2n$-dimensional symplectic manifold satisfying the assumptions of Theorem 11.8.5. Assume, in addition, that the G-action be free and proper. Then, the following holds.*

1. *There exists an open neighbourhood U of the origin in $\mathfrak{g}^*_\mu$ and an anti-equivariant symplectomorphism Φ from a G-invariant open neighbourhood W of Σ in M onto $G \times U$, endowed with the symplectic form*

$$\omega_{(g,\nu)}\big((\mathrm{L}'_g A_1, \sigma_1), (\mathrm{L}'_g A_2, \sigma_2)\big) = \langle \sigma_1, A_2 \rangle - \langle \sigma_2, A_1 \rangle - \big\langle \mu + \nu, [A_1, A_2] \big\rangle$$

and the action of G by left translation on the factor G. Φ maps $G \cdot \Sigma$ to $G \times \{0\}$ and satisfies

$$J \circ \Phi^{-1}(a, \nu) = \mathrm{Ad}^*(a)(\mu + \nu). \tag{11.8.12}$$

2. *There exist smooth functions $I_1, \ldots, I_{2n-p}$ on W in involution whose flows are complete and generate an action of $\mathbb{R}^{2n-p}$ on W with the common stabilizer $2\pi\mathbb{Z}^l \times \{0\} \subset \mathbb{R}^{2n-p}$. The orbits of this action are the level set components of J in W and the level sets of the mapping $I = (I_1, \ldots, I_{2n-p})$ are the G-orbits in W.*
3. *On some neighbourhood W_0 of Σ in W, which can be chosen to be a union of level set components of J, there exist smooth functions*[40] *$\vartheta^1, \ldots, \vartheta^{2n-p}$, $q^1, \ldots, q^{p-n}$ and $p_1, \ldots, p_{p-n}$ such that*

$$\omega = \sum_{i=1}^{2n-p} \mathrm{d}I_i \wedge \mathrm{d}\vartheta^i + \sum_{j=1}^{p-n} \mathrm{d}p_j \wedge \mathrm{d}q^j. \tag{11.8.13}$$

The functions I_i, q^j and p_j parameterize the level set components of J in W_0 and the functions ϑ^i provide coordinates on the latter.

As in the situation of the Nekhoroshev Theorem, by a slight abuse of language, the coordinates provided by this theorem are referred to as generalized action and angle coordinates. Note that the $\vartheta^{l+1}, \ldots, \vartheta^{2n-p}$ do not provide angle variables.

Proof 1. First, we use the anti-isomorphism of Lie group actions given by the identical mapping of M and the inversion mapping of G to turn the right action Ψ on

[40] With $\vartheta^1, \ldots, \vartheta^l$ being multi-valued mod 2π.

$M^{\mathscr{H}}$ into a left action Ψ^{L}. Both actions have the same orbits and the same invariant subsets. For Ψ^{L}, the Symplectic Tubular Neighbourhood Theorem 10.4.4 yields a Ψ^{L}-equivariant symplectomorphism Φ of an invariant open neighbourhood W of the orbit of some $m \in \Sigma$ onto the subset $\Phi(W) = G \times_{G_m} (\tilde{\mathfrak{m}}^* \oplus \tilde{V})$ of the twisted product $E = G \times_{G_m} (\mathfrak{m}^* \oplus V)$, where $\mathfrak{m}$ is a vector space complement of $\mathfrak{g}_m$ in $\mathfrak{g}_\mu$, cf. (10.2.6), and V is some symplectic slice at m, and where $\tilde{\mathfrak{m}}^*$ and $\tilde{V}$ are open neighbourhoods of the origin in $\mathfrak{m}^*$ and V, respectively. Since G_m is trivial, we have $\mathfrak{m} = \mathfrak{g}_\mu$. Then, (10.2.11) and condition (11.8.11) imply $V = 0$, so that we end up with $E = G \times U$, where U is some open neighbourhood of the origin in $\mathfrak{g}_\mu^*$. Since Σ is contained in the orbit of m, it is contained in W and we have $G \cdot \Sigma = G \cdot m$. Hence, it is mapped onto the zero section of E. Finally, in this special situation, the symplectic form on $\Phi(W)$ is given by (10.4.12) and the momentum mapping is given by the normal form (10.4.15).

2. Choose a basis $\{E_1, \ldots, E_{2n-p}\}$ in $\mathfrak{g}_\mu$ and define $I_i : W \to \mathbb{R}$ by

$$I_i(m) := \big\langle \mathrm{pr}_{\mathfrak{g}_\mu^*} \circ \Phi(m), E_i \big\rangle,$$

where $\mathrm{pr}_{\mathfrak{g}_\mu^*} : G \times \mathfrak{g}_\mu^* \to \mathfrak{g}_\mu^*$ denotes the natural projection. Obviously, the level sets of $I = (I_1, \ldots, I_{2n-p})$ coincide with the G-orbits in W. To determine the Hamiltonian vector field X_{I_i}, we pass to $G \times U$ via Φ. Denote $\tilde{X}_{I_i} := \Phi_* X_{I_i}$. For $(a, \nu) \in G \times \mathfrak{g}_\mu^*$ and a tangent vector (B_a, σ) with $B \in \mathfrak{g}$ and $\sigma \in \mathfrak{g}_\mu^*$ we compute

$$(\mathrm{d}I_i)_{(a,\nu)}(B_a, \sigma) = \langle \sigma, E_i \rangle.$$

With the ansatz $(\tilde{X}_{I_i})_{(a,\nu)} = (A_a, \rho)$, the defining equation for $\tilde{X}_{I_i}$ reads

$$\langle \rho, B \rangle - \langle \sigma, A \rangle - \big\langle \mu + \nu, [A, B] \big\rangle = -\langle \sigma, E_i \rangle$$

for all $B \in \mathfrak{g}$ and $\sigma \in \mathfrak{g}_\mu^*$. Since $\mathfrak{g}_\mu$ is the stabilizer of μ, we have $\mathrm{ad}^*(E_i)\mu = 0$. Since $\mathfrak{g}_\mu$ is Abelian, $\mathrm{ad}^*(E_i)\nu = 0$. Hence, $A = E_i$, $\rho = 0$ is a solution, and it is, therefore, the unique solution. Thus, $(\tilde{X}_{I_i})_{(a,\nu)} = ((E_i)_a, 0)$ and the flow is complete and given by

$$\big(s, (a, \nu)\big) \mapsto \big(a \exp(s E_i), \nu\big).$$

Since $\mathfrak{g}_\mu$ is Abelian, the vector fields $\tilde{X}_{I_i}$ commute pairwise. Hence, according to (6.1.5), their flows define an action $\tilde{\Psi}^I$ of $\mathbb{R}^{2n-p}$ on $G \times U$. Due to Proposition 5.3.10, this action can be written in the form

$$\tilde{\Psi}^I_{\mathbf{t}}(a, \nu) = \left(a \exp\left(\sum_{i=1}^{2n-p} t_i E_i \right), \nu \right), \quad \mathbf{t} = (t_1, \ldots, t_{2n-p}). \tag{11.8.14}$$

The corresponding orbit mapping of the point $(\mathbb{1}, 0)$ induces a Lie group homomorphism from $\mathbb{R}^{2n-p}$ to G_μ^0. Since G_μ^0 is isomorphic to $\mathrm{T}^l \times \mathbb{R}^{2n-p-l}$, the basis elements E_i can be chosen so that the kernel of this homomorphism is $2\pi\mathbb{Z}^l \times \{0\}$. Then, this is also the common stabilizer of all points of $G \times \mathfrak{g}_\mu^*$.

Next, we return to W. The Hamiltonian vector fields X_{I_i} commute and hence their flows define an action Ψ^I of $\mathbb{R}^{2n-p}$ given by $\Psi^I = \Phi \circ \tilde{\Psi}^I \circ \Phi^{-1}$. Then, (11.8.14) reads

$$\Psi^I_{\mathbf{t}} \circ \Phi(a, \nu) = \Phi\left(a \exp\left(\sum_{i=1}^{2n-p} t_i E_i\right), \nu\right). \tag{11.8.15}$$

We show that the orbits of Ψ^I coincide with the level set components of J in W. Let $m = \Phi(a, \nu) \in W$ and let $\Sigma(m)$ be the level set component of J containing m. By (11.8.12),

$$J(m) = \mathrm{Ad}^*(a)(\mu + \nu).$$

Since $\nu \in \mathfrak{g}^*_\mu$ and $\mathfrak{g}_\mu$ is Abelian, $\mathfrak{g}_{J(m)} = \mathrm{Ad}(a)\mathfrak{g}_{\mu+\nu} \supset \mathrm{Ad}(a)\mathfrak{g}_\mu$. Since $\mathfrak{g}_\mu$ has minimal dimension, we can choose U so that $\mathfrak{g}_{\mu+\nu}$ has minimal dimension for all $\nu \in U$. First, this implies $\mathfrak{g}_{\mu+\nu} = \mathfrak{g}_\mu$ and hence $\mathfrak{g}_{J(m)} = \mathrm{Ad}(a)\mathfrak{g}_\mu$. Second, in view of the first assertion, this implies that $\Sigma(m)$ is the orbit of m under the action of the identity connected component $G^0_{J(m)}$ of $G_{J(m)}$. By the previous argument, $G^0_{J(m)} = aG^0_\mu a^{-1}$. Finally, since for $b \in G^0_\mu$ we have $\Psi_{aba^{-1}}(m) = \Phi(ab^{-1}, \nu)$, (11.8.15) implies that the orbit of $aG^0_\mu a^{-1}$ through m coincides with the Ψ^I-orbit. Finally, the functions I_i are in involution, because their Hamiltonian vector fields commute.

3. This follows by the same arguments as in the Nekhoroshev Theorem 11.8.3. □

Remark 11.8.8 If the functions $H_1, \dots, H_p$ are constants of motion with respect to a given Hamiltonian function on (M, ω), we have a corollary completely analogous to Corollary 11.8.4. In particular, in the generalized action and angle variables provided by the theorem, the dynamics of the Hamiltonian system is given by (11.8.6).

Remark 11.8.9

1. The Mishchenko-Fomenko Theorem can be generalized in various directions. In particular, one can weaken the assumption that the functions H_i form a finite-dimensional Lie algebra by requiring that

$$\{H_i, H_j\} = P_{ij}(H),$$

   where P_{ij} is a matrix of constant rank, see e.g. [89], [91], [261] and the references therein. In [261] a nice discussion of non-commutative integrability of the Kepler problem can be found. We also recommend the survey article [154].
2. One can prove that non-commutative integrability implies ordinary integrability, see [214] for a restricted class of Lie algebras (including semisimple ones) and [258] for the general case.
3. If we choose W small enough so that for every level set of J it contains at most one connected component, the space of level set components in W may be identified with $J(W) \subset \mathfrak{g}^*$ and the corresponding natural projection with

$J : W \to J(W) \subset \mathfrak{g}^*$. On the other hand, combining the G-equivariant symplectomorphism Φ with the natural projection $\mathrm{pr}_{\mathfrak{g}^*_\mu} : G \times \mathfrak{g}^*_\mu \to \mathfrak{g}^*_\mu$, we obtain a smooth mapping

$$\mathscr{I} := \mathrm{pr}_{\mathfrak{g}^*_\mu} \circ \Phi : W \to \mathfrak{g}^*_\mu .$$

By definition, $\mathscr{I} = \sum_{i=1}^{2n-p} I_i E_i$. Since the fibres of this mapping are foliated by level set components of J, there exists a smooth mapping $\pi : J(W) \to \mathfrak{g}^*_\mu$ such that the diagram

$$\begin{array}{ccc} & W & \\ {}^{\mathscr{I}}\swarrow & & \searrow^{J} \\ \mathfrak{g}^*_\mu & \xleftarrow[\pi]{} & J(W) \subset \mathfrak{g}^* \end{array} \tag{11.8.16}$$

commutes. One says that $J_{\restriction W}$ and $\mathscr{I}$ define a bifibration of W. The fibres of $J_{\restriction W}$ are isotropic and the fibres of $\mathscr{I}$ are coisotropic. We show that $\pi^{-1}(\nu)$ coincides with the coadjoint orbit through $\mu + \nu$. Indeed, using (11.8.12) and the equivariance of J we obtain

$$\pi^{-1}(\nu) = J\big(\mathscr{I}(\nu)\big) = J\big(G \cdot \Phi^{-1}(\mathbb{1}, \nu)\big) = \mathrm{Ad}^*(G)\big(J\big(\Phi^{-1}(\mathbb{1}, \nu)\big)\big) = \mathscr{O}_{\mu+\nu} .$$

In particular, the fibres of π coincide with symplectic leaves of the Lie-Poisson structure of $\mathfrak{g}^*$, discussed in Example 8.2.18/3. Via J, the generalized action and angle variables ϑ^i, q^j, I_i, p_j induce coordinates $\tilde{q}^j$, $\tilde{I}_i$, $\tilde{p}_j$ on $\mathfrak{g}^*$ and, via $\mathscr{I}$, coordinates $\hat{I}_i$ on $\mathfrak{g}^*_\mu$, which in this context is often referred to as the action space. Let us add that the above bifibration defines a dual pair in the sense of Weinstein [310]. For more details we refer to [89] and Chap. 11 of [232].

To conclude the discussion of non-commutative integrability, let us consider a $2n$-dimensional Hamiltonian system which is both non-commutatively integrable, with momentum mapping J, and integrable in the ordinary sense, with the independent constants of motion in involution $\mathscr{H} = (H_1, \dots, H_n)$. Let $m \in M$ be a regular point of both J and $\mathscr{H}$ and assume that the level set components of m with respect to J and to $\mathscr{H}$ are compact. On the one hand, according to Theorem 11.4.3, m possesses an open neighbourhood U which is foliated by n-dimensional tori which are invariant under the dynamics, thus giving rise to ordinary action and angle variables. On the other hand, according to Theorem 11.8.7, the dynamics in U further reduces to lower-dimensional tori. In general, the decomposition into n-tori is not compatible with the topology of the non-commutatively integrable system. The following example illustrates this point.

Example 11.8.10 (Symmetric Euler top) We take up Example 11.5.4. Recall that by means of the left trivialization of $\mathrm{T}^*\mathrm{SO}(3)$ and the natural isomorphism $\mathfrak{so}(3) \cong \mathbb{R}^3$ of Example 5.2.8, the phase space of the symmetric Euler top can be identified with

$M = \mathrm{SO}(3) \times \mathbb{R}^3$. In these variables, the action of SO(3) on M induced by left translation is given by

$$\Psi_b(a, \mathbf{L}) = (ba, \mathbf{L}).$$

In order to reserve the symbols G and J for the Andoyer variables, in this example, we denote the Lie group acting effectively by $\hat{G}$ and the corresponding momentum mapping by $\hat{J}$. Consider the Andoyer function L on M, given by

$$L(a, \mathbf{L}) := \mathbf{L} \cdot \mathbf{e}_3,$$

and the components of the momentum mapping of the action of SO(3) on M by left translation, given by

$$K(a, \mathbf{L}) = a\mathbf{L}.$$

We have

$$\{K_i, K_j\} = \varepsilon_{ij}{}^k K_k, \qquad \{L, K_i\} = 0, \tag{11.8.17}$$

where the last equation is due to the fact that X_{K_i} are Killing vector fields for the SO(3)-action, whereas L is invariant. Thus, the functions L, K_1, K_2, K_3 span a Lie subalgebra $\mathfrak{g}$ of $C^\infty(M)$. It is isomorphic to $\mathbb{R} \oplus \mathfrak{so}(3)$ via the assignment $L \mapsto (1, 0)$, $K_i \mapsto \mathrm{I}_i^{\mathbb{R}}$, cf. Example 5.2.8. The corresponding momentum mapping is

$$\hat{J} = (L, K_1, K_2, K_3)$$

and the associated simply connected Lie group is $\tilde{G} = \mathbb{R} \times \mathrm{SU}(2)$. An easy computation (Exercise 11.8.4) shows that its action on M is given by

$$\Psi_{(\alpha,u)}(a, \mathbf{L}) = \big(\phi(u)a, R_\alpha \mathbf{L}\big), \qquad \alpha \in \mathbb{R}, \quad u \in \mathrm{SU}(2), \tag{11.8.18}$$

where $\phi : \mathrm{SU}(2) \to \mathrm{SO}(3)$ is the covering homomorphism of Example 5.1.11 and R_α denotes rotation about $\mathbf{e}_3$ by the angle α. The subset of regular points of $\hat{J}$ is given by

$$M^{\hat{J}} = \big\{(a, \mathbf{L}) \in \mathrm{SO}(3) \times \mathbb{R}^3 : \mathbf{L} \times \mathbf{e}_3 \neq 0\big\} = \mathrm{SO}(3) \times \big(\mathbb{R}^3 \setminus \mathbb{R}\mathbf{e}_3\big).$$

The points of $M^{\hat{J}}$ have the common stabilizer $\tilde{G}_{(a,\mathbf{L})} = 2\pi\mathbb{Z} \times \{\pm\mathbb{1}\}$, hence the group acting effectively is

$$\hat{G} = \mathrm{U}(1) \times \mathrm{SO}(3)$$

and this action is free and proper. The values of $\hat{J}_r$ are given by

$$\mu = (\lambda, \mathbf{k}) \in \mathbb{R} \times \mathbb{R}^3 \setminus \{0\}, \qquad |\lambda| < \|\mathbf{k}\|.$$

The stabilizer subgroup $\hat{G}_\mu$ is the direct product of U(1) with the subgroup of SO(3) of rotations about $\mathbf{k}$, hence $\hat{G}_\mu \cong \mathrm{T}^2$. Under the identification $\hat{\mathfrak{g}} \cong \mathbb{R} \oplus \mathfrak{so}(3)$, the stabilizer subalgebra $\hat{\mathfrak{g}}_\mu$ is given by $\mathbb{R} \oplus \mathbb{R}\mathbf{k} \subset \mathbb{R} \oplus \mathbb{R}^3$. In terms of the original representation of $\hat{\mathfrak{g}}$ in terms of functions, $\hat{\mathfrak{g}}_\mu$ is the subalgebra spanned by the two functions L and $k^i K_i$. The dimension condition (11.8.11) is obviously satisfied for all values μ of $\hat{J}_r$. Hence, Theorem 11.8.7 yields that the level set components of $\hat{J}_r$ are 2-tori. The level set components can also be obtained directly by acting with $\hat{G}_\mu$ on a point $(a, \mathbf{L})$ with $\mathbf{L} \cdot \mathbf{e}_3 = \lambda$ and $a\mathbf{L} = \mathbf{k}$, which amounts to left translations of a by rotations about $a\mathbf{L}$ and to rotations of $\mathbf{L}$ about the $\mathbf{e}_3$-axis.

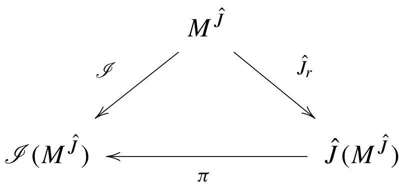
Now, consider the Andoyer variable G, given by

$$G(a, \mathbf{L}) = \|\mathbf{L}\|.$$

We already know that X_L and X_G define an action of $\mathbb{R}^2$ on $M^{\hat{J}}$ with common stabilizer $2\pi\mathbb{Z}^2$. Due to $\{L, G\} = 0$ and $\{G, J_i\} = 0$, this action leaves the level set components of $\hat{J}_r$ invariant and hence the latter are the orbits of this action. Moreover, the level sets of the mapping $\mathscr{I} := (L, G)$ are the orbits of the action of the symmetry group $\mathrm{U}(1) \times \mathrm{SO}(3)$ and the image is given by

$$\mathscr{I}\big(M^{\hat{J}}\big) = \big\{(\lambda, \gamma) \in \mathbb{R}^2 : |\lambda| < \gamma\big\}.$$

Thus, as in the situation of Remark 11.8.9/3, we have a bifibration of the type (11.8.16), which here reads



The induced projection π is given by

$$\pi(\lambda, \mathbf{k}) = (\lambda, \|\mathbf{k}\|)$$

and its fibres are $\{\lambda\} \times \mathrm{S}^2_{\|\mathbf{k}\|}$, where $\mathrm{S}^2_{\|\mathbf{k}\|}$ is the sphere of radius $\|\mathbf{k}\|$ in $\mathbb{R}^3$, representing the (co)adjoint orbit of $\mathbf{k}$. Thus, the action variables L and G are adapted to the foliation of $M^{\hat{J}}$ induced by the symmetry associated with $\hat{\mathfrak{g}}$: the fibre of $\hat{J}_r$ over (λ, γ) is a bundle with fibre T^2 over $\{\lambda\} \times \mathrm{S}^2_{\|\mathbf{k}\|}$. For every λ, this bundle is a direct product of a trivial S^1-bundle over the point λ with a nontrivial S^1-bundle over S^2 of the type $\mathrm{SO}(3) \to \mathrm{S}^2$, that is, the $\mathrm{SO}(3) \times \mathrm{S}^1$-fibres of $\mathscr{I}$ cannot be decomposed globally into 3-dimensional tori. This can be further illustrated using the remaining Andoyer variable J: by Remark 11.5.5, J is not globally defined on $M^{\hat{J}}$, but only on the open subset

$$M_* = \big\{(a, \mathbf{L}) \in \mathrm{SO}(3) \times \mathbb{R}^3 : \mathbf{L} \times \mathbf{e}_3 \neq 0, \mathbf{n}_3 \times \mathbf{L} \neq 0\big\}. \tag{11.8.19}$$

The Andoyer variables are adapted to the corresponding bundles as follows: (g, l) are angle coordinates in the T^2-fibres of $\hat{J}_{\restriction M_*}$, (G, L, J, j) are coordinates on

$\hat{J}(M_*)$, (g, l, J, j) are coordinates on the fibres of $\mathscr{I}_{\restriction M_*}$ and (G, L) are coordinates on $\mathscr{I}(M_*)$. By the condition $\mathbf{n}_3 \times \mathbf{L} \neq 0$, the points $(j, J) = (0, G)$ (north pole) and $(j, J) = (\pi, -G)$ (south pole) are excluded, that is, (J, j) yield local coordinates on the 2-sphere with radius G. Altogether, the action and angle variables define a diffeomorphism

$$(G, L, J, g, l, j) : M_* \to N_* \times \mathrm{T}^3, \quad N_* = \big\{(G, L, J) \in \mathbb{R}^3 : |L| < G, |J| < G\big\}.$$

To cover M by action and angle variables, one has to introduce a second Andoyer chart. This is explained in detail in a paper by Fasso [89]. From the above discussion we see that it is quite unnatural to decompose the phase space into 3-tori. The motion of the top takes place on invariant 2-tori and the additional angle variable j has no physical meaning. The definition of the pair (J, j) of variables is of local nature and depends on the choice of the inertial frame.

Exercises

11.8.1 Complete the proof of Theorem 11.8.1 by showing that, with respect to the Poisson bivector field Π, $(D^\beta)^0$ is orthogonal to E.
Hint. Use Proposition 5.4 in [181, §I.3].

11.8.2 Verify Formula (11.8.7) and use this to prove that the momentum mapping J defined by (11.8.8) is equivariant.

11.8.3 Prove Formula (11.8.9).

11.8.4 In Example 11.8.10, verify that the effective action induced by L, K_1, K_2, K_3 is given by Formula (11.8.18).

11.8.5 Show that the Kepler problem yields a non-commutatively integrable model both for positive and for negative values of the energy. Work out the details for negative energy values.
Hint. Reconsider Example 10.6.3 carefully.

Chapter 12
Hamilton-Jacobi Theory

In this chapter, we present the classical Hamilton-Jacobi theory. This theory has played an enormous role in the development of theoretical and mathematical physics. On the one hand, it builds a bridge between classical mechanics and other branches of physics, in particular, optics. On the other hand, it yields a link between classical and quantum theory. In Sect. 12.1 we start with deriving the Hamilton-Jacobi equation and give a proof of the classical Jacobi Theorem, which yields a powerful tool for solving the dynamical equations of a Hamiltonian system. We interpret the Hamilton-Jacobi equation geometrically as an equation for a Lagrangian submanifold of phase space[1] which is contained in the coisotropic submanifold given by a level set of the Hamiltonian. Using this geometric picture, one can extract a general method for solving initial value problems for arbitrary first order partial differential equations of the Hamilton-Jacobi type. This method is based on the fact that solutions are generated by the characteristics of the underlying Hamiltonian system. That is why this procedure is called the method of characteristics. It will be discussed in detail in Sect. 12.2. In Sect. 12.3 we generalize this method to the case of systems of partial differential equations of the Hamilton-Jacobi type.

It turns out that one can go beyond the case where a solution is generated by a single function on configuration space. This is interesting both from the mathematical and from the physical point of view. To do so, instead of single generating functions, one must consider families depending on additional parameters. Such families are called Morse families. In Sects. 12.4 and 12.5 we develop the theory of Morse families in a systematic way. In Sect. 12.6 we present the theory of critical points[2] of Lagrangian submanifolds in cotangent bundles, including a topological characterization in terms of the Maslov class and a description of the topological data in terms of generating Morse families.

In Sects. 12.7 and 12.8 we discuss applications in the spirit of geometric asymptotics. First, we study the short wave asymptotics in lowest order for the Helmholtz

[1] In this chapter, the phase space will always be the cotangent bundle of some configuration space.

[2] Points where the Lagrangian submanifold is not transversal to the fibres.

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7_12,

equation. This leads to the eikonal equation of geometric optics. We discuss classes of solutions of this equation including the formation of caustics. In Sect. 12.8, we study the transport equation. We present a detailed study of its geometry and, on this basis, derive first order short wave asymptotic solutions for a class of first-order partial differential equations. In this analysis a key role is played by a consistency condition of topological type, which for good reasons is called the Bohr-Sommerfeld quantization condition. We discuss applications to the Helmholtz and to the Schrödinger equations.

12.1 The Hamilton-Jacobi Equation

The basic idea of Hamilton-Jacobi theory consists in finding a time-dependent symplectomorphism transforming the system to equilibrium. Let (M, ω, H) be a Hamiltonian system and let $(\tilde{M}, \tilde{\omega}, \tilde{H})$ be its extension to the time-dependent phase space, that is,

$$\tilde{M} = \mathrm{T}^*\mathbb{R} \times M, \qquad \tilde{\omega} = \omega - \mathrm{d}E \wedge \mathrm{d}t, \qquad \tilde{H} = H - E,$$

see Sect. 9.3. Let $\tilde{\Sigma} = \tilde{H}^{-1}(0)$ and let $\tilde{\Phi}$ be a time-dependent canonical transformation of $\tilde{M}$. By Proposition 9.3.4, $\tilde{\Phi}$ induces a time-dependent canonical transformation of M, that is, a smooth mapping $\Phi : M \times \mathbb{R} \to M$ such that $\Phi(\cdot, t)$ is a canonical transformation of M for all t. In Darboux coordinates q^i and p_i on M, Φ is given by

$$(\mathbf{q}, \mathbf{p}, t) \mapsto \Phi(\mathbf{q}, \mathbf{p}, t) = \big(\bar{q}(\mathbf{q}, \mathbf{p}, t), \bar{p}(\mathbf{q}, \mathbf{p}, t), t\big).$$

It transforms the system to equilibrium iff the new variables $\bar{q}^i$ and $\bar{p}_i$ are constants of motion, that is,

$$\dot{\bar{q}}^i = 0, \qquad \dot{\bar{p}}_i = 0.$$

In this case, the Hamilton equations imply

$$\frac{\partial \bar{H}}{\partial \bar{q}^i} = 0, \qquad \frac{\partial \bar{H}}{\partial \bar{p}_i} = 0. \tag{12.1.1}$$

To find a local generating function S of the first kind for Φ, we must solve

$$\big(\bar{p}_i \mathrm{d}\bar{q}^i - \bar{H}\mathrm{d}t\big) - \big(p_i \mathrm{d}q^i - H\mathrm{d}t\big) = -\mathrm{d}S \tag{12.1.2}$$

on the graph $\Gamma_\Phi \subset (M \times \mathbb{R}) \times M$ of Φ, cf. Sect. 8.8. Comparison of coefficients yields the equations

$$\bar{p}_i = -\frac{\partial S}{\partial \bar{q}^i}, \qquad p_i = \frac{\partial S}{\partial q^i}, \qquad \bar{H} = H + \frac{\partial S}{\partial t} \tag{12.1.3}$$

on Γ_Φ. By (12.1.1), $\bar{H}$ depends on t only. It can, therefore, be absorbed into S. Then,

$$H + \frac{\partial S}{\partial t} = 0 \tag{12.1.4}$$

and by coordinatizing Γ_Φ by q^i and $\bar{q}^i$, from (12.1.3) we obtain

$$H\left(\mathbf{q}, \frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}, \bar{\mathbf{q}}, t), t\right) + \frac{\partial S}{\partial t}(\mathbf{q}, \bar{\mathbf{q}}, t) = 0, \tag{12.1.5}$$

where H stands for the local representative of the Hamiltonian function in the Darboux coordinates q^i, p_i. This is the time-dependent Hamilton-Jacobi equation. Here, $\bar{\mathbf{q}}$ plays the role of a parameter labelling the solutions. It can, therefore, be omitted. If H is not explicitly time-dependent, one can separate the time variable by means of the ansatz $S(\mathbf{q}, \bar{\mathbf{q}}, t) = S(\mathbf{q}, \bar{\mathbf{q}}) + T(t)$:

$$\frac{\mathrm{d}T}{\mathrm{d}t} = -H\left(\mathbf{q}, \frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}, \bar{\mathbf{q}})\right).$$

Since the right hand side of this equation does not depend on t, both sides must be equal to a constant, say c. Thus, we have $T(t) = -c(t - t_0)$ and

$$H\left(\mathbf{q}, \frac{\partial S}{\partial q}(\mathbf{q}, \bar{\mathbf{q}})\right) = c. \tag{12.1.6}$$

This is the time-independent Hamilton-Jacobi equation. Again, $\bar{\mathbf{q}}$ appears as a parameter labelling the solutions. Let us summarize. If the time-dependent canonical transformation generated by $S(\mathbf{q}, \bar{\mathbf{q}}, t)$ transforms the system to equilibrium, then $S(\mathbf{q}, \bar{\mathbf{q}}, t)$ fulfils the first order partial differential equation (12.1.5) for every $\bar{\mathbf{q}}$, that is, the $\bar{q}^i$ play the role of parameters for a family of solutions.[3] In the sequel, such a family will be called a complete integral for H, provided it fulfils a certain regularity condition. We will show that complete integrals of (12.1.5) are in one-to-one correspondence with solutions of the Hamilton equations ($2n$ ordinary differential equations of first order) for H. This is the famous Jacobi Theorem, yielding a powerful solution scheme. On the other hand, forgetting about the above derivation, we can view (12.1.5) as an equation defined by H for the function $S = S(\mathbf{q}, t)$. This type of equations occurs in various branches of physics, notably in optics. We will see that the initial value problem for this equation can be solved by means of the flow of the Hamiltonian vector field X_H. This is the method of characteristics, which will be discussed in detail below.

[3]The same statement holds true for a generating function $S(\mathbf{q}, \bar{\mathbf{p}}, t)$ of the second kind with parameters $\bar{p}_i$.

Remark 12.1.1

1. To understand the physical meaning of S, let us consider a mechanical system with configuration space $Q = \mathbb{R}^n$ and Hamiltonian function $H : \mathrm{T}^*Q \to \mathbb{R}$. Let $S = S(\mathbf{q}, t)$ be a solution of the associated Hamilton-Jacobi equation. Via S, every curve $t \mapsto \mathbf{q}(t)$ in Q generates a curve $t \mapsto (\mathbf{q}(t), \mathbf{p}(t))$ in T^*Q by
$$p_i(t) = \frac{\partial S}{\partial q^i}\big(\mathbf{q}(t)\big).$$
Using (12.1.4), for the total derivative of S along $\mathbf{q}(t)$ we find
$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} S\big(\mathbf{q}(t)\big) &= \frac{\partial S}{\partial q^i}\big(\mathbf{q}(t)\big)\dot{q}^i(t) + \frac{\partial S}{\partial t}\big(\mathbf{q}(t)\big) \\ &= p_i(t)\dot{q}^i(t) - H\big(\mathbf{q}(t), \mathbf{p}(t)\big) \\ &= L\big(\mathbf{q}(t), \dot{\mathbf{q}}(t)\big),\end{aligned} \tag{12.1.7}$$
with L denoting the Lagrange function of the system, cf. (9.1.15). Integrating both sides of (12.1.7) from t_0 to $t > t_0$, we obtain
$$S\big(\mathbf{q}(t), t\big) - S\big(\mathbf{q}(t_0), t_0\big) = \int_{t_0}^{t} L\big(\mathbf{q}(t), \dot{\mathbf{q}}(t)\big)\, \mathrm{d}t.$$
Thus, up to an additive constant, $S(q, t)$ coincides with the physical action function along $t \to \mathbf{q}(t)$, which is the projection of the integral curve to Q.
2. Obviously, one way to transform a Hamiltonian system (M, ω, H) to equilibrium is given by the flow Φ of the Hamiltonian vector field X_H, cf. Remark 8.2.5/1. In this case, the constants of motion $\bar{\mathbf{q}}$ and $\bar{\mathbf{p}}$ coincide with the initial values $\mathbf{q}(t_0)$ and $\mathbf{p}(t_0)$ for some given time $t = t_0$. Note that in a neighbourhood of the point
$$\big(\big(\mathbf{q}(t_0), \mathbf{p}(t_0), t_0\big), (\bar{\mathbf{q}}, \bar{\mathbf{p}}, t_0)\big) \in \Gamma_\Phi \subset \tilde{M} \times \tilde{M}$$
we have
$$\det\left(\frac{\partial p_i}{\partial \bar{p}_j}\right) \neq 0,$$
so that the coordinates q^i and $\bar{p}_i$ define a local chart on Γ_Φ. Thus, in this context it is reasonable to use a generating function of the second kind.

Now, we give the geometrical interpretation of the Hamilton-Jacobi equation. Let $M = \mathrm{T}^*Q$ and assume for simplicity that H does not explicitly depend on time; the general case is completely analogous. Under this assumption, we can confine our attention to the time-independent Hamilton-Jacobi equation (12.1.6). This equation can be rewritten as
$$H \circ \mathrm{d}S = c \tag{12.1.8}$$

with $c \in \mathbb{R}$ and with the parameter $\bar{\mathbf{q}}$ omitted. Assume that c is a regular value of H. Then, $H^{-1}(c)$ is an embedded submanifold of T^*Q of codimension 1. By Proposition 7.2.4/2, it is coisotropic. On the other hand, the image of the differential $\mathrm{d}S$, which will be denoted by the same symbol, is a Lagrangian submanifold. Thus, we arrive at the following geometric interpretation of the Hamilton-Jacobi equation: its solutions are Lagrangian submanifolds contained in a given coisotropic submanifold $\mathscr{C}$ of T^*Q on which the Hamiltonian vector field X_H has no zeros.

Furthermore, the integral curves of X_H can be interpreted geometrically as the characteristics of $\mathscr{C}$: to see this, recall from Sect. 8.5 that the characteristics of $\mathscr{C}$ are the integral manifolds of the characteristic distribution $D^{\omega\mathscr{C}}$. Since the latter has rank 1 and since X_H has no zeros on $\mathscr{C}$, Lemma 8.5.4/2 implies that $D^{\omega\mathscr{C}}$ is spanned by X_H. It follows that the integral curves of X_H coincide with the characteristics of $\mathscr{C}$, indeed.[4]

First, we show that the dynamics of an autonomous Hamiltonian system reduces to the image of the differential $\mathrm{d}S$ of a function S on Q iff this function solves the Hamilton-Jacobi equation.

Proposition 12.1.2 *Let* $(\mathrm{T}^*Q, \omega, H)$ *be an autonomous Hamiltonian system, with* Q *being connected, and let* $S\colon Q \to \mathbb{R}$ *be a smooth function. The following statements are equivalent.*

1. S *is a solution of the time-independent Hamilton-Jacobi equation* (12.1.8).
2. *The image of the differential* $\mathrm{d}S$ *is invariant under the flow of* X_H.

Proof For $x \in Q$ and $Y \in \mathrm{T}_x Q$, we have

$$\omega_{\mathrm{d}S(x)}\big(X_H, (\mathrm{d}S)'Y\big) = -\big\langle \mathrm{d}H, (\mathrm{d}S)'Y\big\rangle = -Y(H \circ \mathrm{d}S). \tag{12.1.9}$$

Since Q is connected, S solves (12.1.8) iff the right hand side of (12.1.9) vanishes for all $Y \in \mathrm{T}Q$. On the other hand, since $\mathrm{d}S$ is Lagrange, we have

$$(\mathrm{d}S)'\mathrm{T}Q = \mathrm{T}(\mathrm{d}S) = \big(\mathrm{T}(\mathrm{d}S)\big)^{\omega}.$$

Thus, X_H takes values in $\mathrm{T}(\mathrm{d}S)$, and hence $\mathrm{d}S$ is invariant under the flow of X_H, iff the left hand side of the above equation vanishes for all $Y \in \mathrm{T}Q$. □

Remark 12.1.3

1. That point 1 implies point 2 follows also from the more general statement of Proposition 8.5.3.
2. An analogous statement holds for explicitly time-dependent Hamiltonian functions, because every solution $S = S(\mathbf{q}, t)$ of the time-dependent Hamilton-Jacobi equation defines a Lagrangian submanifold $\mathrm{d}S$ of $\mathrm{T}^*(\mathbb{R} \times Q)$. We leave the details to the reader (Exercise 12.1.1).

[4] For the case where $\mathscr{C}$ is an energy surface, this has already been discussed in Sect. 9.1.

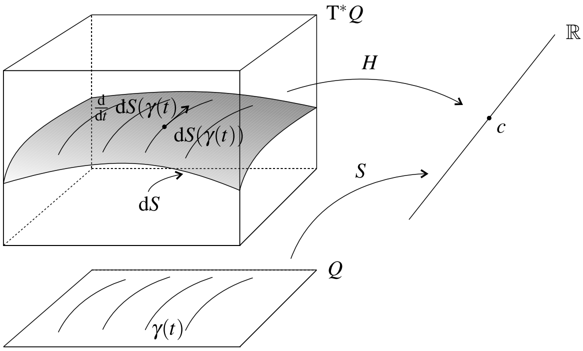


Fig. 12.1 Geometric meaning of a solution S of the time-independent Hamilton-Jacobi equation

Figure 12.1 illustrates the geometric content of Proposition 12.1.2: every solution S of the Hamilton-Jacobi equation yields a reduction of the dynamics of the Hamiltonian system to the Lagrangian submanifold $\mathrm{d}S$. This submanifold is a union of integral curves of the Hamiltonian vector field X_H. This way, we have obtained a fundamental relation between systems of first order ordinary differential equations (the Hamilton equations) and first order partial differential equations of the Hamilton-Jacobi type. On the one hand, this relation can be used to study the dynamics of Hamiltonian systems. On the other hand it can be used to solve initial value problems for partial differential equations of this type.

In the remainder of this section we discuss the first of these two aspects. The discussion is based on the notion of a complete integral. In order to keep in touch with the standard mechanics language, we give a definition in terms of local coordinates here. In Sect. 12.3, we will give a coordinate-free definition in a more general situation (Definition 12.3.3).

Definition 12.1.4 (Complete integral) Let (M, ω, H) be a $2n$-dimensional Hamiltonian system and let q^i, p_i be Darboux coordinates. An n-parameter[5] family $S(\mathbf{q}, \bar{\mathbf{q}}, t)$ of solutions of the time-dependent Hamilton-Jacobi equation (12.1.5), with H being expressed in terms of the Darboux coordinates q^i, p_i, is called a complete integral for H if

$$\det\left(\frac{\partial^2 S}{\partial q^i \partial \bar{q}^j}\right) \neq 0. \tag{12.1.10}$$

Remark 12.1.5 From this definition one derives the notion of a first integral for the time-independent Hamilton-Jacobi equation. Since S does not depend on t here, the conjugate variable E is a constant of motion which can be expressed in terms of the remaining constants. Therefore, one has only $n-1$ independent constants of motion a_j and Condition (12.1.10) is replaced by the requirement that the matrix

[5] The notation $\bar{q}^i$ for the parameters is a matter of convention, their physical meaning depends on the concrete context.

$$\frac{\partial^2 S}{\partial q^i \partial a^j}, \quad 1 \le i \le n, 1 \le j \le n-1, \tag{12.1.11}$$

be of maximal rank.

The following theorem states that finding a complete integral is equivalent to solving the Hamilton equations. Here, we give the formulation and the proof in local coordinates. In Sect. 12.3 we will come back to this theorem in a more general context. There, we will present a coordinate-free proof.

Theorem 12.1.6 (Jacobi) *Let (M, ω, H) be a $2n$-dimensional Hamiltonian system, let q^i, p_i be Darboux coordinates and let $S(\mathbf{q}, \bar{\mathbf{q}}, t)$ be a complete integral of the corresponding time-dependent Hamilton-Jacobi equation. Then, via the relations*

$$p_i = \frac{\partial S}{\partial q^i}(q, \bar{q}, t), \qquad \bar{p}_i = -\frac{\partial S}{\partial \bar{q}^i}(q, \bar{q}, t), \tag{12.1.12}$$

S defines a time-dependent canonical transformation $(q, p) \mapsto (\bar{q}, \bar{p})$, which assigns to every set of constants of motion $(\bar{\mathbf{q}}, \bar{\mathbf{p}})$ a solution of the Hamilton equations given in the coordinates q^i and p_i by

$$t \mapsto \big(q(\bar{\mathbf{q}}, \bar{\mathbf{p}}, t), p(\bar{\mathbf{q}}, \bar{\mathbf{p}}, t)\big). \tag{12.1.13}$$

Proof By (12.1.10), S defines via (12.1.12) a canonical transformation of the first kind, indeed. We show that the curve $t \mapsto (\mathbf{q}(t), \mathbf{p}(t))$ in $\mathbb{R}^{2n}$ given by (12.1.13) satisfies

$$\dot{q}^j(t) = \frac{\partial H}{\partial p_j}\big(\mathbf{q}(t), \mathbf{p}(t), t\big), \qquad \dot{p}_i(t) = -\frac{\partial H}{\partial q^i}\big(\mathbf{q}(t), \mathbf{p}(t), t\big). \tag{12.1.14}$$

By (12.1.12), the mappings $\mathbf{q}(t)$ and $\mathbf{p}(t)$ are defined by the relations

$$p_i(t) = \frac{\partial S}{\partial q^i}\big(\mathbf{q}(t), \bar{\mathbf{q}}, t\big), \qquad \bar{p}_i = -\frac{\partial S}{\partial \bar{q}^i}\big(\mathbf{q}(t), \bar{\mathbf{q}}, t\big) \tag{12.1.15}$$

where $\bar{\mathbf{q}}$ and $\bar{\mathbf{p}}$ are fixed. The second of these relations implies

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial S}{\partial \bar{q}^i}\big(\mathbf{q}(t), \bar{\mathbf{q}}, t\big)\right) = \frac{\partial^2 S}{\partial q^j \partial \bar{q}^i}\big(\mathbf{q}(t), \bar{\mathbf{q}}, t\big)\dot{q}^j(t) + \frac{\partial^2 S}{\partial t \partial \bar{q}^i}\big(\mathbf{q}(t), \bar{\mathbf{q}}, t\big).$$

By the Hamilton-Jacobi equation and by (12.1.15), the second term yields

$$-\frac{\partial}{\partial \bar{q}^i}\left(H\left(\mathbf{q}(t), \frac{\partial S}{\partial q^i}\big(\mathbf{q}(t), \bar{\mathbf{q}}, t\big), t\right)\right) = -\frac{\partial H}{\partial p_j}\big(\mathbf{q}(t), \mathbf{p}(t), t\big)\frac{\partial^2 S}{\partial q^j \partial \bar{q}^i}\big(\mathbf{q}(t), \bar{\mathbf{q}}, t\big).$$

In view of (12.1.10), this yields the first equation in (12.1.14). The first relation in (12.1.15) implies

$$\dot{p}_i(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial S}{\partial q^i}\big(\mathbf{q}(t), \bar{\mathbf{q}}, t\big)\right) = \frac{\partial^2 S}{\partial q^i \partial q^j}\big(\mathbf{q}(t), \bar{\mathbf{q}}, t\big)\dot{q}^j(t) + \frac{\partial^2 S}{\partial q^i \partial t}\big(\mathbf{q}(t), \bar{\mathbf{q}}, t\big).$$

For the second term, the Hamilton-Jacobi equation and (12.1.15) yield

$$
\begin{aligned}
&-\frac{\partial}{\partial q^i}\left(H\left(\mathbf{q}(t), \frac{\partial S}{\partial q^i}\big(\mathbf{q}(t), \bar{\mathbf{q}}, t\big), t\right)\right)\\
&\quad = -\frac{\partial H}{\partial q^i}\big(\mathbf{q}(t), \mathbf{p}(t), t\big) - \frac{\partial H}{\partial p_j}\big(\mathbf{q}(t), \mathbf{p}(t), t\big)\frac{\partial^2 S}{\partial q^j \partial q^i}\big(\mathbf{q}(t), \bar{\mathbf{q}}, t\big),
\end{aligned}
$$

where on the right hand side, the partial derivatives of H are taken in the coordinates q^i, p_i, t (as opposed to the coordinates q^i, $\bar{q}^i$, t on the left hand side). Thus, using the first equation of (12.1.14), which was already shown, we obtain the second equation in (12.1.14). □

The only systematic way for finding a complete integral is provided by the method of separation of variables. Above we have applied this method to separate the time variable in the (time-dependent) Hamilton-Jacobi equation for a Hamiltonian function H which does not explicitly depend on time, thus arriving at the time-independent Hamilton-Jacobi equation. It can also be applied to the other variables, as soon as the Hamiltonian function consists of a sum of terms which depend on disjoint sets of coordinates. This is illustrated by the following example.

Example 12.1.7 (Central Force Field) Let us find a complete integral for the motion of a particle in a central force field, given by a spherically symmetric potential V. In spherical coordinates r, ϑ, ϕ, the Hamiltonian takes the form

$$
H = \frac{1}{2m}\left(p_r^2 + \frac{p_\vartheta^2}{r^2} + \frac{p_\phi^2}{r^2 \sin^2\vartheta}\right) + V(r),
$$

cf. Example 11.5.2. Thus, the Hamilton-Jacobi equation reads

$$
\frac{1}{2m}\left(\left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial \vartheta}\right)^2 + \frac{1}{r^2\sin^2\vartheta}\left(\frac{\partial S}{\partial \phi}\right)^2\right) + V(r) = E.
$$

We plug in the separation ansatz $S(r, \vartheta, \phi) = S_r(r) + S_\vartheta(\vartheta) + S_\phi(\phi)$:

$$
\frac{1}{2m}\left(\left(\frac{\mathrm{d}S_r}{\mathrm{d}r}\right)^2 + \frac{1}{r^2}\left(\frac{\mathrm{d}S_\vartheta}{\mathrm{d}\vartheta}\right)^2 + \frac{1}{r^2\sin^2\vartheta}\left(\frac{\mathrm{d}S_\phi}{\mathrm{d}\phi}\right)^2\right) + V(r) = E. \qquad (12.1.16)
$$

Since $\frac{\mathrm{d}S_\phi}{\mathrm{d}\phi}$ is the only quantity depending on ϕ,

$$
\frac{\mathrm{d}S_\phi}{\mathrm{d}\phi} = \Phi
$$

for some integration constant $\Phi \in \mathbb{R}$. Then, (12.1.16) can be rewritten as

$$
r^2\left(\frac{\mathrm{d}S_r}{\mathrm{d}r}\right)^2 + 2mr^2\big(V(r) - E\big) = -\left(\frac{\mathrm{d}S_\vartheta}{\mathrm{d}\vartheta}\right)^2 - \frac{\Phi^2}{\sin^2\vartheta}.
$$

Since the left hand side depends on r only and the right hand side on ϑ,

$$\left(\frac{\mathrm{d}S_\vartheta}{\mathrm{d}\vartheta}\right)^2 + \frac{\Phi^2}{\sin^2\vartheta} = \Theta^2, \qquad \left(\frac{\mathrm{d}S_r}{\mathrm{d}r}\right)^2 = 2m\big(E - V(r)\big) - \frac{\Theta^2}{r^2}$$

with a further integration constant $\Theta \in \mathbb{R}$. The constants E, Θ, Φ play the role of the parameters $\bar{q}^i$ in Theorem 12.1.6. Thus,

$$p_r \equiv \frac{\partial S}{\partial r} = \frac{\mathrm{d}S_r}{\mathrm{d}r} = \sqrt{2m\big(E - V(r)\big) - \frac{\Theta^2}{r^2}},$$

$$p_\vartheta \equiv \frac{\partial S}{\partial \vartheta} = \frac{\mathrm{d}S_\vartheta}{\mathrm{d}\vartheta} = \sqrt{\Theta^2 - \frac{\Phi^2}{\sin^2\vartheta}},$$

$$p_\phi \equiv \frac{\partial S}{\partial \phi} = \frac{\mathrm{d}S_\phi}{\mathrm{d}\phi} = \Phi$$

and hence

$$S_r(r) = \int \sqrt{2m\big(E - V(r)\big) - \frac{\Theta^2}{r^2}}\,\mathrm{d}r,$$

$$S_\vartheta(\vartheta) = \int \sqrt{\Theta^2 - \frac{\Phi^2}{\sin^2\vartheta}}\,\mathrm{d}\vartheta,$$

$$S_\phi(\phi) = \Phi\phi.$$

Thus, we arrive at the complete integral

$$S(t, r, \phi, \vartheta; E, \Phi, \Theta) = -Et + \Phi \cdot \phi + S_r(r, E, \Theta) + S_\vartheta(\vartheta, \Phi, \Theta),$$

and the following equations describing the dynamics:

$$t_0 \equiv \bar{p}_E = -\frac{\partial S}{\partial E} = t - \frac{\partial S_r}{\partial E} = t - \int \frac{m}{\sqrt{2m(E - V(r)) - \frac{\Theta^2}{r^2}}}\mathrm{d}r, \tag{12.1.17}$$

$$\bar{p}_\Theta = -\frac{\partial S}{\partial \Theta} = -\frac{\partial S_r}{\partial \Theta} - \frac{\partial S_\vartheta}{\partial \Theta}, \tag{12.1.18}$$

$$\bar{p}_\Phi = -\frac{\partial S}{\partial \Phi} = -\frac{\partial S_\vartheta}{\partial \Phi} - \phi. \tag{12.1.19}$$

Exercises

12.1.1 Prove Proposition 12.1.2 for the time-dependent case.

12.1.2 Analyze Eqs. (12.1.17)–(12.1.19) for the case of the Kepler potential. Show that they yield the Kepler orbits.

12.1.3 Solve the Hamilton-Jacobi equation for the planar two-centre problem (Example 11.1.4).
Hint. Use elliptic coordinates.

12.2 The Method of Characteristics

In this section, we show how to solve the initial value problem for the Hamilton-Jacobi equation using the method of characteristics.

Let Q be a manifold of dimension n and let $H : \mathrm{T}^*Q \to \mathbb{R}$ be a smooth function for which 0 is a regular value. We consider the level set $\mathscr{C} := H^{-1}(0)$. Recall that the characteristic distribution $D^{\omega\mathscr{C}}$ of $\mathscr{C}$ is spanned by the Hamiltonian vector field X_H. Let D be an embedded submanifold of Q of dimension m and let $S_0 : D \to \mathbb{R}$ be a smooth function. Solving the initial value problem for the Hamilton-Jacobi equation defined by H in the analytic sense consists in finding a smooth function S which is defined on some neighbourhood of D and fulfils

$$H \circ \mathrm{d}S = 0, \qquad S_{\restriction D} = S_0. \tag{12.2.1}$$

To solve this problem, we consider the associated Hamiltonian system $(\mathrm{T}^*Q, \omega, H)$ and perform the following steps.

1. We determine the Lagrangian submanifold of T^*Q given by the canonical lift $\widehat{(D, S_0)}$ of the pair (D, S_0), cf. Example 8.3.8/4. According to the Transversal Mapping Theorem 1.8.2, if $\widehat{(D, S_0)}$ is transversal to $\mathscr{C}$, the intersection

$$\mathscr{S}_0 := \widehat{(D, S_0)} \cap \mathscr{C}$$

   is an embedded isotropic submanifold of dimension $n-1$ of T^*Q. If $\mathscr{S}_0$ is transversal in $\mathscr{C}$ to the integral curves of X_H, we say that $\mathscr{S}_0$ is in non-characteristic position and call it an admissible initial condition, or a submanifold of Cauchy data. For dimensional reasons, and since X_H has no zeros on $\mathscr{C}$, $\mathscr{S}_0$ is transversal to the integral curves of X_H iff

$$\mathrm{T}_\xi \mathscr{S}_0 \cap D^{\omega\mathscr{C}}_\xi = \{0\} \tag{12.2.2}$$

   for all $\xi \in \mathscr{S}_0$.
2. By means of the flow of X_H, from $\mathscr{S}_0$ we generate a Lagrangian immersion $\Psi : \Lambda \to \mathrm{T}^*Q$ satisfying $H \circ \Psi = 0$. Locally, this immersion induces Lagrangian submanifolds $\mathscr{S}$.
3. If $\mathscr{S}$ intersects the fibres of T^*Q transversally and at most once, Proposition 8.3.10 yields a local function S on Q satisfying $\mathrm{d}S = \mathscr{S}$ and hence (12.2.1).

We refer to the Lagrangian immersion (Λ, Ψ) as a generalized solution, to the Lagrangian submanifold $\mathscr{S}$ as a geometric solution and to the function S as an analytic solution of the initial value problem (12.2.1). Figure 12.2 illustrates the geometry of the problem.

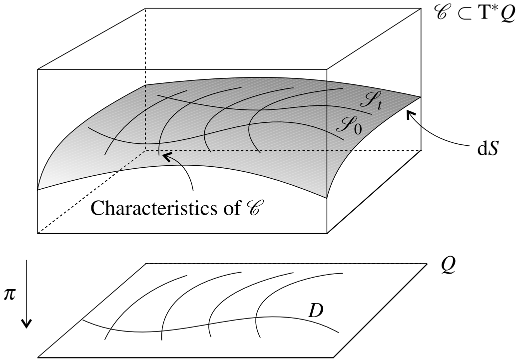


Fig. 12.2 Continuation of $\mathscr{S}_0$ by the help of the flow of X_H

Theorem 12.2.1 (Method of Characteristics) *Let $\mathscr{S}_0$ be an admissible initial condition for the initial value problem* (12.2.1)*. Let Φ be the flow of X_H and let $\mathscr{D} \subset \mathbb{R} \times \mathrm{T}^*Q$ be the domain of Φ. Then, the restriction Ψ of Φ to $\Lambda = (\mathbb{R} \times \mathscr{S}_0) \cap \mathscr{D}$ is a Lagrangian immersion. (Λ, Ψ) is a generalized solution of* (12.2.1).

Proof Obviously, Λ is an open subset of $\mathbb{R} \times \mathscr{S}_0$ and $\Psi : \Lambda \to \mathrm{T}^*Q$ is a smooth mapping. Since $\Psi(0, \xi) = \xi$ for all $\xi \in \mathscr{S}_0$, $\Psi(\Lambda)$ contains $\mathscr{S}_0$. Tangent vectors at $(t, \xi) \in \Lambda$ are of the form $(\lambda \frac{\mathrm{d}}{\mathrm{d}t}, Y)$, where $\lambda \in \mathbb{R}$, $\frac{\mathrm{d}}{\mathrm{d}t}$ denotes the standard vector field on $\mathbb{R}$ and $Y \in \mathrm{T}_\xi \mathscr{S}_0$. The tangent mapping of Ψ is given by

$$\Psi'_{(t,\xi)}\left(\lambda \frac{\mathrm{d}}{\mathrm{d}t}, Y\right) = (\Phi_t)'_\xi (\lambda X_H + Y). \tag{12.2.3}$$

Since X_H has no zeros on $\mathscr{C}$ and since $(\Phi_t)'_\xi$ is a bijection, (12.2.2) implies that Ψ is an immersion. To prove that (Λ, Ψ) is Lagrange, it suffices to show that it is isotropic: for $\lambda, \mu \in \mathbb{R}$ and $X, Y \in \mathrm{T}_\xi \mathscr{S}_0$ we calculate

$$\begin{aligned}
&\big(\Psi^*\omega\big)_{(t,\xi)}\left(\left(\lambda \frac{\mathrm{d}}{\mathrm{d}t}, X\right), \left(\mu \frac{\mathrm{d}}{\mathrm{d}t}, Y\right)\right) \\
&\quad = \omega_{\Phi_t(\xi)}\big((\Phi_t)'_\xi(\lambda X_H + X), (\Phi_t)'_\xi(\mu X_H + Y)\big) \\
&\quad = \omega_\xi(\lambda X_H + X, \mu X_H + Y) \\
&\quad = 0,
\end{aligned}$$

because Φ_t is symplectic, $\mathscr{S}_0$ is isotropic and X_H is characteristic for $\mathscr{C}$. □

Remark 12.2.2 Theorem 12.2.1 extends to the case of an arbitrary symplectic manifold M and an arbitrary isotropic and non-characteristic submanifold $\mathscr{S}_0$ of M.

Corollary 12.2.3 *Let (Λ, Ψ) be a generalized solution of the initial value problem* (12.2.1).

1. *There exists an open connected neighbourhood of* $\{0\} \times \mathscr{S}_0$ *in* Λ *on which* Ψ *is injective and hence defines a Lagrangian submanifold* $\mathscr{S}$ *which constitutes a local geometric solution.*
2. *The geometric solution is locally unique in the following sense. If* $\mathscr{S}_1$ *and* $\mathscr{S}_2$ *are solutions for the initial condition* $\mathscr{S}_0$, *there exists a connected subset of* $\mathscr{S}_1 \cap \mathscr{S}_2$ *which contains* $\mathscr{S}_0$ *and which is open in both* $\mathscr{S}_1$ *and* $\mathscr{S}_2$. *This follows from the fact that every solution is a union of open segments of integral curves of* X_H.

It may happen that the Lagrangian submanifold $\mathscr{S}$ obtained from $\mathscr{S}_0$ intersects some of the fibres of T^*Q several times, whereas other fibres are not intersected at all. It may also happen that $\mathscr{S}$ is tangent to a fibre. That is why the third step, the reconstruction of an analytic solution, again can be performed only locally. Using Proposition 8.3.10, we obtain

Proposition 12.2.4 *Let* $\mathscr{S}_0$ *be an admissible initial condition for the initial value problem* (12.2.1), *defined by the pair* (D, S_0). *Assume that* D *is contractible and that* $\mathscr{S}_0$ *is transversal to the fibres of* T^*Q *and intersects every fibre at most once. Let* $\mathscr{S}$ *be a geometric solution. Then, there exists a neighbourhood* U *of* D *in* Q *and a smooth function* $S : U \to \mathbb{R}$ *such that* $\mathrm{d}S(U) \subset \mathscr{S}$. *The function* S *is an analytic solution of the initial value problem* (12.2.1).

Proof Since $\mathscr{S}_0$ is transversal to the fibres of T^*Q and intersects every fibre at most once, there exists an open neighbourhood V of $\mathscr{S}_0$ in $\mathscr{S}$ for which this is still true. Since V is open in $\mathscr{S}$, it is a Lagrangian submanifold of T^*Q. Since $\mathscr{S}_0$ is contractible, we can choose V to be contractible as well. Then, Proposition 8.3.10 yields the existence of U and S. Since $\mathscr{S} \subset \mathscr{C}$, we have $H \circ \mathrm{d}S = 0$, that is, S is a solution of the Hamilton-Jacobi equation. To check the initial condition, we observe that $\mathscr{S}_0 = \mathrm{d}S(D)$. On the other hand, by the definition of the canonical lift in Example 8.3.8/4, for every $\xi \in \mathscr{S}_0$ and $X \in \mathrm{T}_{\pi(\xi)}D$ we have

$$\langle \xi, X \rangle = \langle \mathrm{d}S_0, X \rangle .$$

It follows that $\langle \mathrm{d}S(x), X \rangle = \langle \mathrm{d}S_0(x), X \rangle$ for all $x \in D$ and all $X \in \mathrm{T}_x D$ and hence $\mathrm{d}(S_{\restriction D}) = \mathrm{d}S_0$. By adding an appropriate constant we obtain $S_{\restriction D} = S_0$. □

Remark 12.2.5 Recall from the proof of Proposition 8.3.10 that an analytic solution S can be derived from $\mathscr{S}$ as follows: choose an open subset $V \subset \mathscr{S}$ which is transversal to the fibres of T^*Q and intersects each fibre at most once. Then, $U = \pi(V)$ is an open subset of Q and V defines a closed 1-form α on U by $\alpha \circ \pi = \mathrm{id}_V$. Now, the generating function S is determined from $\mathrm{d}S = \alpha$.

As we have seen, analytic solutions exist only on open subsets of Q over which the geometric solution is transversal to the fibres of T^*Q. In Sect. 12.4 we will show that every Lagrangian submanifold can be locally generated by a more general object, a so-called Morse family. In this sense, local analytic solutions always exist.

12.3 Generalized Hamilton-Jacobi Equations

In this section, we discuss the following natural generalization of the initial value problem for the Hamilton-Jacobi equation: instead of (12.2.1) we consider a system of partial differential equations,

$$C \circ \mathrm{d}S = 0, \tag{12.3.1}$$

defined by a smooth mapping $C\colon \mathrm{T}^*Q \to \mathbb{R}^k$ for which 0 is a regular value. If the submanifold $\mathscr{C} = C^{-1}(0) \subset \mathrm{T}^*Q$ is coisotropic, (12.3.1) is called a generalized Hamilton-Jacobi equation. By Remark 8.5.7/2, every coisotropic submanifold can be locally represented as a level set. Thus, even more generally, every coisotropic submanifold $\mathscr{C} \subset \mathrm{T}^*Q$ yields a generalized Hamilton-Jacobi equation.

Let $n = \dim Q$ and $k = \operatorname{codim} \mathscr{C}$. An admissible initial condition for (12.3.1) is an $(n-k)$-dimensional submanifold $\mathscr{S}_0$ of $\mathscr{C}$ which is isotropic in T^*Q and non-characteristic in $\mathscr{C}$, meaning that it is transversal to the characteristics of $\mathscr{C}$. For dimensional reasons, this is equivalent to the requirement

$$\mathrm{T}_\xi \mathscr{S}_0 \cap (\mathrm{T}_\xi \mathscr{C})^\omega = \{0\} \tag{12.3.2}$$

for all $\xi \in \mathscr{S}_0$. As in Sect. 12.2, we look for generalized solutions (Lagrangian immersions), geometric solutions (Lagrangian submanifolds) and analytic solutions (local functions on Q).

Theorem 12.3.1 *Let $1 \le k \le n-1$ and let $C\colon \mathrm{T}^*Q \to \mathbb{R}^k$ be a smooth mapping. Assume that 0 is a regular value of C and that the level set $\mathscr{C} = C^{-1}(0)$ is coisotropic. For every admissible initial condition $\mathscr{S}_0 \subset \mathscr{C}$, there exists a generalized solution of the generalized Hamilton-Jacobi equation* (12.3.1).

Proof Let $C_i : \mathrm{T}^*Q \to \mathbb{R}$ denote the components of the mapping C and let X_i denote the Hamiltonian vector fields generated by these functions. Since 0 is a regular value of C, the vector fields X_i are pointwise linearly independent on $\mathscr{C}$. By Proposition 8.5.6, they span the characteristic distribution $D^{\omega\mathscr{C}}$ of $\mathscr{C}$. Let Φ^i denote the flows of the X_i. There exists an open neighbourhood Λ of $\{0\} \times \mathscr{S}_0$ in $\mathbb{R}^k \times \mathscr{S}_0$ such that the mapping

$$\Psi : \Lambda \to \mathrm{T}^*Q, \qquad \Psi(\mathbf{t}, \xi) := \left(\Phi^1_{t_1} \circ \cdots \circ \Phi^k_{t_k}\right)(\xi) \tag{12.3.3}$$

is defined: indeed, the right hand side is defined on an open neighbourhood of $\{0\} \times \mathrm{T}^*Q$ in $\mathbb{R}^k \times \mathrm{T}^*Q$ and Λ can be obtained by intersecting this neighbourhood with $\mathbb{R}^k \times \mathscr{S}_0$. Since $\Psi(0, \xi) = \xi$ for all $\xi \in \mathscr{S}_0$, the image $\Psi(\Lambda)$ contains $\mathscr{S}_0$.

Next, we prove that Λ can be chosen so that Ψ is an immersion. For that purpose, it suffices to show that the tangent mapping of Ψ at $(0, \xi)$ is injective for all $\xi \in \mathscr{S}_0$. Writing tangent vectors of Λ at $(0, \xi)$ in the form $(\mathbf{x}, X)$ with $\mathbf{x} \in \mathbb{R}^k$ and $X \in \mathrm{T}_\xi \mathscr{S}_0$, we find

$$\Psi'_{(0,\xi)}(\mathbf{x}, X) = X + \sum_{i=1}^k x_i X_i(\xi). \tag{12.3.4}$$

Since the vector fields X_i are pointwise linearly independent on $\mathscr{S}_0$ and since (12.3.2) holds, $\Psi'_{(0,\xi)}$ is injective for all $\xi \in \mathscr{S}_0$, indeed.

Finally, we show that (Λ, Ψ) is Lagrange. Since Λ has dimension $k + \dim \mathscr{S}_0 = n$, it suffices to prove that Ψ is isotropic. For $\xi \in \mathscr{S}_0$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^k$ and $X, Y \in \mathrm{T}_\xi \mathscr{S}_0$, we have

$$\omega_\xi\left(X + \sum_{i=1}^k x_i X_i(\xi), Y + \sum_{i=1}^k y_i X_i(\xi)\right) = 0,$$

because the vector fields X_i are characteristic for $\mathscr{C}$ and $\mathscr{S}_0$ is isotropic. Hence, (12.3.4) implies that $\operatorname{im} \Psi'_{(0,\xi)}$ is an isotropic subspace of $\mathrm{T}_\xi(\mathrm{T}^* Q)$ for all $\xi \in \mathscr{S}_0$. Since $\Psi_\mathbf{t}$ is symplectic, to see that $\operatorname{im} \Psi'_{(\mathbf{t},\xi)}$ is isotropic for all $(\mathbf{t}, \xi) \in \Lambda$, it is sufficient to show that

$$\operatorname{im} \Psi'_{(\mathbf{t},\xi)} = (\Psi_\mathbf{t})'_\xi\big(\operatorname{im} \Psi'_{(0,\xi)}\big). \tag{12.3.5}$$

This can be seen as follows. On the one hand, since the vector fields X_i are characteristic, the mapping $\mathbf{t} \mapsto \Psi(\mathbf{t}, \xi)$ takes values in the integral manifold N_ξ of the characteristic distribution $D^{\omega\mathscr{C}}$ of $\mathscr{C}$ through ξ. Hence,

$$\operatorname{im} \Psi'_{(\mathbf{t},\xi)} = D^{\omega\mathscr{C}}_{\Psi(\mathbf{t},\xi)} + (\Psi_\mathbf{t})'_\xi \mathrm{T}_\xi \mathscr{S}_0.$$

On the other hand, for the same reason, the mapping $\Psi_\mathbf{t}$ restricts to a local diffeomorphism of N_ξ. Hence,

$$D^{\omega\mathscr{C}}_{\Psi(\mathbf{t},\xi)} = (\Psi_\mathbf{t})'_\xi D^{\omega\mathscr{C}}_\xi.$$

Since (12.3.4) implies $\operatorname{im} \Psi'_{(0,\xi)} = D^{\omega\mathscr{C}}_\xi + \mathrm{T}_\xi \mathscr{S}_0$, this proves (12.3.5) and hence the theorem. □

Remark 12.3.2

1. The geometric picture shown in Fig. 12.2 carries over to the present situation, but with modified dimensions of the submanifolds involved: $\mathscr{C}$ has dimension $2n - k$, $\mathscr{S}_0$ has dimension $n - k$, the characteristics of $\mathscr{C}$ have dimension k and thus the solution has again dimension n.
2. As for $k = 1$, with a generalized solution one can associate geometric solutions, that is, Lagrangian submanifolds. Moreover, one has an analogous uniqueness statement: if $\mathscr{S}_1$ and $\mathscr{S}_2$ are two geometric solutions of (12.3.1) for the initial condition $\mathscr{S}_0$, there exists a connected subset of $\mathscr{S}_1 \cap \mathscr{S}_2$ which contains $\mathscr{S}_0$ and which is open in both $\mathscr{S}_1$ and $\mathscr{S}_2$. This follows from Proposition 8.5.3.
3. If the initial condition $\mathscr{S}_0$ is transversal to the fibres of $\mathrm{T}^* Q$ and intersects each fibre at most once, there exists a local generating function, that is, an analytic solution.

In the remainder of this section, we consider the situation where $\mathscr{C}$ is a general coisotropic submanifold (not necessarily given by a family of functions).[6] In Sect. 12.1 we have proved the Jacobi Theorem, which states that finding a complete integral of the Hamilton-Jacobi equation is equivalent to solving the initial value problem for the Hamilton equations. Our aim is to prove that this statement is also true in the present context. Our presentation is along the lines of [181]. First, we must generalize the notion of a complete integral.

For that purpose, the space of the constants of motion $\bar{q}^i$, which in the case of the ordinary time-independent Hamilton-Jacobi equation has dimension $n-1$, is replaced by a manifold A of dimension $n-k$, called the parameter manifold. Thus, we build the $(2n-k)$-dimensional manifold $Q \times A$ and endow the cotangent bundle $\mathrm{T}^*(Q \times A) \cong \mathrm{T}^*Q \times \mathrm{T}^*A$ with the modified symplectic form

$$\mathrm{pr}_{\mathrm{T}^*Q}^* \, \mathrm{d}\theta_Q - \mathrm{pr}_{\mathrm{T}^*A}^* \, \mathrm{d}\theta_A, \qquad (12.3.6)$$

where θ_Q and θ_A are the canonical 1-forms on the corresponding cotangent bundles and $\mathrm{pr}_{\mathrm{T}^*Q}$ and $\mathrm{pr}_{\mathrm{T}^*A}$ are the canonical projections onto the factors T^*Q and T^*A, respectively. For an open subset $U \subset Q \times A$ and $x \in Q$, $a \in A$, we define

$$U_a := \{x \in Q \colon (x,a) \in U\}, \qquad U_x := \{a \in A \colon (x,a) \in U\}.$$

For a smooth function $S \colon U \to \mathbb{R}$, we denote

$$S_a \colon U_a \to \mathbb{R}, \qquad S_x \colon U_x \to \mathbb{R}, \qquad S_a(x) = S_x(a) := S(x,a),$$

and define the partial differentials

$$\mathrm{d}^Q S := \mathrm{pr}_{\mathrm{T}^*Q} \circ \mathrm{d}S \colon U \to \mathrm{T}^*Q, \qquad \mathrm{d}^A S := \mathrm{pr}_{\mathrm{T}^*A} \circ \mathrm{d}S \colon U \to \mathrm{T}^*A.$$

Then, for all $(x,a) \in U$ we have

$$\mathrm{d}^Q S(x,a) = \mathrm{d}S_a(x), \qquad \mathrm{d}^A S(x,a) = \mathrm{d}S_x(a).$$

Definition 12.3.3 (Complete integral) Let $\mathscr{C} \subset \mathrm{T}^*Q$ be a coisotropic submanifold of codimension k, $1 \le k \le n-1$. Let A be an $(n-k)$-dimensional manifold and let $U \subset Q \times A$ be a connected open subset. A function $S \colon U \to \mathbb{R}$ is called a complete integral of the generalized Hamilton-Jacobi equation (12.2.1) if

1. every nonempty subset U_a is connected and satisfies $\mathrm{d}S_a(U_a) \subset \mathscr{C}$,
2. the mapping $\mathrm{d}^Q S$ is a diffeomorphism from U onto an open subset of $\mathscr{C}$.

A is referred to as the parameter manifold of S and U is referred to as the domain of S. In case $d^Q S(U) = \mathscr{C}$, S is said to be a global complete integral.

[6] Theorem 12.3.1 extends to this more general situation. For the proof one has to apply the Tubular Neighbourhood Theorem for embedded submanifolds to $\mathscr{S}_0$, see Remark 6.4.7.

Remark 12.3.4

1. The functions S_a form a family of analytic solutions of the generalized Hamilton-Jacobi equation. The connected Lagrangian submanifolds $\mathrm{d}S_a(U_a)$ establish a foliation of the open subset $\mathrm{d}^Q S(U)$ of $\mathscr{C}$ by geometric solutions.
2. For the local discussion of solutions, condition 2 may be weakened to the requirement that the rank of $\mathrm{d}^Q S$ be $2n-k$. Let us analyze this requirement in Darboux coordinates $q^1, \ldots, q^n,\ p_1, \ldots, p^n$ on T^*Q and $\bar{q}^1, \ldots, \bar{q}^{n-k},\ \bar{p}_1, \ldots, \bar{p}_{n-k}$ on T^*A. In these coordinates, $\mathrm{d}^Q S$ is given by

$$(\mathbf{q}, \bar{\mathbf{q}}) \mapsto \left(\mathbf{q}, \mathbf{p} = \frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}, \bar{\mathbf{q}})\right).$$

It has rank $2n-k$ iff the matrix

$$\frac{\partial^2 S}{\partial q^i \partial \bar{q}^j}, \quad 1 \le i \le n,\ 1 \le j \le n-k, \tag{12.3.7}$$

has rank $n-k$. In particular, for $k=1$, this requirement is equivalent to (12.1.11). Let us add that $\mathrm{d}^A S$ is given by

$$(\mathbf{q}, \bar{\mathbf{q}}) \mapsto \left(\bar{\mathbf{q}}, \bar{\mathbf{p}} = \frac{\partial S}{\partial \bar{\mathbf{q}}}(\mathbf{q}, \bar{\mathbf{q}})\right)$$

and that it has rank $2(n-k)$ iff the matrix (12.3.7) has rank $n-k$. Thus, condition 2 implies that $\mathrm{d}^A S$ has rank $2(n-k)$ everywhere. It is, therefore, a submersion. In particular, $\mathrm{d}^A S(U)$ is an open subset of T^*A.

Theorem 12.3.5 (Generalized Jacobi Theorem) *Let $\mathscr{C}$ be a coisotropic submanifold of T^*Q and let S be a complete integral of the generalized Hamilton-Jacobi equation defined by $\mathscr{C}$, with parameter manifold A and domain U. Consider the mapping*

$$\varphi : U \to \mathrm{T}^*Q \times \mathrm{T}^*A, \qquad \varphi(x,a) := \left(\mathrm{d}^Q S(x,a), -\mathrm{d}^A S(x,a)\right).$$

Then, the following holds.

1. *$\Lambda = \varphi(U)$ is an embedded Lagrangian submanifold of $\mathrm{T}^*Q \times \mathrm{T}^*A$ endowed with the symplectic form (12.3.6). One has*

$$\mathrm{pr}_{\mathrm{T}^*Q}(\Lambda) = \mathrm{d}^Q S(U), \qquad \mathrm{pr}_{\mathrm{T}^*A}(\Lambda) = -\mathrm{d}^A S(U).$$

2. *There exists a unique mapping $\pi :\ \mathrm{pr}_{\mathrm{T}^*Q}(\Lambda) \to \mathrm{pr}_{\mathrm{T}^*A}(\Lambda)$ satisfying the relation $\pi \circ \mathrm{d}^Q S = -\mathrm{d}^A S$. This mapping is a surjective submersion and Λ is its graph.*
3. *The distribution $\ker \pi'$ coincides with the characteristic distribution of the coisotropic submanifold $\mathrm{pr}_{\mathrm{T}^*Q}(\Lambda)$ of T^*Q. In particular, for every $\alpha \in \mathrm{pr}_{\mathrm{T}^*A}(\Lambda)$, the connected components of $\pi^{-1}(\alpha)$ are characteristics of $\mathrm{pr}_{\mathrm{T}^*Q}(\Lambda)$.*

Proof 1. Λ is the image of the submanifold $\mathrm{d}S(U)$ of $\mathrm{T}^*Q \times \mathrm{T}^*A$ under the diffeomorphism $(\xi, \eta) \mapsto (\xi, -\eta)$. Since $\mathrm{d}S(U)$ is Lagrange with respect to the cotangent bundle symplectic form $\mathrm{pr}^*_{\mathrm{T}^*Q}\,\mathrm{d}\theta_Q + \mathrm{pr}^*_{\mathrm{T}^*A}\,\mathrm{d}\theta_A$, cf. Example 8.3.8/2, Λ is Lagrange with respect to the modified symplectic form (12.3.6). The relations $\mathrm{pr}_{\mathrm{T}^*Q}(\Lambda) = \mathrm{d}^Q S(U)$ and $\mathrm{pr}_{\mathrm{T}^*A}(\Lambda) = -\mathrm{d}^A S(U)$ are obvious.

2. Since the mapping $\mathrm{d}^Q S$ is a diffeomorphism onto its image, one can define a mapping

$$\pi:\ \mathrm{pr}_{\mathrm{T}^*Q}(\Lambda) \to \mathrm{pr}_{\mathrm{T}^*A}(\Lambda), \qquad \pi = -\mathrm{d}^A S \circ \left(\mathrm{d}^Q S\right)^{-1}.$$

This mapping is surjective and its graph, when viewed as a subset of $\mathrm{T}^*Q \times \mathrm{T}^*A$, coincides with Λ. By Remark 12.3.4/2, $\mathrm{d}^A S$ is a submersion. Hence, so is π.

3. Since Λ is the graph of π, under the natural identification of $\mathrm{T}(\mathrm{T}^*Q \times \mathrm{T}^*A)$ with $\mathrm{T}(\mathrm{T}^*Q) \times \mathrm{T}(\mathrm{T}^*A)$, tangent vectors of Λ correspond to pairs $(X, \pi'(X))$ with $X \in \mathrm{T}(\mathrm{pr}_{\mathrm{T}^*Q}(\Lambda))$. Since Λ is Lagrange with respect to the symplectic form (12.3.6), for all $\xi \in \mathrm{pr}_{\mathrm{T}^*Q}(\Lambda)$ and $X, Y \in \mathrm{T}_\xi(\mathrm{pr}_{\mathrm{T}^*Q}(\Lambda))$, we have

$$\begin{aligned} 0 &= \left(\mathrm{pr}^*_{\mathrm{T}^*Q}\,\mathrm{d}\theta_Q - \mathrm{pr}^*_{\mathrm{T}^*A}\,\mathrm{d}\theta_A\right)\left(\left(X, \pi'(X)\right), \left(Y, \pi'(Y)\right)\right) \\ &= \mathrm{d}\theta_Q(X, Y) - \mathrm{d}\theta_A\left(\pi'(X), \pi'(Y)\right). \end{aligned} \tag{12.3.8}$$

Thus, if $\pi'(X) = 0$, then $\mathrm{d}\theta_Q(X, Y) = 0$ for all Y and thus $X \lrcorner \mathrm{d}\theta_Q = 0$. Conversely, if $X \lrcorner \mathrm{d}\theta_Q = 0$, then $\mathrm{d}\theta_A(\pi'(X), \pi'(Y)) = 0$ for all Y. Since π'_ξ is surjective onto $\mathrm{T}_{\pi(\xi)}(\mathrm{T}^*A)$ and since $\mathrm{d}\theta_A$ is non-degenerate, we conclude $\pi'(X) = 0$. □

Remark 12.3.6

1. If S is a global complete integral, then $\mathrm{pr}_{\mathrm{T}^*Q}(\Lambda) = \mathscr{C}$.
2. Point 3 of Theorem 12.3.5 implies that for every $\alpha \in \mathrm{T}^*A$, the complete integral S yields a solution of the generalized Hamilton-Jacobi equation defined by the coisotropic submanifold $\mathscr{C}$. Geometrically, it is given by a connected component of $\pi^{-1}(\alpha)$. In particular, for every $a \in A$, the Lagrangian submanifold $\mathrm{d}S_a(U_a) \subset \mathrm{d}^Q S(U) \subset \mathscr{C}$ is foliated by the (k-dimensional) connected components of the level sets $\pi^{-1}(\alpha)$ with $\mathrm{pr}_{\mathrm{T}^*A}(\alpha) = a$. Thus, the elements of T^*A play the role of invariants labelling the solutions (up to connected components), and π is the mapping which assigns to every solution the corresponding invariants.

   In Darboux coordinates q^i, p_i on T^*Q and $\bar{q}^j$, $\bar{p}_j$ on T^*A, the Lagrangian submanifold Λ consists of the pairs $((\mathbf{q}, \mathbf{p}), (\bar{\mathbf{q}}, \bar{\mathbf{p}}))$ satisfying

$$\mathbf{p} = \frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}, \bar{\mathbf{q}}), \qquad \bar{\mathbf{p}} = -\frac{\partial S}{\partial \bar{\mathbf{q}}}(\mathbf{q}, \bar{\mathbf{q}}). \tag{12.3.9}$$

   The local representative of π is given by $(\mathbf{q}, \mathbf{p}) \mapsto (\bar{q}(\mathbf{q}, \mathbf{p}), \bar{p}(\mathbf{q}, \mathbf{p}))$, where the mapping $(\mathbf{q}, \mathbf{p}) \mapsto \bar{q}(\mathbf{q}, \mathbf{p})$ is obtained by solving the first equation in (12.3.9) for $\bar{\mathbf{q}}$ and the mapping $(\mathbf{q}, \mathbf{p}) \mapsto \bar{p}(\mathbf{q}, \mathbf{p})$ is obtained by inserting the resulting expression for $\bar{\mathbf{q}}$ into the second equation. Thus, Theorem 12.3.5 generalizes

the time-independent version of the ordinary Jacobi Theorem, obtained from Theorem 12.1.6 by separating the time variable and by removing from the $2n$ constants of motion $\bar{q}^i$, $\bar{p}_i$ the energy E and the initial time t_0. This way, the parameter manifold A generalizes the space of the constants of motion $\bar{q}^i$.
3. Equation (12.3.8) implies $\pi^*\mathrm{d}\theta_A = j^*\mathrm{d}\theta_Q$, where $j : \mathrm{pr}_{\mathrm{T}^*Q}(\Lambda) \to \mathrm{T}^*Q$ is the natural inclusion mapping. Thus, π defines a symplectic reduction, cf. Sect. 8.7.
4. If the rank of the restriction of the natural projection $\mathrm{T}^*Q \to Q$ to the isotropic submanifold $\mathscr{C}$ is less than n, a complete integral cannot exist. In this case, one can use the concept of Morse families, to be introduced in the next section. This leads to a further generalization of the Jacobi Theorem [40, 181].

For physical problems leading to a generalized Hamilton-Jacobi equation, the reader may consult the original papers by Tulczyjew and Benenti [40]. An interesting example of this type is provided by the motion of a charged particle, formulated in a gauge-invariant way in the 5-dimensional Kaluza-Klein space [290, 292]. Here, one has two equations, namely a mass and a charge condition, that is, one deals with the case of codimension 2.

12.4 Morse Families

In this section we show that the concept of a generating function for a fibre-transversal Lagrangian submanifold of a cotangent bundle can be generalized in such a way that it applies to any Lagrangian submanifold of this bundle. This leads to the notion of a Morse family, which generalizes that of a Morse function, cf. Sect. 8.9. Morse families were introduced by Hörmander [141], who called them phase functions. We also refer to [305] for the first completely intrinsic presentation and to [181], where in particular the relation to symplectic reduction is discussed. To our knowledge, in physics, Morse families have been used for the first time by Benenti and Tulczyjew [40].

Let B and Q be manifolds and let

$$\pi : B \to Q$$

be a submersion. Let $\pi_B : \mathrm{T}^*B \to B$ and $\pi_Q : \mathrm{T}^*Q \to Q$ be the canonical projections and let θ_B and θ_Q be the canonical 1-forms in T^*B and T^*Q, respectively. Consider the induced (pull-back) bundle

$$B \times_Q \mathrm{T}^*Q \equiv \pi^*\big(\mathrm{T}^*Q\big) = \big\{(b, \xi) \in B \times \mathrm{T}^*Q : \pi(b) = \pi_Q(\xi)\big\}.$$

Recall from Sect. 2.6 that $B \times_Q \mathrm{T}^*Q$ is a vector bundle over B with projection

$$\rho : B \times_Q \mathrm{T}^*Q \to B, \qquad \rho(b, \xi) := b,$$

and fibres $\rho^{-1}(b) = \mathrm{T}^*_{\pi(b)}Q$. By Proposition 2.6.1, $B \times_Q \mathrm{T}^*Q$ is an embedded submanifold of $B \times \mathrm{T}^*Q$ and the natural projection $B \times \mathrm{T}^*Q \to \mathrm{T}^*Q$ restricts to a vector bundle morphism

$$\varphi\colon B\times_Q \mathrm{T}^*Q\to \mathrm{T}^*Q,\qquad \varphi(b,\xi):=\xi, \tag{12.4.1}$$

which covers π,

$$\pi_Q\circ\varphi=\pi\circ\rho. \tag{12.4.2}$$

Since π is a submersion, $\ker\pi'$ is a vertical subbundle of $\mathrm{T}B$ (that is, a regular distribution on B) and the annihilator $(\ker\pi')^0$ is a vertical subbundle of T^*B, cf. Examples 2.7.7 and 2.7.8. We denote

$$\mathrm{V}B:=\ker\pi',\qquad \mathrm{V}^0B:=\big(\ker\pi'\big)^0.$$

In the present context, V^0B is often referred to as the conormal bundle associated with $\mathrm{V}B$. By Example 8.5.8/2, it is a coisotropic submanifold of T^*B.

Lemma 12.4.1 *The mapping*

$$\psi\colon B\times_Q \mathrm{T}^*Q\to \mathrm{T}^*B,\qquad \psi(b,\xi):=\big(\pi_b'\big)^{\mathrm{T}}(\xi) \tag{12.4.3}$$

is an injective vertical vector bundle morphism with image V^0B. *It satisfies*

$$\psi^*\theta_B=\varphi^*\theta_Q. \tag{12.4.4}$$

By Proposition 2.7.4, it follows that ψ is an embedding. Thus, as vector bundles,

$$B\times_Q \mathrm{T}^*Q\cong \mathrm{V}^0B. \tag{12.4.5}$$

Proof By definition, ψ preserves the fibres and is fibrewise linear. Moreover,

$$\pi_B\circ\psi=\rho, \tag{12.4.6}$$

so that ψ is vertical. Since the tangent mapping $\pi_b'\colon \mathrm{T}_bB\to \mathrm{T}_{\pi(b)}Q$ is surjective for every $b\in B$, the dual mapping $(\pi_b')^{\mathrm{T}}\colon \mathrm{T}^*_{\pi(b)}Q\to \mathrm{T}^*_bB$ is injective and one has

$$\mathrm{im}\big(\pi_b'\big)^{\mathrm{T}}=\big(\ker\big(\pi_b'\big)\big)^0=\mathrm{V}^0_bB.$$

Thus, ψ is an injective vector bundle morphism with image V^0B.

To prove (12.4.4), let $(b,\xi)\in B\times_Q \mathrm{T}^*Q$. We choose a section

$$\sigma\colon B\to B\times_Q \mathrm{T}^*Q$$

such that $\sigma(b)=\xi$ and consider the induced decomposition of the tangent space

$$\mathrm{T}_{(b,\xi)}\big(B\times_Q \mathrm{T}^*Q\big)=\ker\rho'_{(b,\xi)}\oplus\sigma'(\mathrm{T}_bB).$$

First, we show that both θ_B and θ_Q vanish on the vertical component $\ker\rho'_{(b,\xi)}$. For that purpose, take $Z\in\ker\rho'_{(b,\xi)}$. Then, (12.4.2) implies

$$\begin{aligned}
\left(\varphi^*\theta_Q\right)_{(b,\xi)}(Z) &= (\theta_Q)_{\varphi(b,\xi)}\left(\varphi'(Z)\right)\\
&= \left\langle \varphi(b,\xi), (\pi_Q\circ\varphi)'(Z)\right\rangle\\
&= \left\langle \varphi(b,\xi), \pi'\circ\rho'(Z)\right\rangle\\
&= 0
\end{aligned}$$

and (12.4.6) yields

$$\begin{aligned}
\left(\psi^*\theta_B\right)_{(b,\xi)}(Z) &= (\theta_B)_{\psi(b,\xi)}\left(\psi'(Z)\right)\\
&= \left\langle \psi(b,\xi), (\pi_B\circ\psi)'(Z)\right\rangle\\
&= \left\langle \psi(b,\xi), \rho'(Z)\right\rangle\\
&= 0.
\end{aligned}$$

Next, we prove that $\psi^*\theta_B$ and $\varphi^*\theta_Q$ coincide on the transversal component $\sigma'(\mathrm{T}_b B)$. Using $\pi_Q\circ\varphi\circ\sigma=\pi$ and $\pi_B\circ\psi\circ\sigma=\mathrm{id}_B$, for every $X\in\mathrm{T}_bB$ we have

$$\left(\varphi^*\theta_Q\right)_{(b,\xi)}\left(\sigma'X\right) = (\theta_Q)_{\varphi(b,\xi)}\left(\varphi'\circ\sigma'(X)\right) = \left\langle \xi, \pi'_Q\circ\varphi'\circ\sigma'(X)\right\rangle = \left\langle \xi, \pi'_b X\right\rangle$$

and

$$\begin{aligned}
\left(\psi^*\theta_B\right)_{(b,\xi)}\left(\sigma'X\right) &= (\theta_B)_{\psi(b,\xi)}\left(\psi'\circ\sigma'(X)\right)\\
&= \left\langle \left(\pi'_b\right)^{\mathrm{T}}\xi, \pi'_B\circ\psi'\circ\sigma'(X)\right\rangle\\
&= \left\langle \left(\pi'_b\right)^{\mathrm{T}}\xi, X\right\rangle.
\end{aligned}$$

□

Definition 12.4.2 (Morse family) Let $\pi: B\to Q$ be a submersion. A smooth function $S\colon B\to\mathbb{R}$ is called a Morse family along π if $\mathrm{d}S(B)$ is transversal to V^0B in T^*B, that is, if

$$\mathrm{T}_\eta\left(\mathrm{T}^*B\right) = \mathrm{T}_\eta\left(\mathrm{d}S(B)\right) + \mathrm{T}_\eta\left(\mathrm{V}^0B\right) \tag{12.4.7}$$

for all $\eta\in\mathrm{d}S(B)\cap\mathrm{V}^0B$. The triple (B,π,S) is referred to as a Morse family over Q.

Remark 12.4.3 The notion of a Morse family generalizes the notion of a Morse function. Indeed, if S is a Morse family, then $S_{\restriction\pi^{-1}(x)}$ is a Morse function for every $x\in Q$, because the intersection of the transversality condition (12.4.7) with $\mathrm{T}_\eta(\mathrm{T}^*\pi^{-1}(x))$ yields the transversality condition (8.9.1) for Morse functions. Thus, S yields a family of Morse functions, parameterized by the points of Q.

In the sequel, let us assume that the intersection $\mathrm{d}S(B)\cap\mathrm{V}^0B$ is nonempty. By Theorem 1.8.2, $\mathrm{d}S(B)\cap\mathrm{V}^0B$ is an embedded submanifold of T^*B and one has

$$\mathrm{T}_\eta\left(\mathrm{d}S(B)\cap\mathrm{V}^0B\right) = \mathrm{T}_\eta\left(\mathrm{d}S(B)\right)\cap\mathrm{T}_\eta\left(\mathrm{V}^0B\right).$$

Since $\mathrm{d}S(B) \cap \mathrm{V}^0 B$ is contained in the embedded Lagrangian submanifold $\mathrm{d}S(B)$ of T^*B, it is isotropic in T^*B. Since, in addition, the restriction of the canonical projection $\pi_B\colon \mathrm{T}^*B \to B$ to $\mathrm{d}S(B)$ is a diffeomorphism with inverse $\mathrm{d}S$, the projection

$$B_S := \pi_B\big(\mathrm{d}S(B) \cap \mathrm{V}^0 B\big)$$

is an embedded submanifold of B and the induced mapping

$$(\mathrm{d}S)_{\restriction B_S} : B_S \to \mathrm{d}S(B) \cap \mathrm{V}^0 B$$

is a diffeomorphism. Let us calculate the dimension of B_S. With $\dim B = n + k$ and $\dim Q = n$, (12.4.7) implies

$$\begin{aligned}\dim\big(\mathrm{d}S(B) \cap \mathrm{V}^0 B\big) &= \dim\big(\mathrm{d}S(B)\big) + \dim \mathrm{V}^0 B - \dim\big(\mathrm{T}^*B\big)\\ &= (n+k) + (2n+k) - 2(n+k)\\ &= \dim Q.\end{aligned}$$

Now, consider the vector bundle morphism $\psi : B \times_Q \mathrm{T}^*Q \to \mathrm{T}^*B$ of Lemma 12.4.1. Since it restricts to an isomorphism from $B \times_Q \mathrm{T}^*Q$ onto $\mathrm{V}^0 B$, there is a unique vector bundle morphism $\lambda\colon \mathrm{V}^0 B \to \mathrm{T}^*Q$ such that

$$\varphi = \lambda \circ \psi, \tag{12.4.8}$$

where ψ is understood as a mapping to $\mathrm{V}^0 B$. Obviously, λ covers π. Let

$$\lambda_S\colon \mathrm{d}S(B) \cap \mathrm{V}^0 B \to \mathrm{T}^*Q$$

be the restriction of λ to the isotropic submanifold $\mathrm{d}S(B) \cap \mathrm{V}^0 B$ and define

$$\Lambda_S := \lambda_S \circ (\mathrm{d}S)_{\restriction B_S}\colon B_S \to \mathrm{T}^*Q.$$

In the sequel, we will need the following two natural inclusion mappings:

$$i : \mathrm{d}S(B) \cap \mathrm{V}^0 B \to \mathrm{V}^0 B, \qquad j_0 : \mathrm{V}^0 B \to \mathrm{T}^*B.$$

Lemma 12.4.4 *Let (B, π, S) be a Morse family over Q.*

1. *We have*

$$\lambda^*\theta_Q = j_0^*\theta_B. \tag{12.4.9}$$

In particular, λ is a strict symplectic reduction.

2. *The mappings λ_S and Λ_S are Lagrangian immersions fulfilling*

$$\lambda_S^*\theta_Q = i^* \circ j_0^*(\theta_B), \qquad \Lambda_S^*\theta_Q = \mathrm{d}(S_{\restriction B_S}). \tag{12.4.10}$$

Proof 1. By (12.4.4), we have $\psi^* \circ \lambda^*(\theta_Q) = \psi^* \circ j_0^*(\theta_B)$. Since ψ is a diffeomorphism, this yields the assertion.

2. Using that $\mathrm{d}S(B)$ and $\mathrm{V}^0 B$ are transversal and that $\mathrm{d}S(B) \cap \mathrm{V}^0 B$ is isotropic, as well as (8.5.1), (8.7.2) and Proposition 7.2.1/5, for $\eta \in \mathrm{d}S(B) \cap \mathrm{V}^0 B$ we obtain

$$\begin{aligned}
\ker(\lambda_S)'_\eta &= \mathrm{T}_\eta\big(\mathrm{d}S(B) \cap \mathrm{V}^0 B\big) \cap \ker \lambda'_\eta \\
&= \mathrm{T}_\eta\big(\mathrm{d}S(B)\big) \cap \mathrm{T}_\eta\big(\mathrm{V}^0 B\big) \cap \big(\mathrm{T}_\eta\big(\mathrm{V}^0 B\big)\big)^{\omega_B} \\
&= \mathrm{T}_\eta\big(\mathrm{d}S(B)\big) \cap \big(\mathrm{T}_\eta\big(\mathrm{V}^0 B\big)\big)^{\omega_B} \\
&= \big(\mathrm{T}_\eta\big(\mathrm{d}S(B)\big) + \mathrm{T}_\eta\big(\mathrm{V}^0 B\big)\big)^{\omega_B} \\
&= \big(\mathrm{T}_\eta\big(\mathrm{T}^* B\big)\big)^{\omega_B} \\
&= \{0\}.
\end{aligned}$$

Thus, λ_S is an immersion. Since $\mathrm{d}S(B) \cap \mathrm{V}^0 B$ has the dimension of Q, to see that λ_S is Lagrange, it suffices to show that it is isotropic. Indeed, using point 1 and $\lambda_S = \lambda \circ i$, we find

$$\lambda_S^*(\theta_Q) = i^* \circ j_0^*(\theta_B) \tag{12.4.11}$$

and hence $\lambda_S^*(\mathrm{d}\theta_Q) = i^* \circ j_0^*(\mathrm{d}\theta_B) = 0$, because $\mathrm{d}S(B) \cap \mathrm{V}^0 B$ is isotropic. Finally, applying $((\mathrm{d}S)_{\restriction B_S})^*$ to (12.4.11) and using $j_0 \circ i \circ (\mathrm{d}S)_{\restriction B_S} = \mathrm{d}S \circ k$, where $k : B_S \to B$ denotes the natural inclusion mapping, we obtain

$$\Lambda_S^* \theta_Q = k^* \circ (\mathrm{d}S)^*(\theta_B) = k^* \mathrm{d}S = \mathrm{d}(S_{\restriction B_S}).$$

In the last step we have used point 1 of Remark 8.3.3. □

Definition 12.4.5 (Fibre-critical submanifold) Let (B, π, S) be a Morse family over Q. The embedded submanifold $B_S \subset B$ is called the fibre-critical submanifold of S. The mapping $\Lambda_S \colon B_S \to \mathrm{T}^* Q$ is called the Lagrangian immersion generated by S.

Remark 12.4.6 We give a local description of Morse families. Let x^i, $i = 1, \dots, n$ and y^α, $\alpha = 1, \dots, r$ be local coordinates on B adapted to π, that is, π is given by

$$\pi(\mathbf{x}, \mathbf{y}) = \mathbf{x}.$$

Denote the corresponding fibre coordinates in $\mathrm{T}^* B$ by p_i, $\tilde{p}_\alpha$ and in $\mathrm{T}^* Q$ by p_i. Then, the conormal bundle $\mathrm{V}^0 B$ is defined by the condition

$$\tilde{p}_\alpha = 0, \quad \alpha = 1, \dots, r,$$

and the morphism λ has the form

$$\lambda(\mathbf{x}, \mathbf{y}, \mathbf{p}, 0) = (\mathbf{x}, \mathbf{p}).$$

The mapping $\mathrm{d}S$ is given by

$$\mathrm{d}S(\mathbf{x},\mathbf{y}) = \left(\mathbf{x},\mathbf{y},\frac{\partial S}{\partial \mathbf{x}}(\mathbf{x},\mathbf{y}),\frac{\partial S}{\partial \mathbf{y}}(\mathbf{x},\mathbf{y})\right).$$

Thus, the submanifold $\mathrm{d}S(B)\cap \mathrm{V}^0 B$ is defined by the $(n+2r)$ equations

$$\frac{\partial S}{\partial x^i}(\mathbf{x},\mathbf{y}) = p_i, \qquad \frac{\partial S}{\partial y^\alpha}(\mathbf{x},\mathbf{y}) = 0, \qquad \tilde{p}_\alpha = 0, \tag{12.4.12}$$

and the fibre-critical submanifold B_S is given by the r equations

$$\frac{\partial S}{\partial y^\alpha}(\mathbf{x},\mathbf{y}) = 0. \tag{12.4.13}$$

In this language, transversality of $\mathrm{d}S(B)$ and $\mathrm{V}^0 B$ means that the $(n+r)\times r$-matrix

$$\left(\frac{\partial^2 S}{\partial x^i \partial y^\alpha},\frac{\partial^2 S}{\partial y^\beta \partial y^\alpha}\right)$$

must have rank r at all points $(\mathbf{x},\mathbf{y})$ fulfilling (12.4.13) (Exercise 12.4.2). The Lagrangian immersion Λ_S is then given by

$$\Lambda_S(\mathbf{x},\mathbf{y}) = \left(\mathbf{x},\frac{\partial S}{\partial \mathbf{x}}(\mathbf{x},\mathbf{y})\right), \tag{12.4.14}$$

with $(\mathbf{x},\mathbf{y})$ fulfilling (12.4.13). Finally, in terms of the induced fibre coordinates $\dot{x}^i$, $\dot{y}^\alpha$ on $\mathrm{T}B$ and $\dot{x}^i$, $\dot{p}_i$ on $\mathrm{T}(\mathrm{T}^*Q)$, the tangent space $\mathrm{T}_{(\mathbf{x},\mathbf{y})}B_S \subset \mathrm{T}_{(\mathbf{x},\mathbf{y})}B$ is given by

$$\frac{\partial^2 S}{\partial x^i \partial y^\alpha}(\mathbf{x},\mathbf{y})\dot{x}^i + \frac{\partial^2 S}{\partial y^\alpha \partial y^\beta}(\mathbf{x},\mathbf{y})\dot{y}^\beta = 0 \tag{12.4.15}$$

and the tangent space $\mathrm{T}_{\Lambda_S(\mathbf{x},\mathbf{y})}(\Lambda_S(B_S)) \subset \mathrm{T}_{\Lambda_S(\mathbf{x},\mathbf{y})}(\mathrm{T}^*Q)$ consists of the pairs $(\dot{\mathbf{x}},\dot{\mathbf{p}})$, where

$$\dot{p}_i = \frac{\partial^2 S}{\partial x^i \partial x^j}(\mathbf{x},\mathbf{y})\dot{x}^j + \frac{\partial^2 S}{\partial x^i \partial y^\alpha}(\mathbf{x},\mathbf{y})\dot{y}^\alpha,$$

and $\dot{x}^i$, $\dot{y}^\alpha$ fulfil (12.4.15).

Example 12.4.7 Let $B=\mathbb{R}^2$ and $Q=\mathbb{R}$ and let x, y, p_x, p_y and x, p_x be global Darboux coordinates on $\mathrm{T}^*B=\mathbb{R}^4$ and $\mathrm{T}^*Q=\mathbb{R}^2$, respectively. We consider the surjective submersion

$$\pi:\mathbb{R}^2\to\mathbb{R}, \qquad \pi(x,y)=x$$

and the smooth function

$$S:\mathbb{R}^2\to\mathbb{R}, \qquad S(x,y):=\frac{1}{3}y^3+\left(x^2-1\right)y. \tag{12.4.16}$$

The conormal bundle $V^0 B$ is defined by the condition $p_y = 0$ and the morphism λ has the form

$$\lambda(x, y, p_x, 0) = (x, p_x).$$

The mapping $\mathrm{d}S$ is given by

$$\mathrm{d}S(x, y) = \left(x, y, \frac{\partial S}{\partial x}(x, y), \frac{\partial S}{\partial y}(x, y)\right) = \big(x, y, 2xy, x^2 + y^2 - 1\big).$$

Consequently, the submanifold $\mathrm{d}S(B) \cap V^0 B$ is given by

$$2xy = p_x, \qquad x^2 + y^2 = 1, \qquad p_y = 0$$

and the fibre-critical submanifold B_S by

$$x^2 + y^2 = 1. \tag{12.4.17}$$

Thus, B_S is the unit sphere in $\mathbb{R}^2$. To check the transversality condition, we calculate

$$\left(\frac{\partial^2 S}{\partial x \partial y}(x, y), \frac{\partial^2 S}{\partial y^2}(x, y)\right) = (2x, 2y).$$

This matrix has rank 1 for all (x, y) which fulfil equation (12.4.17). Therefore, S is a Morse family. The Lagrangian immersion $\Lambda_S = \lambda_S \circ \mathrm{d}S$ is given by

$$\Lambda_S(x, y) = \left(x, \frac{\partial S}{\partial x}(x, y)\right) = (x, 2xy),$$

with (x, y) fulfilling (12.4.17). Parameterizing B_S by

$$(x, y) = \big(\cos(2\pi t), \sin(2\pi t)\big)$$

with $t \in \mathbb{R}$, we obtain the following parameterization of Λ_S:

$$\Lambda_S\big(\cos(2\pi t), \sin(2\pi t)\big) = \big(\cos(2\pi t), \sin(4\pi t)\big).$$

We see that the image is a figure eight immersion with the self intersection point $\Lambda_S(0, 1) = \Lambda_S(0, -1)$. Thus, (B_S, Λ_S) is locally, but not globally, a submanifold.[7]

This example can be generalized as follows. Let

$$\pi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n, \qquad \pi(\mathbf{x}, y) = \mathbf{x},$$

let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function for which 0 is a regular value, and let

$$S : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}, \qquad S(\mathbf{x}, y) := \frac{1}{3} y^3 + f(\mathbf{x}) y.$$

[7] According to Example 1.6.6, $\Lambda_S(B_S)$ can nevertheless be equipped with a submanifold structure, though in two inequivalent ways. Both of them are Lagrangian.

We encourage the reader to analyze this class of examples along the lines of the above discussion (Exercise 12.4.3).

Example 12.4.8 Let $f : Q \to \mathbb{R}^r$ be a smooth mapping for which 0 is a regular value, let $P := f^{-1}(0)$ and let $F : P \to \mathbb{R}$ be smooth function. We show that the canonical lift[8] $\widehat{(P, F)}$ of the pair (P, F) coincides with the image of the Lagrangian immersion Λ_S generated by the Morse family

$$S\colon Q \times \mathbb{R}^r \to \mathbb{R}, \qquad S(x, \mathbf{y}) := f(x) \cdot \mathbf{y} + F(x),$$

along the natural projection $\pi : Q \times \mathbb{R}^r \to Q$. Let $\xi \in \mathrm{T}^*Q$. To see that $\xi \in \Lambda_S(B_S)$ iff $\xi \in \widehat{(P, F)}$, we choose coordinates x^i on Q at $\pi_Q(\xi)$. By Remark 12.4.6, the fibre-critical submanifold B_S of S is given by

$$\frac{\partial S}{\partial y^\alpha}(\mathbf{x}, \mathbf{y}) = f_\alpha(\mathbf{x}) = 0.$$

Hence, it coincides with $P \times \mathbb{R}^r$. On B_S, we have

$$\left(\frac{\partial^2 S}{\partial x^i \partial y^\alpha}(\mathbf{x}, \mathbf{y}), \frac{\partial^2 S}{\partial y^\alpha \partial y^\beta}(\mathbf{x}, \mathbf{y})\right) = \left(\frac{\partial f_\alpha}{\partial x^i}(\mathbf{x}), 0\right).$$

Since this matrix has rank r, S is a Morse family along π and thus B_S is an embedded submanifold of $Q \times \mathbb{R}^r$. The Lagrangian immersion Λ_S generated by S is given in coordinates by

$$\Lambda_S(\mathbf{x}, \mathbf{y}) = \left(\mathbf{x}, y^\alpha \frac{\partial f_\alpha}{\partial \mathbf{x}}(\mathbf{x}) + \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x})\right).$$

Hence, its image consists of the points ξ in T^*Q whose coordinates $(\mathbf{x}, \mathbf{p})$ satisfy

$$\mathbf{p} = y^\alpha \frac{\partial f_\alpha}{\partial \mathbf{x}}(\mathbf{x}) + \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}), \qquad f(\mathbf{x}) = 0$$

for some $\mathbf{y} \in \mathbb{R}^r$. On the other hand, every $X \in \mathrm{T}_{\pi_Q(\xi)}P$ fulfils $\frac{\partial f_\alpha}{\partial x^i}(\mathbf{x}) X^i = 0$, where $X = X^i \partial_{x^i}$. Thus, we obtain

$$\langle \xi, X \rangle = X^i \left(y^\alpha \frac{\partial f_\alpha}{\partial x^i}(\mathbf{x}) + \frac{\partial F}{\partial x^i}(\mathbf{x})\right) = X^i \frac{\partial F}{\partial x^i}(\mathbf{x}) = \big\langle \mathrm{d}F\big(\pi_Q(\xi)\big), X\big\rangle,$$

so that $\Lambda_S(B_S) = \widehat{(P, F)}$, indeed.

We have seen above that every Morse family for which $\mathrm{d}S(Q) \cap \mathrm{V}^0 B$ is nonempty yields a Lagrangian embedding and hence, locally, Lagrangian submanifolds of T^*Q. The following theorem implies that, conversely, every Lagrangian

[8] See Example 8.3.8/4.

submanifold of a cotangent bundle can be locally obtained as the image of a Lagrangian immersion generated by a Morse family. The theorem belongs to Hörmander [141] and Maslov [197], see also [115, 181, 305].

Let $\pi_L : \mathrm{T}^*L \to L$ be the canonical projection and denote

$$\mathrm{V}(\mathrm{T}^*Q) := \ker(\pi'_Q) \subset \mathrm{T}(\mathrm{T}^*Q).$$

Theorem 12.4.9 *Let (L, ι) be an embedded Lagrangian submanifold of $(\mathrm{T}^*Q, \mathrm{d}\theta_Q)$ and let θ_L denote the canonical 1-form on T^*L. Assume that the following conditions hold.*

1. *The 1-form $\iota^*\theta_Q$ on L is exact.*
2. *There exists a Lagrangian subbundle F of $(\mathrm{T}(\mathrm{T}^*Q))_{\restriction L}$ which is transversal to both $\mathrm{T}L$ and $(\mathrm{V}(\mathrm{T}^*Q))_{\restriction L}$.*

Then, there exists a Morse family (B, π, S) over Q such that (B_S, Λ_S) is an embedded submanifold equivalent to (L, ι).

Proof Our proof is along the lines of Weinstein [305] and Libermann and Marle [181].

Assume that conditions 1 and 2 are fulfilled. Then, F is a Lagrangian complement of $\mathrm{T}L$ in the symplectic vector bundle $(\mathrm{T}(\mathrm{T}^*Q))_{\restriction L}$ and Theorem 8.6.4 yields a symplectomorphism Φ of an open neighbourhood U of L in T^*Q onto an open neighbourhood $\Phi(U)$ of the zero section s_0 in T^*L which maps L to s_0 and satisfies

$$\Phi'_\xi(F_\xi) = \mathrm{T}_{\Phi(\xi)}\big(\mathrm{T}^*_\xi L\big) = \ker(\pi_L)'_{\Phi(\xi)} \tag{12.4.18}$$

for all $\xi \in L$. By condition 2, $\mathrm{V}_\xi(\mathrm{T}^*Q)$ is a complement of F_ξ in $\mathrm{T}_\xi(\mathrm{T}^*Q)$ for every $\xi \in L$. Hence, (12.4.18) implies that $\Phi'_\xi(\mathrm{V}_\xi(\mathrm{T}^*Q))$ is a complement of $\ker(\pi_L)'_{\Phi(\xi)}$ for all $\xi \in L$. By shrinking U we can achieve that this remains true for all $\xi \in U$. Then,

$$\pi'_L \circ \Phi'\big(\mathrm{V}_\xi\big(\mathrm{T}^*Q\big)\big) = (\pi_L)'_{\Phi(\xi)}\big(\mathrm{T}_{\Phi(\xi)}\big(\mathrm{T}^*L\big)\big) = \mathrm{T}_\xi L \tag{12.4.19}$$

for all $\xi \in U$. Since $\Phi \circ \iota = s_0$ and $\pi_L \circ \Phi \circ \iota = \mathrm{id}_L$, the 1-form α on U defined by

$$\alpha := \theta_Q - (\iota \circ \pi_L \circ \Phi)^*\theta_Q - \Phi^*\theta_L \tag{12.4.20}$$

satisfies $\iota^*\alpha = 0$. Then, the generalized Poincaré Lemma 4.3.14 implies that U can be shrunk so that $\alpha = \mathrm{d}h$ for some smooth function h on U satisfying $h \circ \iota = 0$. Moreover, by condition 1, there exists a smooth function f on L with $\iota^*\theta_Q = \mathrm{d}f$. We choose $B = U \subset \mathrm{T}^*Q$ and $\pi := (\pi_Q)_{\restriction U} : U \to Q$ and define $S\colon U \to R$ by

$$S = h + \Phi^* \circ \pi_L^* f.$$

Then, $S \circ \iota = f$ and

$$\mathrm{d}S = \theta_Q - \Phi^*\theta_L. \tag{12.4.21}$$

We show that S defines a Morse family along π which satisfies $\Lambda_S(B_S) = \iota(L)$. Let us start with proving the latter relation. One can check that the vector bundle morphism $\lambda : \mathrm{V}^0 B \to \mathrm{T}^* Q$ is given by the restriction to $\mathrm{V}^0 B$ of the canonical projection $\mathrm{T}^*(\mathrm{T}^* Q) \to \mathrm{T}^* Q$. Hence, $\Lambda_S = \lambda \circ (\mathrm{d}S)_{\restriction B_S}$ is given by the natural inclusion mapping of $B_S \subset B \subset \mathrm{T}^* Q$ and thus we have to show that $B_S = L$. Let $\xi \in B$. By definition of B_S, $\xi \in B_S$ iff $\langle \mathrm{d}S(\xi), Z\rangle = 0$ for all $Z \in \mathrm{V}_\xi B$. By (12.4.21), for $Z \in \mathrm{V}_\xi B$, we have

$$\big\langle \mathrm{d}S(\xi), Z\big\rangle = -\big\langle \big(\Phi^* \theta_L\big)_\xi, Z\big\rangle = -\big\langle \Phi(\xi), \pi'_L \circ \Phi'(Z)\big\rangle. \tag{12.4.22}$$

Hence, if $\xi \in L$, then $\Phi(\xi)$ belongs to the zero section. Then, $\langle \mathrm{d}S(\xi), Z\rangle = 0$ and thus $\xi \in B_S$. Conversely, if $\xi \in B_S$, then $\langle \mathrm{d}S(\xi), Z\rangle = 0$ for all $Z \in \mathrm{V}_\xi B$ and (12.4.19) implies that $\Phi(\xi)$ belongs to the zero section of $\mathrm{T}^* L$. This implies $\xi \in L$.

It remains to show that $\mathrm{d}S(B)$ is transversal to $\mathrm{V}^0 B$, that is, that (12.4.7) holds for all $\eta = \mathrm{d}S(\xi)$, $\xi \in L$. Since by condition 2, F_ξ is transversal to $\mathrm{T}_\xi L$ in $\mathrm{T}_\xi(\mathrm{T}^* Q) = \mathrm{T}_\xi B$, we have

$$\mathrm{T}_{\mathrm{d}S(\xi)}\big(\mathrm{d}S(B)\big) = (\mathrm{d}S)'_\xi(\mathrm{T}_\xi B) = (\mathrm{d}S)'_\xi(\mathrm{T}_\xi L + F_\xi).$$

By injectivity of $(\mathrm{d}S)'_\xi$, the image $(\mathrm{d}S)'_\xi(F_\xi)$ is a subspace of $\mathrm{T}_{\mathrm{d}S(\xi)}(\mathrm{d}S(B))$ of dimension n. Since $\mathrm{T}_{\mathrm{d}S(\xi)}(\mathrm{V}^0 B)$ has dimension $3n$ and $\mathrm{T}_{\mathrm{d}S(\xi)}(\mathrm{T}^* B)$ has dimension $4n$, it therefore suffices to show that

$$(\mathrm{d}S)'_\xi(F_\xi) \cap \mathrm{T}_{\mathrm{d}S(\xi)}\big(\mathrm{V}^0 B\big) = \{0\}. \tag{12.4.23}$$

For that purpose, let $X \in F_\xi$ such that $(\mathrm{d}S)'_\xi X \in \mathrm{T}_{\mathrm{d}S(\xi)}(\mathrm{V}^0 B)$ and let γ be a curve through ξ representing X. By (12.4.18), γ may be chosen so that $\Phi(\gamma(t))$ is contained in $\mathrm{T}^*_\xi L$. As a first step, we show that $(\mathrm{d}S)'_\xi X \in \mathrm{T}_{\mathrm{d}S(\xi)}(\mathrm{V}^0 B)$ implies that

$$\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0}\big\langle \mathrm{d}S \circ \gamma(t), Z \circ \gamma(t)\big\rangle = 0 \tag{12.4.24}$$

for all vector fields Z on B taking values in $\mathrm{V}B$. Indeed, we may view Z as a function $Z : \mathrm{T}^* B \to \mathbb{R}$ and we may assume that it is a submersion. Then, $\mathrm{V}^0 B$ is contained in the submanifold $Z^{-1}(0)$ of $\mathrm{T}B$ and hence

$$\mathrm{T}_{\mathrm{d}S(\xi)}\big(\mathrm{V}^0 B\big) \subset \mathrm{T}_{\mathrm{d}S(\xi)}\big(Z^{-1}(0)\big) = \ker Z'_{\mathrm{d}S(\xi)}.$$

Thus, if $(\mathrm{d}S)'_\xi X \in \mathrm{T}_{\mathrm{d}S(\xi)}(\mathrm{V}^0 B)$, then

$$\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0}\big\langle \mathrm{d}S \circ \gamma(t), Z \circ \gamma(t)\big\rangle = \frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} Z\big(\mathrm{d}S \circ \gamma(t)\big) = Z'_{\mathrm{d}S(\xi)} \circ (\mathrm{d}S)'_\xi(X) = 0.$$

On the other hand, using (12.4.22) we calculate

$$\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0}\big\langle \mathrm{d}S \circ \gamma(t), Z \circ \gamma(t)\big\rangle = -\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0}\big\langle \Phi \circ \gamma(t), \pi'_L \circ \Phi' \circ Z \circ \gamma(t)\big\rangle.$$

By our choice of γ, the expression on the right hand side can be interpreted as arising from the pairing $\mathrm{T}^*_\xi B \times \mathrm{T}_\xi B \to \mathbb{R}$. Hence, we may apply the product rule to obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}_{\restriction 0} \big\langle \mathrm{d}S \circ \gamma(t), Z \circ \gamma(t) \big\rangle = -\big\langle \Phi'_\xi X, \pi'_L \circ \Phi'\big(Z(\xi)\big)\big\rangle,$$

where $\Phi'_\xi X$ is viewed as an element of $\mathrm{T}^*_\xi L$ via the identification of $\mathrm{T}_{\Phi(\xi)}(\mathrm{T}^*_\xi L)$ with $\mathrm{T}^*_\xi L$. The second contribution from the product rule vanishes, because $\Phi(\xi)$ belongs to the zero section. Since by (12.4.24), the right hand side vanishes for all $Z(\xi) \in \mathrm{V}_\xi B$, (12.4.19) implies that $\Phi'_\xi X = 0$. Since Φ'_ξ is injective, we conclude that $X = 0$. This proves (12.4.23) and hence completes the proof of the theorem. □

More generally, consider a Lagrangian immersion $\iota : L \to \mathrm{T}^* Q$. Since every point in L possesses an open neighbourhood for which points 1 and 2 of Theorem 12.4.9 are fulfilled, this theorem implies

Corollary 12.4.10 *Locally, every Lagrangian immersion is generated by a Morse family.*

Exercises

12.4.1 In Example 8.5.8/2 we have shown that $(\mathrm{V}Q)^0$ is a coisotropic submanifold of $\mathrm{T}^* Q$. Use Lemma 12.4.1 to give an alternative proof of this fact.
Hint. Use Proposition 7.2.4 and the inequality (8.1.6).

12.4.2 Prove the statements of Remark 12.4.6.

12.4.3 Complete Example 12.4.7 by showing that for the class of functions under consideration, S is a Morse family.

12.5 Stable Equivalence

In this section we derive a partition of Morse families into classes which locally generate the same Lagrangian immersion. This will lead us to the notion of stable equivalence. The results below belong to Hörmander and Weinstein.

The local generation concept alluded to above is the following. Let $\iota : L \to \mathrm{T}^* Q$ be a Lagrangian immersion and let $\xi \in L$. We say that a Morse family (B, π, S) over Q generates (L, ι) at ξ if there exists an open subset U of B_S and an open neighbourhood V of ξ in L such that $\Lambda_S(U) = \iota(V)$. The class of Morse families generating (L, ι) at ξ will be denoted by $\mathscr{L}(L, \iota, \xi)$. According to Corollary 12.4.10, every Lagrangian immersion with target space $\mathrm{T}^* Q$ is generated at an arbitrary point by some Morse family over Q.

By the local nature of the generation concept under consideration, we can restrict our attention to the case where B is an open subset of $Q \times \mathbb{R}^r$ for some r, referred to as the fibre dimension of (B, π, S), and where π is given by the restriction of the projection to the factor Q. Points of Q and $\mathbb{R}^r$ will be denoted by x and $\mathbf{y}$, respectively, and the dimension of Q will be denoted by n. For the first and second derivatives of $S\colon B \to \mathbb{R}$ we introduce the simplified block matrix notation

$$S' = \big[S'_x, S'_y\big], \qquad S'' = \begin{bmatrix} S''_{xx} & S''_{xy} \\ S''_{yx} & S''_{yy} \end{bmatrix}.$$

For all $(x, \mathbf{y}) \in B_S$, we have $S'_y(x, \mathbf{y}) = 0$. For such points, the fibrewise bilinear form S''_{yy} on $(\mathrm{V}B)_{\restriction B_S}$ is referred to as the fibre Hessian[9] of S. For $(x, \mathbf{y}) \in B_S$,

$$\mathrm{T}_{(x,\mathbf{y})} B_S = \big\{(X, \mathbf{Y}) \in \mathrm{T}_x Q \oplus \mathbb{R}^r : S''_{yx}(x, \mathbf{y})X + S''_{yy}(x, \mathbf{y})\mathbf{Y} = 0\big\} \tag{12.5.1}$$

and for $(X, \mathbf{Y}) \in \mathrm{T}_{(x,\mathbf{y})} B_S$,

$$(\varLambda_S)'_{(x,\mathbf{y})}(X, \mathbf{Y}) = S''_{xx}(x, \mathbf{y})X + S''_{xy}(x, \mathbf{y})\mathbf{Y}. \tag{12.5.2}$$

Since $\varLambda_S$ covers π, the vector $S''_{xy}(x, \mathbf{y})\mathbf{Y}$ is tangent to the fibre $\mathrm{T}^*_x Q$ at ξ and thus belongs to the intersection $\mathrm{T}_\xi(\mathrm{T}^*_x Q) \cap \mathrm{T}_\xi L$. Finally, we introduce the notation

$$\varPi := \pi_Q \circ \iota : L \to Q. \tag{12.5.3}$$

Lemma 12.5.1 *Let $\iota : L \to \mathrm{T}^* Q$ be a Lagrangian immersion, let $\xi \in L$ and let (B, π, S) be a Morse family of fibre dimension r generating (L, ι) at ξ. For $(x, \mathbf{y}) \in B_S$ such that $\xi := \varLambda_S(x, \mathbf{y})$, we have*

$$r - \operatorname{rank} S''_{yy}(x, \mathbf{y}) = n - \operatorname{rank} \varPi'_\xi. \tag{12.5.4}$$

Proof Let $(X, \mathbf{Y}) \in \mathrm{T}_{(x,\mathbf{y})} B_S$ and $Z := (\varLambda_S)'_{(x,\mathbf{y})}(X, \mathbf{Y})$. Then, $\varPi'_\xi(Z) = X$. Since $r - \operatorname{rank} S''_{yy}(x, \mathbf{y}) = \dim\ker S''_{yy}(x, \mathbf{y})$ and $n - \operatorname{rank} \varPi'_\xi = \dim\ker \varPi'_\xi$, it is enough to compare the dimensions of the kernels. For that purpose, we set $X = 0$ in (12.5.1) and (12.5.2) and read off the system of equations

$$Z = S''_{xy}(x, \mathbf{y})\mathbf{Y}, \qquad S''_{yy}(x, \mathbf{y})\mathbf{Y} = 0.$$

If $\dim\ker S''_{yy}(x, \mathbf{y}) = 0$, then $\mathbf{Y} = 0$ and $Z = 0$ is the only solution and we obtain $\dim\ker \varPi'_\xi = 0$. If $\dim\ker S''_{yy}(x, \mathbf{y}) \neq 0$, since S is a Morse family, $S''_{xy}(x, \mathbf{y})$ has maximal rank and thus yields an isomorphism between $\ker S''_{yy}(x, \mathbf{y})$ and $\ker \varPi'_\xi$. □

Next, we use the Morse-Bott Lemma in the formulation of Corollary 8.9.13 to prove the Splitting Lemma for Morse families. We say that a mapping φ between open subsets of $Q \times \mathbb{R}^r$ preserves the fibres if $\mathrm{pr}_Q \circ \varphi(x, \mathbf{y}) = x$ for all points $(x, \mathbf{y})$ in the domain of φ.

Theorem 12.5.2 (Splitting Lemma) *Let (B, π, S) be a Morse family over Q of fibre dimension $k + l$ and let $(x_0, \mathbf{y}_0) \in B_S$. Write $\mathbf{y}_0 = (\hat{\mathbf{y}}_0, \tilde{\mathbf{y}}_0)$ with $\hat{\mathbf{y}}_0 \in \mathbb{R}^l$ and $\tilde{\mathbf{y}}_0 \in \mathbb{R}^k$.*

[9] For an intrinsic definition, see [36, §4.3].

If the fibre Hessian $S''_{yy}(x_0, \mathbf{y}_0)$ *has rank* k, *there exist open neighbourhoods* $\hat{B}$ *of* $(x_0, \hat{\mathbf{y}}_0)$ *in* $Q \times \mathbb{R}^l$ *and* V *of* $\tilde{\mathbf{y}}_0$ *in* $\mathbb{R}^k$, *a fibre-preserving diffeomorphism* φ *from* $\hat{B} \times V$ *into* B, *a smooth function* $\hat{S} : \hat{B} \to \mathbb{R}$ *and a non-degenerate quadratic form* Q *on* $\mathbb{R}^k$ *such that*

$$S \circ \varphi(x, \hat{\mathbf{y}}, \tilde{\mathbf{y}}) = \hat{S}(x, \hat{\mathbf{y}}) + \mathsf{Q}(\tilde{\mathbf{y}}). \tag{12.5.5}$$

$\hat{S}$ *is a Morse family along the submersion* $\hat{\pi} : \hat{B} \subset Q \times \mathbb{R}^l \to Q$ *whose fibre Hessian* $\hat{S}''_{\hat{y}\hat{y}}$ *vanishes at* $(x_0, \hat{\mathbf{y}}_0)$ *and which generates the same Lagrangian submanifold at* $\Lambda_S(x_0, \mathbf{y}_0)$ *as* S, *that is,* $\Lambda_{\hat{S}}(\hat{B}_{\hat{S}}) = \Lambda_S(B_S \cap (\hat{B} \times V))$.

Proof Let $(x_0, \mathbf{y}_0) \in B_S$. Up to a linear transformation which can be absorbed in φ, we may assume $\det S''_{\tilde{y}\tilde{y}}(x_0, \mathbf{y}_0) \neq 0$. Therefore, among the defining equations $S'_y = 0$ of B_S we can use the last k equations $S'_{\tilde{y}} = 0$ for locally determining

$$\tilde{\mathbf{y}} = \psi(x, \hat{\mathbf{y}})$$

in terms of a smooth mapping $\psi : \hat{B} \to \mathbb{R}^k$. Then, the subset $\tilde{B}_S \subset B$, defined by the equation $S'_{\tilde{y}} = 0$, is locally given by the graph

$$\big(x, \hat{\mathbf{y}}, \psi(x, \hat{\mathbf{y}})\big).$$

By the local diffeomorphism

$$\phi(x, \mathbf{y}) := \big(x, \hat{\mathbf{y}}, \tilde{\mathbf{y}} - \psi(x, \hat{\mathbf{y}})\big)$$

this graph is mapped to $\hat{B} \times \{0\}$. Let $i : \hat{B} \to \hat{B} \times \mathbb{R}^k$ be the corresponding inclusion mapping and $\mathrm{pr}_1 : \hat{B} \times \mathbb{R}^k \to \hat{B}$ the natural projection. Denote

$$\hat{S} := S \circ \phi^{-1} \circ i : \hat{B} \to \mathbb{R}$$

and define

$$\tilde{S} : \hat{B} \times \mathbb{R}^k \to \mathbb{R}, \qquad \tilde{S} := S \circ \phi^{-1} - \hat{S} \circ \mathrm{pr}_1 .$$

This is a Morse-Bott function with critical submanifold $\hat{B} \times \{0\}$. Now, the Morse-Bott Lemma 8.9.13 yields a diffeomorphism Φ and a non-degenerate quadratic form Q on $\mathbb{R}^k$ such that

$$\tilde{S} \circ \Phi(x, \hat{\mathbf{y}}, \tilde{\mathbf{y}}) = \mathsf{Q}(\tilde{\mathbf{y}}), \qquad \mathrm{pr}_1 \circ \Phi(x, \hat{\mathbf{y}}, \tilde{\mathbf{y}}) = (x, \hat{\mathbf{y}}).$$

Then, $\varphi := \phi^{-1} \circ \Phi$ yields the desired diffeomorphism. By construction, we have

$$\hat{S}(x, \hat{\mathbf{y}}) = S\big(x, \hat{\mathbf{y}}, \psi(x, \hat{\mathbf{y}})\big). \tag{12.5.6}$$

By assumption, $\dim\ker S''_{yy}(x_0, \mathbf{y}_0) = l$. Since S is a Morse family, $S''_{xy}(x_0, \mathbf{y}_0)$ has maximal rank. Hence, also $\hat{S}''_{x\hat{y}}(x_0, \hat{\mathbf{y}}_0)$ has maximal rank l, that is, $\hat{S}$ is a Morse family. To show that $\hat{S}$ generates the same Lagrangian submanifold at $(x_0, \mathbf{y}_0)$ as S, we can use the local description of Remark 12.4.6. This way, from (12.5.6) we read off that $\hat{B}_{\hat{S}} = B_S \cap (\hat{B} \times V)$. Then, equality of the images follows. □

Remark 12.5.3

1. Let $\iota : L \to \mathrm{T}^* Q$ be a Lagrangian immersion and let $\xi \in L$. Lemma 12.5.1 states that

$$r \geq n - \operatorname{rank} \Pi'_{\xi} = \dim \ker \Pi'_{\xi}, \tag{12.5.7}$$

that is, the right hand side yields a lower bound for the fibre dimension of a Morse family generating (L, ι) at ξ. Moreover, equality holds in (12.5.7) iff $S''_{yy}(x, \mathbf{y}) = 0$ for some $(x, \mathbf{y}) \in B_S$ such that $\Lambda_S(x, \mathbf{y}) = \xi$. A Morse family generating (L, ι) at ξ which fulfils this condition will be said to be reduced. In view of Corollary 12.4.10, Theorem 12.5.2 states that a reduced family always exists, that is, that the lower bound provided by (12.5.7) is sharp.

2. Let (B, π, S) be a reduced Morse family over Q of fibre dimension r generating (L, ι) at $\xi \in L$ and let $(x, \mathbf{y}) \in B_S$ such that $\Lambda_S(x, \mathbf{y}) = \xi$. Then, elements Z of $\iota'(\mathrm{T}_{\xi} L)$ are characterized by

$$Z = S''_{xx}(x, \mathbf{y}) X + S''_{xy}(x, \mathbf{y}) \mathbf{Y}, \qquad S''_{yx}(x, \mathbf{y}) X = 0, \tag{12.5.8}$$

where $(X, \mathbf{Y}) \in \mathrm{T}_{(x,\mathbf{y})} B = \mathrm{T}_x Q \oplus \mathbb{R}^r$. Since S is a Morse family, $S''_{xy}(x, \mathbf{y})$ must have maximal rank at $(x, \mathbf{y})$. Thus, by (12.5.8), we have the direct sum decomposition

$$\iota'(\mathrm{T}_{\xi} L) = \operatorname{im} S''_{xy}(x, \mathbf{y}) \oplus S''_{xx}\big(\ker S''_{yx}(x, \mathbf{y})\big), \tag{12.5.9}$$

where $\operatorname{im} S''_{xy}(x, \mathbf{y}) = \mathrm{T}_{\xi}(\mathrm{T}^*_x Q) \cap \iota'(\mathrm{T}_{\xi} L)$. Moreover, Eq. (12.5.2) implies that a tangent vector $(0, \mathbf{Y}) \in \mathrm{T}_{(x,\mathbf{y})} \pi^{-1}(x)$ is tangent to the fibre-critical submanifold B_S iff it is contained in the kernel of $S''_{yy}(x, \mathbf{y})$. Since $S''_{yy}(x, \mathbf{y}) = 0$, all vectors tangent to the fibres are contained in the kernel of this mapping and we obtain

$$\mathrm{T}_{(x,\mathbf{y})} \pi^{-1}(x) \subset \mathrm{T}_{(x,\mathbf{y})} B_S. \tag{12.5.10}$$

The Splitting Lemma can be interpreted as an operation on $\mathscr{L}(L, \iota, \xi)$, building from a given element (B, π, S) the reduced element $(\hat{B}, \hat{\pi}, \hat{S})$. The following operations produce Morse families belonging to $\mathscr{L}(L, \iota, \xi)$, too.

(a) Addition: choose $c \in \mathbb{R}$ and take $(B, \pi, S + c)$,
(b) Composition: choose a submersion $\tilde{\pi} : \tilde{B} \to Q$ and a fibre-preserving diffeomorphism $\varphi : \tilde{B} \to B$ and take $(\tilde{B}, \tilde{\pi}, S \circ \varphi)$,
(c) Suspension: choose a non-degenerate bilinear form Q on $\mathbb{R}^k$ and take $(\tilde{B}, \tilde{\pi}, \tilde{S})$ with $\tilde{B} = B \times \mathbb{R}^k$, $\tilde{\pi} = \pi \circ pr_B$ and $\tilde{S} = \mathrm{pr}^*_B S + \mathrm{pr}^*_{\mathbb{R}^k} \mathsf{Q}$.
(d) Restriction: choose an open subset $\tilde{B}$ of B containing a point of $\Lambda_S^{-1}(\iota(\xi))$ and take $(\tilde{B}, \pi_{\restriction \tilde{B}}, S_{\restriction \tilde{B}})$.

These operations generate an equivalence relation in $\mathscr{L}(L, \iota, \xi)$, called stable equivalence.[10] Theorem 12.5.2 implies

Corollary 12.5.4 *Every Morse family generating (L, ι) at ξ is stably equivalent to a reduced Morse family with that property.*

Now, we can state the main result of this section.

Theorem 12.5.5 *Any two Morse families generating (L, ι) at ξ are stably equivalent.*

Proof The proof is along the lines of the proof of Proposition 5.4 of [115] and Theorem 4.18 in [36]. By Corollary 12.5.4, it is enough to show that any two reduced Morse families in $\mathscr{L}(L, \iota, \xi)$ are stably equivalent. Since this is a local statement, we may assume $Q = \mathbb{R}^n$. Let S and $\tilde{S}$ be reduced Morse families defined on the open subsets B and $\tilde{B}$ of $\mathbb{R}^n \times \mathbb{R}^l$, respectively. By shrinking B and $\tilde{B}$ if necessary, we may assume that $\Lambda_S(B_S) = \Lambda_{\tilde{S}}(\tilde{B}_{\tilde{S}})$.

We will proceed in two steps. In the first step, we construct a fibre-preserving diffeomorphism $\psi : \tilde{B} \to B$ such that the functions $\tilde{S}$ and $S \circ \psi$ coincide on $\tilde{B}_{\tilde{S}}$. In the second step, we use the deformation method of Moser to construct a fibre-preserving diffeomorphism Φ_1 such that $S \circ \psi \circ \Phi_1 = \tilde{S}$ on the whole of $\tilde{B}_{\tilde{S}}$.

To carry out the first step, define mappings $\varphi\colon B \to \mathbb{R}^n \times \mathbb{R}^n$ and $\tilde{\varphi}\colon \tilde{B} \to \mathbb{R}^n \times \mathbb{R}^n$ by

$$\varphi(\mathbf{x}, \mathbf{y}) = \big(\mathbf{x}, S'_x(\mathbf{x}, \mathbf{y})\big), \qquad \tilde{\varphi}(\mathbf{x}, \mathbf{y}) = \big(\mathbf{x}, \tilde{S}'_x(\mathbf{x}, \mathbf{y})\big),$$

respectively. Since both λ_S and $\lambda_{\tilde{S}}$ cover the restriction of the projection to the factor Q, under the identification of T^*Q with $\mathbb{R}^n \times \mathbb{R}^n$, the restrictions $\varphi_{\restriction B_S}$ and $\tilde{\varphi}_{\restriction \tilde{B}_{\tilde{S}}}$ correspond to Λ_S and $\Lambda_{\tilde{S}}$, respectively. Let $(\mathbf{x}_0, \mathbf{y}_0) \in B_S$ and $(\mathbf{x}_0, \tilde{\mathbf{y}}_0) \in \tilde{B}_{\tilde{S}}$ be such that $\Lambda_S(\mathbf{x}_0, \mathbf{y}_0) = \Lambda_{\tilde{S}}(\mathbf{x}_0, \tilde{\mathbf{y}}_0) = \xi$. Since S is reduced, $S''_{xy}(\mathbf{x}_0, \mathbf{y}_0)$ has maximal rank l. Hence, its image

$$W := \operatorname{im} S''_{xy}(\mathbf{x}_0, \mathbf{y}_0)$$

is an l-dimensional subspace of $\mathrm{T}_\xi(\mathrm{T}^*_{\mathbf{x}_0} Q)$, the latter coinciding with $\mathbb{R}^n$ according to the above identification. Denoting the corresponding orthogonal projection by $p : \mathbb{R}^n \to W$, we define

$$F : B \times W \to W, \qquad F(\mathbf{x}, \mathbf{y}, \mathbf{w}) := p \circ S'_x(\mathbf{x}, \mathbf{y}) - \mathbf{w}.$$

Since W is the tangent space of the submanifold embedding[11] $\mathbf{y} \mapsto S'_x(\mathbf{x}_0, \mathbf{y})$ at $\mathbf{y} = \mathbf{y}_0$, it is orthogonal to $S'_x(\mathbf{x}_0, \mathbf{y}_0)$ and we have $p \circ S'_x(\mathbf{x}_0, \mathbf{y}_0) = 0$. Hence,

[10] We refer to Remark 12.6.16 for a comment on the notion of stability in this context.

[11] According to Remark 12.5.3/2, $W = \mathrm{T}_\xi(\mathrm{T}^*_x Q) \cap \iota'(\mathrm{T}_\xi L)$.

$F(\mathbf{x}_0, \mathbf{y}_0, 0) = 0$. Since, by construction, $F'_y(\mathbf{x}_0, \mathbf{y}_0, 0) = p \circ S''_{xy}(\mathbf{x}_0, \mathbf{y}_0)$ is bijective, the Implicit Mapping Theorem yields a smooth mapping g from some neighbourhood of $(\mathbf{x}_0, 0)$ in $\mathbb{R}^n \times W$ to $\mathbb{R}^l$ such that for all $\mathbf{y}$ in a neighbourhood of $\mathbf{y}_0$ in $\mathbb{R}^l$ one has $p \circ S'_x(\mathbf{x}, \mathbf{y}) = \mathbf{w}$ iff $\mathbf{y} = g(\mathbf{x}, \mathbf{w})$. After further shrinking B and $\tilde{B}$ if necessary, we can find an open neighbourhood U of both $\varphi(B)$ and $\tilde{\varphi}(\tilde{B})$ in $\mathbb{R}^n \times \mathbb{R}^n$ such that $\mathrm{pr}_1(U) = \mathrm{pr}_1(B) = \mathrm{pr}_1(\tilde{B})$ and such that the mapping

$$\rho : U \to B, \qquad \rho(\mathbf{x}, \mathbf{z}) := \big(\mathbf{x}, g\big(\mathbf{x}, p(\mathbf{z})\big)\big),$$

is defined. Then, $g(\mathbf{x}, p \circ S'_x(\mathbf{x}, \mathbf{y})) = \mathbf{y}$ for all $(\mathbf{x}, \mathbf{y}) \in B$ and hence $\rho \circ \varphi = \mathrm{id}_B$. Now, the desired fibre-preserving diffeomorphism is given by

$$\psi := \rho \circ \tilde{\varphi} : \tilde{B} \to B.$$

Indeed, due to $\tilde{\varphi}(\tilde{B}_{\tilde{S}}) = \varphi(B_S)$, it satisfies $\psi(\tilde{B}_{\tilde{S}}) = B_S$, and for $(\mathbf{x}, \mathbf{y}) \in \tilde{B}_{\tilde{S}}$ we have

$$\Lambda_S\big(\psi(\mathbf{x}, \mathbf{y})\big) = \varphi \circ \rho\big(\tilde{\varphi}(\mathbf{x}, \mathbf{y})\big) = \Lambda_{\tilde{S}}(\mathbf{x}, \mathbf{y}),$$

because points in $\varphi(B)$ are mapped identically under $\varphi \circ \rho$. Thus, $\Lambda_S \circ \psi_{\restriction \tilde{B}_{\tilde{S}}} = \Lambda_{\tilde{S}}$ and (12.4.10) yields

$$\mathrm{d}(\tilde{S}_{\restriction \tilde{B}_{\tilde{S}}}) = \Lambda^*_{\tilde{S}}(\theta_Q) = (\psi_{\restriction \tilde{B}_{\tilde{S}}})^* \circ \Lambda^*_S(\theta_Q) = \mathrm{d}(S_{\restriction B_S} \circ \psi_{\restriction \tilde{B}_{\tilde{S}}}),$$

so that on $\tilde{B}_{\tilde{S}}$, the functions $S \circ \psi$ and $\tilde{S}$ differ by a constant[12]. By absorbing this constant in S, we may assume that $S \circ \psi$ and $\tilde{S}$ coincide on $\tilde{B}_{\tilde{S}}$.

Now, we turn to the second step. We denote $S_0 = \tilde{S}$ and $S_1 = S \circ \psi$. Then, $B_{S_0} = B_{S_1} = \tilde{B}_{\tilde{S}}$ and $S_1 - S_0$ vanishes on $\tilde{B}_{\tilde{S}}$ up to second order. Thus, by a version of the Taylor Theorem, see Exercise 12.5.1, there exist smooth functions $h^{\alpha\beta}$ on $\tilde{B}$ such that

$$S_1 - S_0 = \sum_{\alpha,\beta} h^{\alpha\beta} \frac{\partial S_0}{\partial y^\alpha} \frac{\partial S_0}{\partial y^\beta}. \tag{12.5.11}$$

We put $S_t = S_0 + t(S_1 - S_0)$ and seek for a time-dependent vector field X_t whose flow Φ_t preserves the fibres, maps $\tilde{B}_{\tilde{S}}$ identically and fulfils $\Phi_t^* S_t = S_0$ for all $t \in [0, 1]$. The latter equality holds iff

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}(S_t \circ \Phi_t) = \Phi_t^* \frac{\mathrm{d}S_t}{\mathrm{d}t} + \Phi_t^*\big(X_t(S_t)\big) = \Phi_t^*\big(X_t(S_t) + S_1 - S_0\big),$$

that is, iff

$$X_t(S_t) + S_1 - S_0 = 0. \tag{12.5.12}$$

[12] We may of course assume B_S and hence $\tilde{B}_{\tilde{S}}$ to be connected.

Plugging in (12.5.11), together with the ansatz

$$X_t = \sum_{\alpha,\beta} f^{\alpha\beta} \frac{\partial S_0}{\partial y^\alpha} \frac{\partial}{\partial y^\beta},$$

we obtain

$$0 = \sum_{\alpha,\beta} h^{\alpha\beta} \frac{\partial S_0}{\partial y^\alpha} \frac{\partial S_0}{\partial y^\beta} + \sum_{\alpha,\beta} f^{\alpha\beta} \frac{\partial S_0}{\partial y^\alpha} \frac{\partial}{\partial y^\beta} \left(S_0 + t \sum_{\rho,\tau} h^{\rho\tau} \frac{\partial S_0}{\partial y^\rho} \frac{\partial S_0}{\partial y^\tau} \right).$$

This equation is fulfilled if

$$0 = H + F(1+G),$$

where $H \equiv h^{\alpha\beta}$ and $F \equiv f^{\alpha\beta}$, and where G denotes a matrix-valued function which vanishes on $\tilde{B}_{\tilde{S}}$ for all values of t. In a neighbourhood of $\tilde{B}_{\tilde{S}}$ we can solve this equation with respect to F and thus determine X_t and Φ_t. By construction, Φ_t preserves the fibres and fulfils $S_1 \circ \Phi_1 = S_0$. Moreover, since $H = 0$ on $\tilde{B}_{\tilde{S}}$, we have $F = 0$ and hence $X_t = 0$ there. Hence, Φ_t maps $\tilde{B}_{\tilde{S}}$ identically. □

Remark 12.5.6

1. A slightly more direct, but not shorter proof of Theorem 12.5.5 can be obtained by using the following local description of Morse families, see [78, 141] and Exercise 12.5.2. Let $L \subset \mathrm{T}^*Q$ be Lagrange. Then, for every $\xi_0 \in L$ there exists a neighbourhood $L_0 \subset L$, together with appropriately chosen local coordinates x^i, y^i on T^*Q, and a smooth function $\mathbf{y} \mapsto H(\mathbf{y})$ such that L_0 is generated by the Morse family
$$S(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y} - H(\mathbf{y}). \tag{12.5.13}$$
Then, L_0 consists of the pairs $(\mathbf{x} = H'(\mathbf{y}), \mathbf{y})$.
2. We reformulate the results of the above discussion in terms of Lagrangian submanifold germs. For that purpose, let us refer to a Morse family (B, π, S) over Q such that $\Lambda_S(B_S)$ contains some given point $\xi \in \mathrm{T}^*Q$ as a Morse family at ξ. The operations of addition, composition, suspension and restriction, discussed prior to Theorem 12.5.5, naturally apply to Morse families at ξ and thus the notion of stable equivalence naturally extends to these families.

   A germ of immersions to a manifold M at a point $m \in M$ is an equivalence class of immersions to M such that m belongs to their images with respect to the following equivalence relation: two immersions (P, φ) and $(\tilde{P}, \tilde{\varphi})$ of M are equivalent at m if there exist open subsets $U \subset P$ and $\tilde{U} \subset \tilde{P}$ such that $m \in \varphi(U) = \tilde{\varphi}(\tilde{U})$. Theorem 12.5.5 states that germs of Lagrangian immersions at $\xi \in \mathrm{T}^*Q$ bijectively correspond to stable equivalence classes of Morse families at ξ. In particular, Morse families at ξ are stably equivalent iff they generate the same germ of Lagrangian immersions at ξ, that is, iff there exist open subsets $U \subset B_S$ and $\tilde{U} \subset \tilde{B}_{\tilde{S}}$ such that $\xi \in \Lambda_S(U) = \Lambda_{\tilde{S}}(\tilde{U})$. Thus, the latter may be taken as a geometric definition of stable equivalence.

The concept of Morse family is of fundamental importance in geometric asymptotics, to be dealt with in Sects. 12.7 and 12.8. There is a variety of other physical applications, in particular in thermodynamics and in the statics of mechanical systems, see [38, 40, 41, 293, 294], as well as [152].

Exercises

12.5.1 Prove the following version of the Taylor Theorem. Let f be a smooth function on a manifold M and let $g : M \to \mathbb{R}^r$ be a smooth submersion. If f and $\mathrm{d}f$ vanish on $g^{-1}(0)$, there exist functions f^{ij} on a neighbourhood of $g^{-1}(0)$ in M such that

$$f = \sum_{i,j} f^{ij} g_i g_j .$$

12.5.2 Prove the statements of Remark 12.5.6/1.
Hint. Use the atlas on $\mathscr{L}(\mathbb{R}^n)$ constructed in Sect. 7.6.

12.6 Maslov Class and Caustics

Let Q be a manifold of dimension n and let $\iota : L \to \mathrm{T}^*Q$ be a Lagrangian immersion. In this section, we study the intersection properties of the tangent spaces of this immersion with the vertical distribution on T^*Q. In this context, we will find a topological invariant which plays an important role in geometric asymptotics.

Definition 12.6.1 (Caustic) The set of critical points of $\Pi := \pi_Q \circ \iota \colon L \to Q$ is called the singular subset of L and is denoted by $\Sigma(L)$. The set of critical values of Π is called the caustic of L and is denoted by $\Gamma(L)$. The points of $\Gamma(L)$ are called focal points.

Remark 12.6.2 Since L and Q have the same dimension, a point $\xi \in L$ is critical iff the tangent mapping Π'_ξ is not injective.

Let $\Sigma_k(L)$ denote the subset of L consisting of the points ξ where Π'_ξ has rank $n-k$. These subsets provide disjoint decompositions

$$L = \Sigma_0(L) \cup \Sigma(L), \qquad \Sigma(L) := \bigcup_{k=1}^{n} \Sigma_k(L). \tag{12.6.1}$$

First, we show that the structure of the Maslov cycle of a Lagrangian subspace, discussed in Sect. 7.6, generalizes to Lagrangian submanifolds in generic position, provided the induced bundle $\iota^*\mathrm{T}(\mathrm{T}^*Q)$ is trivial. This result belongs to Arnold [13]. For the case where this bundle is nontrivial, we refer to [8] and [56]. Let $\mathbb{R}^{2n}$ be endowed with the canonical symplectic structure and let $L_0 := \{0\} \times \mathbb{R}^n$. We denote $\hat{\mathscr{L}}_k(n) := \hat{\mathscr{L}}_k(L_0)$ and $\mathscr{L}(n) := \mathscr{L}(\mathbb{R}^{2n})$, cf. Sect. 7.6. Let

$$\chi : \iota^*\mathrm{T}\big(\mathrm{T}^*Q\big) \to L \times \mathbb{R}^{2n}$$

be a trivialization mapping the vertical distribution to L_0, that is,

$$\chi\big(\mathrm{T}_{\iota(\xi)}\big(\mathrm{T}^*_{\Pi(\xi)}Q\big)\big) = L_0 \tag{12.6.2}$$

for all $\xi \in L$. Via χ, the image of the tangent mapping

$$\iota'_\xi : \mathrm{T}_\xi L \to \mathrm{T}_{\iota(\xi)}\mathrm{T}^*Q$$

is identified with a Lagrangian subspace in $\mathbb{R}^{2n}$. This way, we obtain a mapping

$$F : L \to \mathscr{L}(n), \qquad F(\xi) := \chi\big(\iota'_\xi(\mathrm{T}_\xi L)\big). \tag{12.6.3}$$

We say that L is in generic position if the mapping F defined by (12.6.3) is transversal[13] to every submanifold $\hat{\mathscr{L}}_k(n)$. One can show that every Lagrangian submanifold can be brought to generic position by an arbitrarily small transformation.[14]

Proposition 12.6.3 *Let $\iota : L \to \mathrm{T}^*Q$ be a Lagrangian immersion in generic position such that the induced bundle $\iota^*\mathrm{T}(\mathrm{T}^*Q)$ is trivial. Then, $\Sigma_k(L)$ is either empty or an embedded submanifold of L of codimension $\frac{k(k+1)}{2}$. Moreover, $\Sigma_1(L)$ possesses a natural coorientation compatible with the natural coorientation of $\hat{\mathscr{L}}_1(n)$ under the tangent mapping of F.*

Proof Let $\xi \in \Sigma_k(L)$. Then, rank $\Pi'_\xi = n - k$ and hence $\dim\ker \Pi'_\xi = k$, that is,

$$\dim\big(\iota'_\xi(\mathrm{T}_\xi L) \cap \mathrm{T}_{\iota(\xi)}\big(\mathrm{T}^*_{\Pi(\xi)}Q\big)\big) = k. \tag{12.6.4}$$

By applying the diffeomorphism χ we obtain $\dim(F(\xi) \cap L_0) = k$, that is, $F(\xi) \in \hat{\mathscr{L}}_k(n)$. Thus, $F(\Sigma_k(L)) \subset \hat{\mathscr{L}}_k(n)$. Conversely, if $F(\xi) \in \hat{\mathscr{L}}_k(n)$, then (12.6.4) holds and hence $\xi \in \Sigma_k(L)$. Thus, $F(\Sigma_k(L)) = \hat{\mathscr{L}}_k(n) \cap F(L)$. It follows that

$$\Sigma_k(L) = F^{-1}\big(\hat{\mathscr{L}}_k(n)\big). \tag{12.6.5}$$

Now, Theorem 1.8.2 and Corollary 7.6.11 imply that $\Sigma_k(L)$ is an embedded submanifold of codimension $\frac{k(k+1)}{2}$. Since F'_ξ induces a bijection between the normal spaces $\mathrm{N}_\xi(\Sigma_1(L))$ and $\mathrm{N}_{F(\xi)}(\hat{\mathscr{L}}_1(n))$, the natural coorientation of $\hat{\mathscr{L}}_1(n)$ provided by Proposition 7.7.7 carries over to $\Sigma_1(L)$. □

Remark 12.6.4

1. Proposition 12.6.3 implies that $\Sigma(L)$ has a structure analogous to that of the Maslov cycle $\hat{\mathscr{L}}(L)$ in $\mathscr{L}(n)$, cf. Remark 7.6.12: it is a stratified subset of L, with the stratum $\Sigma_1(L)$ having codimension 1 in L and the other strata having codimension at least 3. This implies that $\Sigma_1(L)$ is open and dense in $\Sigma(L)$.

[13] This means, in particular, that $F(L)$ does not intersect $\hat{\mathscr{L}}_k(n)$ for $n < \frac{k(k+1)}{2}$.

[14] This is a consequence of the Sard Theorem 1.5.18, see Lemma 4.1.3 in [13].

2. Let us describe L explicitly in terms of local coordinates x^i, p_i on T^*Q induced from a local chart on Q. Let $\dot{x}^i$ and $\dot{p}_i$ denote the corresponding fibre coordinates on $\mathrm{T}(\mathrm{T}^*Q)$, let $U \subset L$ be an open subset such that $\iota_{\restriction U}$ is an embedding and $\iota(U)$ is contained in the domain of the coordinates x^i and p_i. Since the local trivialization of $\mathrm{T}(\mathrm{T}^*Q)$ induced by the bundle coordinates x^i, p_i, $\dot{x}^i$ and $\dot{p}_i$ fulfils (12.6.2), it can be used to construct the mapping $F : U \to \mathscr{L}(n)$. By Proposition 7.6.8, for every $\xi \in U$ there exists a subset $K \subset \{1, \dots, n\}$ such that $F(\xi)$ belongs to the domain of the local chart φ_K on $\mathscr{L}(n)$ defined by (7.6.12) and (7.6.16). This local chart assigns to $F(\xi)$ an n-dimensional symmetric matrix, which we denote by $A^K_{ij}(\xi)$. According to (7.6.18) and (7.6.19), in terms of the fibre coordinates $\dot{x}^i$ and $\dot{p}_i$, the Lagrangian subspace $\iota' \mathrm{T}_\xi L$ of $\mathrm{T}_\xi(\mathrm{T}^*Q)$ is given by the n equations

$$\dot{x}^l = \sum_{j \notin K} A^K_{lj}(\xi)\dot{x}^j - \sum_{m \in K} A^K_{lm}(\xi)\dot{p}_m, \quad l \in K, \tag{12.6.6}$$

$$\dot{p}_i = \sum_{j \notin K} A^K_{ij}(\xi)\dot{x}^j - \sum_{m \in K} A^K_{im}(\xi)\dot{p}_m, \quad i \notin K. \tag{12.6.7}$$

Hence, it can be parameterized by the fibre coordinates corresponding to the coordinate functions x^i with $i \notin K$ and p_l with $l \in K$. Therefore, these functions provide a chart κ_K on L in a neighbourhood of ξ. According to (12.6.6) and (12.6.7), the matrix entries $A^K_{ij}(\xi)$ can be expressed in terms of partial derivatives of the local representatives of the remaining coordinate functions x^k in this chart:

$$\begin{aligned} A^K_{lm}(\xi) &= -\frac{\partial x^l}{\partial p_m}\big(\kappa_K(\xi)\big), & A^K_{li}(\xi) &= \frac{\partial x^l}{\partial x^i}\big(\kappa_K(\xi)\big), \\ A^K_{il}(\xi) &= -\frac{\partial p_i}{\partial p_l}\big(\kappa_K(\xi)\big), & A^K_{ij}(\xi) &= \frac{\partial p_i}{\partial x^j}\big(\kappa_K(\xi)\big), \end{aligned} \tag{12.6.8}$$

where $l, m \in K$ and $i, j \notin K$.

3. Consider the singular subset $\Sigma(L)$. For given $K \subset \{1, \dots, n\}$ consisting of k elements and given $\xi \in L$ such that $F(\xi)$ belongs to the domain of the chart φ_K, Proposition 7.6.10 yields that $\xi \in \Sigma_k(L)$ iff $A^K_{lm}(\xi) = 0$ for all $l, m \in K$. In particular, for $K = \{1\}$, an open subset of $\Sigma_1(L)$ is mapped under $\xi \mapsto A^K_{lm}(\xi)$ to the subspace of the vector space of n-dimensional symmetric matrices A given by $A_{11} = 0$. Hence, for all ξ in this subset we have $\frac{\partial x^1}{\partial p_1}(\kappa_K(\xi)) = 0$. According to Remark 7.7.8, in the corresponding chart φ_K on $\mathscr{L}(n)$, the coorientation of $\hat{\mathscr{L}}_1(n)$ points from the side where $A_{11} > 0$ to the side where $A_{11} < 0$. Thus, the first relation in (12.6.8) implies that the induced coorientation of $\Sigma_1(L)$ points from the side where $\frac{\partial x^1}{\partial p_1}(p_1, x^2, \dots, x^n) < 0$ to the side where $\frac{\partial x^1}{\partial p_1}(p_1, x^2, \dots, x^n) > 0$.

Next, we carry over the Maslov index for closed curves in $\mathscr{L}(n)$ and the intersection index for curves in $\mathscr{L}(n)$ with the Maslov cycle $\hat{\mathscr{L}}(n)$ to curves in L.

First, consider the Maslov index μ for closed curves in $\mathscr{L}(n)$, cf. Definition 7.7.1. Using the mapping F, we can define the Maslov index for closed curves γ in L by

$$\mu_L(\gamma) := \mu(F \circ \gamma).$$

Like μ, the Maslov index μ_L defines a homomorphism from $\pi_1(L)$ to $\mathbb{Z}$, denoted by the same symbol. On the other hand, via F, the universal Maslov class μ on $\mathscr{L}(n)$ defined by (7.7.4) induces a 1-form

$$\mu_L := F^* \mu$$

on L, called the Maslov class of (L, ι). By Proposition 7.7.4, for a closed curve γ in L, we have

$$\int_\gamma \mu_L = \int_\gamma F^* \mu = \int_{F \circ \gamma} \mu = \mu(F \circ \gamma) = \mu_L(\gamma). \tag{12.6.9}$$

As explained in Remark 7.7.6/1, integration of 1-forms over closed curves defines an isomorphism from the de-Rham cohomology group $H^1(L)$ to the group $\mathrm{Hom}(\pi_1(L), \mathbb{R})$ and (12.6.9) implies that this isomorphism assigns to the Maslov class μ_L the homomorphism defined by the Maslov index μ_L.

Second, consider the intersection index of curves in $\mathscr{L}(n)$ with respect to the Maslov cycle $\hat{\mathscr{L}}(n)$, cf. Definition 7.7.10. We carry over the terminology of crossings introduced in Sect. 7.7 to the present situation: a real number t such that $\gamma(t) \in \Sigma(L)$ is called a crossing of γ with $\Sigma(L)$. A crossing t is said to be simple if $\gamma(t) \in \Sigma_1(L)$. A simple crossing is said to be transversal if $\dot{\gamma}(t) \notin \mathrm{T}_{\gamma(t)} \Sigma_1(L)$. Depending on whether $\dot{\gamma}(t)$ is positively or negatively oriented with respect to the natural coorientation of $\Sigma_1(L)$ provided by Proposition 12.6.3, a simple transversal crossing is said to be positive or negative. Proposition 12.6.3, the transversality of F and $\hat{\mathscr{L}}(n)$, and the definition of the coorientation on $\Sigma_1(L)$ imply the following.

(a) Crossings of γ with $\Sigma(L)$ are crossings of $F \circ \gamma$ with $\hat{\mathscr{L}}(n)$ and vice versa.
(b) A crossing of γ with $\Sigma(L)$ is, respectively, simple, transversal, positive or negative iff it is so as a crossing of $F \circ \gamma$ with $\hat{\mathscr{L}}(n)$.

Since the complement of $\Sigma_0(L) \cup \Sigma_1(L)$ in L is the closure of the embedded submanifold $\Sigma_2(L)$ which by Proposition 12.6.3 has codimension 3, point 1 of Proposition 7.7.9 carries over to the present situation, yielding that

(c) every curve in L with end points in $\Sigma_0(L)$ is homotopic with fixed end points to a curve which has only simple transversal crossings with $\Sigma(L)$.

Since a homotopy with fixed end points of γ in L induces a homotopy with fixed end points of $F \circ \gamma$ in $\mathscr{L}(n)$, point 2 of Proposition 7.7.9 implies that

(d) if two curves in L which have the same end points in $\Sigma_0(L)$ and which have only simple and transversal crossings with $\Sigma(L)$ are homotopic with fixed end points, their differences between the numbers of positive and negative crossings coincide.

Points (c) and (d) allow for

Definition 12.6.5 (Maslov intersection index for Lagrangian immersions) The Maslov intersection index of a curve γ in L with end points in $\Sigma_0(L)$ is defined by

$$\mathrm{ind}_L(\gamma) := \nu_+ - \nu_-, \tag{12.6.10}$$

where ν_+ is the number of positive crossings and ν_- is the number of negative crossings with $\Sigma(L)$ of a curve which is homotopic with fixed end points to γ and whose crossings are all simple and transversal.

By point (d), the intersection index so defined is invariant under homotopies with fixed end points. It is obviously additive with respect to the composition of curves. By (a) and (b), we have

$$\mathrm{ind}_L(\gamma) = \mathrm{Ind}_{L_0}(F \circ \gamma) \tag{12.6.11}$$

for all curves γ with end points in $\Sigma_0(L)$.

Proposition 12.6.6 *Let $L \subset \mathrm{T}^*Q$ be a Lagrangian submanifold and let $\gamma : [0, 1] \to L$ be a closed curve with $\gamma(0) = \gamma(1) \in \Sigma_0(L)$. Then,*

$$\mu_L(\gamma) = \mathrm{ind}_L(\gamma). \tag{12.6.12}$$

This shows, in particular, that the definition of μ_L does not depend on the choice of the trivialization of $\iota^*\mathrm{T}(\mathrm{T}^*Q)$ in the construction of the mapping F.

Proof Under the assumption that $\gamma(0) = \gamma(1) \in \Sigma_0(L)$, the Lagrangian subspace L_0 of $\mathbb{R}^{2n}$ is transversal to $F \circ \gamma(0)$ and $F \circ \gamma(1)$, so that we can apply Theorem 7.7.11. In view of (12.6.11), this yields

$$\mu_L(\gamma) = \mu(F \circ \gamma) = \mathrm{Ind}_{L_0}(F \circ \gamma) = \mathrm{ind}_L(\gamma). \qquad \square$$

Remark 12.6.7 The intersection index $\mathrm{ind}_L(\gamma)$ can be expressed in terms of the Kashiwara index by a formula analogous to (7.8.9): for $\xi \in L$ we define

$$L_{1\xi} := \mathrm{T}_\xi\big(\mathrm{T}^*_{\Pi(\xi)}Q\big), \qquad L_{2\xi} := \iota'\mathrm{T}_\xi L.$$

We choose a sufficiently fine covering $\{U_j\}$ of L such that over each U_j there exists a smooth choice of an auxiliary subspace $L^j_{3\xi}$ transversal to both $L_{1\xi}$ and $L_{2\xi}$. Let

$\gamma : [a, b] \to L$ be a curve in L and choose $a = t_0 < t_1 < \cdots < t_k = b$ such that $\gamma([t_{j-1}, t_j]) \subset U_j$ for all j. Then, Proposition 7.8.8 implies

$$\operatorname{ind}_L(\gamma) \\ = \frac{1}{2}\sum_{j=1}^{k}\Big(s\big(L_{1\xi}, L_{2\xi}, L_{3\xi}^{j}\big)_{\restriction \xi=\gamma(t_{j-1})} - s\big(L_{1\xi}, L_{2\xi}, L_{3\xi}^{j}\big)_{\restriction \xi=\gamma(t_j)}\Big), \quad (12.6.13)$$

with the Kashiwara index s taken in the symplectic vector space $\mathrm{T}_\xi(\mathrm{T}^*Q)$.

Example 12.6.8

1. Consider $\mathbb{R}^2$, endowed with the canonical symplectic structure and with canonical coordinates q and p. Let L be an embedded submanifold diffeomorphic to S^1. For dimensional reasons, L is Lagrange. Let us calculate the Maslov index of L using (12.6.13). For that purpose, it suffices to consider a closed curve $t \mapsto \gamma(t)$, which runs through L exactly once in the direction of a chosen orientation. For simplicity, assume that L coincides with the circle defined by $q^2 + p^2 = 1$ and that it is oriented clockwise. The singular subset $\Sigma(L)$ consists of the points $\eta_+ = (1, 0)$ and $\eta_- = (-1, 0)$. We choose a covering by connected open subsets $U_1, \ldots, U_4$ such that $\eta_+ \in U_2$ and $\eta_- \in U_4$ and a compatible partition of γ. If we use the clockwise orientation of $\mathbb{R}^2$, the four terms in (12.6.13) are 0 for U_1 and U_3 and $+1$ for U_2 and U_4, respectively. Thus, we obtain the Maslov index $+2$. This can also be understood in the following way. While running through the singularities, the relative position of L_3 to the pair (L_1, L_2) changes. Above η_+ the subspace L_3 lies between L_1 and L_2, whereas beneath η_+ it lies outside of L_1 and L_2 in the sense of the chosen orientation.
2. In a similar way, one can discuss the Lagrangian immersion $\iota : L \to \mathbb{R}^2$ given by $L = \mathrm{S}^1$, realized as the unit circle in $\mathbb{C}$, and

$$\iota : L \to \mathbb{R}^2, \qquad \iota\big(\mathrm{e}^{\mathrm{i}\phi}\big) = \big(\cos(\phi), \sin(2\phi)\big).$$

   This is a figure eight immersion in horizontal position with respect to the canonical projection. It has the two self-intersection points $\pm\mathrm{i}$. The singular subset $\Sigma(L)$ consists of the two points ± 1. The reader can easily convince himself that for a curve γ which runs through L exactly once, Formula (12.6.13) yields $\mu_L(\gamma) = 0$ (Exercise 12.6.1).

Now, let us analyze the Maslov intersection index using the concept of Morse families. For that purpose, let $U \subset L$ be open and let (B, π, S) be a Morse family generating the Lagrangian immersion $\iota_{\restriction U} : U \to \mathrm{T}^*Q$. As in Sect. 12.5, we assume that B is an open subset of $Q \times \mathbb{R}^r$ and we write $(x, \mathbf{y})$ for its elements. We will also use the simplified notation for the second derivatives introduced there. By Lemma 12.5.1, for $(x, \mathbf{y}) \in B_S$ and $\xi \in U$ such that $\iota(\xi) = (x, \mathbf{y})$, we have

$$\dim\ker S''_{yy}(x, \mathbf{y}) = \dim\ker \Pi'_\xi .$$

Thus, Remark 12.6.2 implies that $\iota(\Sigma(U))$ coincides with the image under Λ_S of the set of solutions of the system of equations

$$S'_y = 0, \qquad \det\big(S''_{yy}\big) = 0. \tag{12.6.14}$$

Hence, the caustic $\Gamma(U) = \Pi(\Sigma(U))$ is given by

$$\Gamma(U) = \big\{x \in Q : \det\big(S''_{yy}(x, \mathbf{y})\big) = 0 \text{ and } S'_y(x, \mathbf{y}) = 0 \text{ for some } \mathbf{y} \in \mathbb{R}^r\big\}. \tag{12.6.15}$$

On the other hand, outside $\Sigma(L)$ we have

$$\det\big(S''_{yy}\big) \neq 0,$$

so that for points which are not critical, the first of the equations in (12.6.14) can be solved for the variables y^α. Thus, if U does not intersect $\Sigma(L)$, S can be reduced to a single generating function on Q.

Example 12.6.9 For the Morse family of Example 12.4.7, which generates the Lagrangian immersion of Example 12.6.8/2, the criterion (12.6.14) yields

$$\frac{\partial S}{\partial y}(x, y) = x^2 + y^2 - 1 = 0, \qquad \frac{\partial^2 S}{\partial y^2}(x, y) = 2y = 0.$$

Hence, the critical points are $(x, y) = (\pm 1, 0)$ and the focal points are $x = \pm 1$. This is consistent with what we have found in Example 12.6.8/2.

Next, we derive a formula for the intersection index in terms of generating Morse families. Recall that the index of a quadratic form Q on a vector space V is defined to be the number of negative eigenvalues, that is, the dimension of a maximal subspace of V on which H is negative definite. One has

$$\operatorname{index}(\mathsf{Q}) = \frac{1}{2}\big(\operatorname{rank}(\mathsf{Q}) - \operatorname{sign}(\mathsf{Q})\big), \tag{12.6.16}$$

where $\operatorname{sign}(\mathsf{Q})$ denotes the signature, that is, the number of positive eigenvalues minus the number of negative eigenvalues, counted with multiplicities.

Lemma 12.6.10 *Let $\iota : L \to \mathrm{T}^*Q$ be a Lagrangian immersion. Let there be given*

1. *an open covering $\{U_i\}$ of L such that all the intersections $U_i \cap U_j$ are connected,*
2. *Morse families (B_i, π_i, S_i), with B_i being open subsets of $Q \times \mathbb{R}^{r_i}$, such that the immersions $\iota_i : U_i \to \mathrm{T}^*Q$ and $\Lambda_{S_i} : B_{S_i} \to \mathrm{T}^*Q$ are equivalent.*

Then, the mappings[15]

$$c_{ij} : U_i \cap U_j \to \mathbb{Z}, \qquad c_{ij} := \operatorname{index}\big(\big(S''_i\big)_{y_i y_i}\big) - \operatorname{index}\big(\big(S''_j\big)_{y_j y_j}\big) \tag{12.6.17}$$

are constant for all i, j.

[15] By an abuse of notation, via the diffeomorphism $B_{S_i} \to U_i$ induced by Λ_{S_i}, $\operatorname{index}((S_i)''_{y_i y_i})$ is viewed as a function on U_i.

Proof Let (i, j) be a pair of indices such that $U_i \cap U_j$ is nonempty and let $\xi_0 \in U_i \cap U_j$. Denote $b_{0,l} = \Lambda_{S_l}^{-1}(\xi_0)$, $l = i, j$. Since the Morse families (B_i, π_i, S_i) and (B_j, π_j, S_j) generate $\iota : L \to \mathrm{T}^*Q$ at ξ, for $l = i, j$, a restriction of (B_l, π_l, S_l) to some open neighbourhood of $b_{0,l}$ in B_l arises from some reduced Morse family $(\hat{B}, \hat{\pi}, \hat{S})$ generating (L, ι) at ξ by the following operations, applied in the order they are listed and with additions and restrictions omitted: a composition with a fibre-preserving diffeomorphism ψ_l, a suspension with a non-degenerate quadratic form Q_l on $\mathbb{R}^{k_l}$ and a further composition with a fibre-preserving diffeomorphism φ_l. Thus, in a neighbourhood of $b_{0,l}$,

$$S_l = (\hat{S} \circ \psi_l + \mathsf{Q}_l) \circ \varphi_l, \tag{12.6.18}$$

where we have omitted the natural projections occurring in the suspension. To analyze how the index of the second derivative with respect to the fibre coordinates behaves under the above operations, let (B, π, S) be a Morse family over Q with $B \subset Q \times \mathbb{R}^r$ open. For a non-degenerate quadratic form Q on $\mathbb{R}^k$ we have

$$\operatorname{index}\big((S + \mathsf{Q})''_{yy}\big) = \operatorname{index}\big(S''_{yy}\big) + \operatorname{index}(\mathsf{Q}). \tag{12.6.19}$$

For an open subset $\tilde{B} \subset Q \times \mathbb{R}^r$ and a fibre-preserving diffeomorphism $\varphi : \tilde{B} \to B$ we calculate

$$(S \circ \varphi)''_{yy}(\tilde{b}) = \big(\varphi'_y(\tilde{b})\big)^{\mathrm{T}}\big(S''_{yy}\big(\varphi(\tilde{b})\big)\big)\varphi'_y(\tilde{b}) + \big(S'_y\big(\varphi(\tilde{b})\big)\big)\varphi''_{yy}(\tilde{b}).$$

Since $\varphi(\tilde{B}_{S\circ\varphi}) = B_S$ and since $S'_y = 0$ on B_S, and since a similarity transformation with a non-singular matrix does not change the index of a quadratic form, this implies

$$\operatorname{index}\big((S \circ \varphi)''_{yy}\big) = \operatorname{index}\big(S''_{yy} \circ \varphi\big). \tag{12.6.20}$$

Using (12.6.19) and (12.6.20), from (12.6.18) we obtain

$$\operatorname{index}\big((S_l)''_{y_l y_l}(b_l)\big) = \operatorname{index}\big(\hat{S}''_{\hat{y}\hat{y}}\big(\psi_l \circ \varphi_l(b_l)\big)\big) + \operatorname{index}(\mathsf{Q}_l) \tag{12.6.21}$$

for all b_l in a neighbourhood of $b_{0,l}$, where $l = i, j$. Since $\Lambda_{\hat{S}}(\psi_l \circ \varphi_l(b_l)) = \Lambda_{S_l}(b_l)$, if $\Lambda_{S_i}(b_i) = \Lambda_{S_j}(b_j)$, then $\psi_i \circ \varphi_i(b_i) = \psi_j \circ \varphi_j(b_j)$. Hence, (12.6.21) implies

$$c_{ij}(\xi) = \operatorname{index}(\mathsf{Q}_i) - \operatorname{index}(\mathsf{Q}_j)$$

for all ξ in some neighbourhood of ξ_0. This shows that c_{ij} is constant in a neighbourhood of every point of $U_i \cap U_j$. Since the latter is connected, this implies that c_{ij} is constant. □

Proposition 12.6.11 *Let $\iota : L \to \mathrm{T}^*Q$ be a Lagrangian immersion. Let $\gamma : [0, 1] \to L$ be a curve with end points in $\Sigma_0(L)$ which has only simple and transversal crossings with $\Sigma(L)$. Choose numbers $0 = t_0 < t_1 < \cdots < t_k = 1$ such that there exist*

1. *open subsets* $U_1, \dots, U_k$ *such that* $\gamma([t_{i-1}, t_i]) \subset U_i$ *and* $U_i \cap U_j$ *are connected,*
2. *Morse families* (B_i, π_i, S_i), *with* B_i *being open subsets of* $Q \times \mathbb{R}^{r_i}$, *such that the immersions* $\iota_i : U_i \to \mathrm{T}^* Q$ *and* $\Lambda_{S_i} : B_{S_i} \to \mathrm{T}^* Q$ *are equivalent.*

Then, the intersection index of γ *is given by*

$$\operatorname{ind}_L(\gamma) = \sum_{i=1}^{k} \Big\{ \operatorname{index}\big((S_i)''_{y_i y_i}\big(\gamma(t_i)\big)\big) - \operatorname{index}\big((S_i)''_{y_i y_i}\big(\gamma(t_{i-1})\big)\big) \Big\}. \tag{12.6.22}$$

Proof Using Lemma 12.6.10, for the right hand side of (12.6.22) we find

$$\begin{aligned} &\sum_{i=1}^{k} \Big\{ \operatorname{index}\big((S_i)''_{y_i y_i}\big(\gamma(t_i)\big)\big) - \operatorname{index}\big((S_i)''_{y_i y_i}\big(\gamma(t_{i-1})\big)\big) \Big\} \\ &\quad = \operatorname{index}\big((S_k)''_{y_k y_k}\big(\gamma(1)\big)\big) + \sum_{i=1}^{k-1} c_{i\,i+1}\big(\gamma(t_i)\big) - \operatorname{index}\big((S_0)''_{y_0 y_0}\big(\gamma(0)\big)\big). \end{aligned}$$

Since the c_{ij} are constant on $U_i \cap U_j$, the sum on the right hand side does not depend on the choice of the numbers t_i. In particular, we may choose them in such a way that $\gamma(t_i) \in \Sigma_0(L)$ for all i. If the line segment $\gamma([t_{i-1}, t_i])$ does not intersect $\Sigma_1(L)$, Lemma 12.5.1 implies that the rank of $(S_i)''_{y_i y_i}$ is constant on $\gamma([t_{i-1}, t_i])$. Then, also the signature is constant. By (12.6.16), then

$$\operatorname{index}\big((S_i)''_{y_i y_i}\big(\gamma(t_i)\big)\big) = \operatorname{index}\big((S_i)''_{y_i y_i}\big(\gamma(t_{i-1})\big)\big),$$

so that this line segment does not contribute to the sum on the right hand side of (12.6.22). For the remaining line segments we may assume that each of them contains exactly one crossing. By the Splitting Lemma 12.5.2, we may also assume that the corresponding Morse families (B_i, π_i, S_i) are reduced, which means that they have fibre dimension 1, because the crossings are simple. Then, Lemma 12.5.1 yields that $(S_i)''_{y_i y_i}(\gamma(t)) = 0$ at the crossing and $(S_i)''_{y_i y_i}(\gamma(t)) \neq 0$ outside. Therefore, $(S_i)''_{y_i y_i}(\gamma(t))$ has rank 1 at t_i and t_{i+1}, whereas the signature changes by ± 2 at the crossing. By (12.6.16), the index then changes by ± 1 there. As a consequence, the right hand side of (12.6.22) counts the crossings of γ with $\Sigma_1(L)$, weighted by $+1$ in case $(S_i)''_{y_i y_i}(\gamma(t))$ changes its sign from $+$ to $-$ and weighted by -1 otherwise. It remains to show that this weighting is consistent with the counting of the crossings in the Maslov intersection index (12.6.10), that is, that $(S_i)''_{y_i y_i}(\gamma(t))$ changes its sign from $+$ to $-$ iff γ crosses $\Sigma_1(L)$ in the direction of the coorientation of $\Sigma_1(L)$. Since the argument is independent of the line segment, we may omit the index i. Let t_c be the crossing under consideration and denote $\xi_c = \gamma(t_c)$. For the first part of the argument it is helpful to distinguish between points in $U \subset L$ and points in B_S. Therefore, let $(x_c, y_c) \in B_S$ be such that $\iota(\xi_c) = \Lambda_S(x_c, y_c)$. Since $S''_{yy}(x_c, y_c) = 0$, the bilinear form $S''_{xy}(x_c, y_c)$ must have maximal rank, that is, rank 1. Thus, we can find coordinates x^i on a neighbourhood of $\Pi(\xi) = x_c$ in Q such that

$\frac{\partial^2 S}{\partial x^1 \partial y} \neq 0$ in a neighbourhood of (x_c, y_c) in B. By replacing x^1 by $-x^1$ if necessary, we may assume that

$$\frac{\partial^2 S}{\partial x^1 \partial y} > 0. \tag{12.6.23}$$

Then, the Implicit Function Theorem yields a function $x^1 = x^1(y, x^2, \dots, x^n)$ fulfilling

$$\frac{\partial S}{\partial y}\big(x^1\big(y, x^2, \dots, x^n\big), x^2, \dots, x^n, y\big) = 0.$$

Hence, $(x^1(y, x^2, \dots, x^n), x^2, \dots, x^n, y) \in B_S$, so that y and $x^2, \dots, x^n$ provide coordinates on B_S in a neighbourhood of (x_c, y_c). Since $y \mapsto (x^1(y, x^2, \dots, x^n), x^2, \dots, x^n, y)$ is a curve in B_S, Eq. (12.4.15) implies

$$\frac{\partial^2 S}{\partial y \partial x^1}\frac{\partial x^1}{\partial y} + \frac{\partial^2 S}{\partial y^2} = 0. \tag{12.6.24}$$

Hence, along the line segment, we have $\frac{\partial x^1}{\partial y} = 0$ at the crossing and

$$\operatorname{sign}\left(\frac{\partial x^1}{\partial y}\right) = -\operatorname{sign}\left(\frac{\partial^2 S}{\partial y^2}\right) \tag{12.6.25}$$

outside. Now, consider L. By (12.4.12), in the coordinates x^i and p_i on T^*Q, for points in a neighbourhood of ξ_c in L we have $p_1 = \frac{\partial S}{\partial x^1}$. Using that $y, x^2, \dots, x^n$ provide coordinates on B_S, we find

$$\frac{\partial p_1}{\partial y} = \frac{\partial^2 S}{\partial (x^1)^2}\frac{\partial x^1}{\partial y} + \frac{\partial^2 S}{\partial x^1 \partial y}.$$

Since at ξ_c we have $\frac{\partial x^1}{\partial y} = 0$, the inequality (12.6.23) implies $\frac{\partial p_1}{\partial y} > 0$ in some neighbourhood of ξ_c in L, so that we may take $p_1, x^2, \dots, x^n$ as coordinates on L there. Then,

$$\frac{\partial x^1}{\partial y} = \frac{\partial x^1}{\partial p_1}\frac{\partial p_1}{\partial y}.$$

From this and from (12.6.25) we read off that, along the line segment, we have $\frac{\partial x^1}{\partial p_1} = 0$ at the crossing and

$$\operatorname{sign}\left(\frac{\partial x^1}{\partial p_1}\right) = \operatorname{sign}\left(\frac{\partial x^1}{\partial y}\right) = -\operatorname{sign}\left(\frac{\partial^2 S}{\partial y^2}\right)$$

outside. According to Remark 12.6.4/3, this means that if the line segment crosses $\Sigma_1(L)$ in the direction of the coorientation, that is, if $\frac{\partial x^1}{\partial p_1}$ changes its sign from $-$

to $+$, then $\frac{\partial^2 S}{\partial y^2}$ changes its sign from $+$ to $-$ and, hence, its index changes by $+1$. This completes the proof. □

Lemma 12.6.10 and Proposition 12.6.11 imply

Corollary 12.6.12 *Let $\iota : L \to \mathrm{T}^*Q$ be a Lagrangian immersion and let γ be a curve in L with end points in $\Sigma_0(L)$. Under the assumptions of Proposition 12.6.11, we have*

$$\operatorname{ind}_L(\gamma) = \sum_{i=1}^{k-1} c_{i\,i+1} + \operatorname{index}\big((S_k)''_{y_k y_k}\big(\gamma(1)\big)\big) - \operatorname{index}\big((S_0)''_{y_0 y_0}\big(\gamma(0)\big)\big).$$

If γ is closed, then

$$\mu_L(\gamma) = \sum_{i=1}^{k-1} c_{i\,i+1}.$$

In particular, a Lagrangian immersion generated by a single Morse family has trivial Maslov class, that is, in this case the Maslov index of any closed curve vanishes.

Example 12.6.13

1. Let L be the circle $q^2 + p^2 = 1$ in $\mathbb{R}^2$, cf. Example 12.6.8/1. Choosing p as a coordinate on L in the vicinity of each of the critical points $\eta_\pm = (\pm 1, 0)$, we find $x(p) = \pm\sqrt{1-p^2}$ and hence

$$\frac{\partial x}{\partial p} = \mp \frac{2p}{\sqrt{1-p^2}}.$$

   Thus, at η_+, the coorientation points from the upper half-plane to the lower half-plane, whereas at η_-, it points in the converse direction. Consequently, a curve running once clockwise around L has Maslov index $+2$. This is consistent with what we have found in the above-cited example. Note that Corollary 12.6.12 tells us that the circle cannot be generated by a single Morse family. We encourage the reader to construct a set of generating Morse families (Exercise 12.6.3).
2. Consider the Morse family of Example 12.4.7, which generates the figure eight immersion in $\mathbb{R}^2$ of Example 12.6.8/2. As in point 1, using p as a coordinate on L at the critical points $\eta_\pm = (\pm 1, 0)$, we find that at η_+ the coorientation points from the upper half-plane to the lower half-plane and that at η_- it points in the converse direction. Let us determine the Maslov intersection index of the following four curves. Denote $\zeta_\pm = (0, \pm 1)$ (the self-intersection points) and $\xi_{\pm\pm} := (\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}})$. Define γ_1 to run from ξ_{++} through η_+ to ξ_{+-}, γ_2 from ξ_{+-} through ζ_- to ξ_{--}, γ_3 from ξ_{--} through η_- to ξ_{-+} and γ_4 from ξ_{-+} through ζ_+ back to ξ_{++}. Since γ_1 traverses η_+ in the direction of the coorientation, whereas γ_3 traverses η_- in the direction opposite to the coorientation,

we find

$$\operatorname{ind}_L(\gamma_1) = 1, \qquad \operatorname{ind}_L(\gamma_2) = 0, \qquad \operatorname{ind}_L(\gamma_3) = -1, \qquad \operatorname{ind}_L(\gamma_4) = 0.$$

In particular, the closed curve obtained by composing $\gamma_1, \dots, \gamma_4$ has Maslov index 0. This is consistent with Corollary 12.6.12.

Next, we show that there is a variety of bundle structures over L associated with a given Lagrangian immersion $\iota : L \to \mathrm{T}^* Q$. These bundles contain information about how the pieces of this immersion generated by Morse families are glued together. Let $\mathscr{L}(L, \iota)$ denote the class of Morse families over Q locally generating $\iota : L \to \mathrm{T}^* Q$. Recall that the elements (B, π, S) of $\mathscr{L}(L, \iota)$ are characterized by the property that there exists an open subset U_S of L such that the immersions (B_S, Λ_S) and $(U_S, \iota_{\restriction U_S})$ are equivalent. With any two elements S_1, S_2 of (L, ι) such that $U_{S_1} \cap U_{S_2}$ is nonempty, one can associate a mapping

$$c_{S_1,S_2} : U_{S_1} \cap U_{S_2} \to \mathbb{Z}, \qquad c_{S_1,S_2}(\xi) := \operatorname{index}\big((S_1)''_{y_1 y_1}(\xi)\big) - \operatorname{index}\big((S_2)''_{y_2 y_2}(\xi)\big),$$

called the transition function of S_1 and S_2. Here, y_1 and y_2 are arbitrarily chosen[16] fibre coordinates on B_1 and B_2, respectively. By Lemma 12.6.10, c_{S_1,S_2} is constant on each connected component of $U_{S_1} \cap U_{S_2}$ and hence smooth. Now, take the subset of $L \times \mathscr{L}(L, \iota) \times \mathbb{Z}$ consisting of the elements (ξ, S, k) such that $\xi \in U_S$ and define M_L to be the quotient of this subset by the equivalence relation

$$(\xi_1, S_1, k_1) \sim (\xi_2, S_2, k_2) \quad \text{iff} \quad \xi_1 = \xi_2 \quad \text{and} \quad k_1 - k_2 = c_{S_1,S_2}(\xi_1).$$

From the direct product $L \times \mathscr{L}(L, \iota) \times \mathbb{Z}$, the quotient M_L inherits the natural projection

$$\pi^{M_L} : M_L \to L, \qquad \pi^{M_L}\big([(\xi, S, k)]\big) := \xi.$$

The elements S of $\mathscr{L}(L, \iota)$ define mappings

$$\chi_S : \big(\pi^{M_L}\big)^{-1}(U_S) \to U_s \times \mathbb{Z}, \qquad \chi_S\big([(\xi, \tilde{S}, k)]\big) := \big(\xi, k + c_{S,\tilde{S}}(\xi)\big),$$

which are easily seen to be bijective. The transition mappings are given by

$$\chi_{S_2} \circ \chi_{S_1}^{-1}(\xi, k) = \big(\xi, k + c_{S_2,S_1}(\xi)\big).$$

Since they are smooth, the family $\{\chi_S : S \in \mathscr{L}(L, \iota)\}$ defines on M_L the structure of a smooth manifold, cf. Remark 1.1.10. Second countability thereby carries over from L, because the latter implies that the covering $\{U_S : S \in \mathscr{L}(L, \iota)\}$ contains a countable subcovering. This way, M_L becomes a locally trivial fibre bundle over L

[16]For any two choices, the quadratic forms $S''_{yy}(\xi)$ are similar and hence have the same index.

with canonical projection π^{M_L} and typical fibre $\mathbb{Z}$. Finally, one can check that the mapping

$$\Psi : M_L \times \mathbb{Z} \to M_L, \qquad \Psi\big(\big[(\xi, S, k)\big], l\big) := \big[(\xi, S, k+l)\big],$$

is well-defined and that it endows M_L with the structure of a principal $\mathbb{Z}$-bundle over L. This bundle is called the Maslov principal bundle. With the Maslov principal bundle, there come the following three associated bundles:

(a) $M_L \times_{\mathbb{Z}} \mathbb{Z}_4$ via the induced action of $\mathbb{Z}$ on $\mathbb{Z}_4 = \mathbb{Z}/4\mathbb{Z}$ by translation,
(b) $M_L \times_{\mathbb{Z}} \mathrm{U}(1)$ via the action of $\mathbb{Z}$ on $\mathrm{U}(1)$ defined by $(k, \alpha) \mapsto \mathrm{e}^{\mathrm{i}\frac{\pi}{2}k}\alpha$,
(c) $\mathscr{M}_L := M_L \times_{\mathbb{Z}} \mathbb{C}$ via the corresponding action of $\mathbb{Z}$ on $\mathbb{C}$.

$\mathscr{M}_L$ is called the Maslov line bundle. The embedding

$$\mathbb{Z}_4 \to U(1), \quad k \bmod 4 \mapsto \mathrm{e}^{\mathrm{i}\frac{\pi}{2}k},$$

induces a vertical principal subbundle embedding $M_L \times_{\mathbb{Z}} \mathbb{Z}_4 \to M_L \times_{\mathbb{Z}} \mathrm{U}(1)$. Using sheaf theory one can prove that there exists a family of functions

$$\big\{c_S \in C^\infty(U_S) : S \in \mathscr{L}(L, \iota)\big\}$$

such that $c_{S_1} - c_{S_2} = c_{S_1,S_2}$ on $U_{S_1} \cap U_{S_2}$ [133, Prop. 2.11.1]. Then,

$$\mathrm{e}^{\mathrm{i}\frac{\pi}{2}c_{S_1}} = \mathrm{e}^{\mathrm{i}\frac{\pi}{2}c_{S_1,S_2}}\mathrm{e}^{\mathrm{i}\frac{\pi}{2}c_{S_2}} \tag{12.6.26}$$

and thus the local sections $\mathrm{e}^{\mathrm{i}\frac{\pi}{2}c_S}$ in $M_L \times_{\mathbb{Z}} \mathrm{U}(1)$ combine to a global non-vanishing section. As a consequence, $M_L \times_{\mathbb{Z}} \mathrm{U}(1)$, and hence the Maslov line bundle $\mathscr{M}_L$, is trivial. The triviality of the complex vector bundle $\mathscr{M}_L$ does, however, not imply the existence of a single real-valued generating function for L.

Remark 12.6.14 Let $\{S_i : i \in I\}$ be a countable subset of $\mathscr{L}(L, \iota)$ such that the subsets $U_i \equiv U_{S_i}$ cover L. The covering $\{U_i\}$, together with the family of transition mappings $c_{ij} \equiv c_{S_i,S_j}$ of the corresponding system of local trivializations of the Maslov principal bundle M_L, defines a 1-cocycle on L with values in $\mathbb{Z}$ and thus an element $\hat{\mu}_L$ of the first integer-valued Čech cohomology $H^1_{\check{\mathrm{C}}}(L, \mathbb{Z})$ of L, cf. Remark 2.2.12/2. According to this remark, $\hat{\mu}_L$ uniquely characterizes the principal $\mathbb{Z}$-bundle M_L up to isomorphisms. Corollary 12.6.12 implies that by the natural homomorphism

$$H^1_{\check{\mathrm{C}}}(L, \mathbb{Z}) \to H^1(L, \mathbb{R}),$$

$\hat{\mu}_L$ is mapped to the Maslov class μ_L. Let us add that the transition mappings of the corresponding system of local trivializations of the Maslov line bundle $\mathscr{M}_L$ are given by $\mathrm{e}^{-\mathrm{i}\frac{\pi}{2}c_{ij}}$.

In the remainder of this section, we derive a local normal form for the caustic of a Lagrangian immersion in the simplest case.

Proposition 12.6.15 *Let* $\iota : L \to \mathrm{T}^*Q$ *be a Lagrangian immersion and let* $\xi_0 \in \Sigma_1(L)$. *Denote* $x_0 = \Pi(\xi_0)$. *Assume that*

$$\iota' \mathrm{T}_{\xi_0}\big(\Sigma_1(L)\big) \cap \mathrm{T}_{\iota(\xi_0)}\big(\mathrm{T}^*_{x_0} Q\big) = \{0\}. \tag{12.6.27}$$

Then, there exist smooth functions f *and* g *on a neighbourhood* U *of* x_0 *in* Q *with* $g' \neq 0$ *such that* $\Gamma(L) \cap U$ *coincides with the caustic* $\Gamma(B_S)$ *of the Lagrangian immersion* Λ_S *generated by the Morse family*

$$S : U \times \mathbb{R} \to \mathbb{R}, \qquad S(x, y) = f(x) + g(x)y - \frac{1}{3}y^3. \tag{12.6.28}$$

Clearly, for the normal form (12.6.28), we have

$$B_S = \big\{(x, y) \in U \times \mathbb{R} : g(x) = y^2\big\}, \qquad \Gamma(B_S) = \big\{x \in U : g(x) = 0\big\}. \tag{12.6.29}$$

Proof The proof is along the lines of the proof of Proposition 6.1 in Chap. VII of [115]. Let (B, π, S) be some Morse family generating $\iota : L \to \mathrm{T}^*Q$ at ξ_0. By Lemma 12.5.1 and by the Splitting Lemma 12.5.2, we may assume that B is an open subset of $Q \times \mathbb{R}$. Let $y_0 \in \mathbb{R}$ such that $(x_0, y_0) \in B_S$ and $\Lambda_S(x_0, y_0) = \xi_0$. In a first step, we show that there exist smooth functions $f, g \colon Q \to \mathbb{R}$ and $\chi : \mathbb{R} \to \mathbb{R}$ such that for $(x, y) \in B_S$ we have

$$S(x, y) = f(x) + g(x)\chi(x, y) - \frac{1}{3}\chi(x, y)^3, \qquad \big(\chi(x, y)\big)^2 = g(x), \tag{12.6.30}$$

where $g'(x) \neq 0$ for all x and $\frac{\partial \chi}{\partial y}(y) \neq 0$ for all y. We will give the argument for the case $\dim Q = 1$. The general case can be reduced to this case, see below. By a constant shift of the fibre coordinate y, we may achieve that $y_0 = 0$. Choosing an appropriate coordinate x on Q we may also assume that $x_0 = 0$. Then, since $\Lambda_S(0, 0) = \xi_0$ belongs to $\Sigma_1(L)$, we have $S''_{yy}(0, 0) = 0$ and hence $S''_{xy}(0, 0) \neq 0$. Thus, by the Implicit Function Theorem, the equation $\frac{\partial S}{\partial y}(x, y) = 0$ can be solved for x and hence B_S is given by a smooth function $y \mapsto x(y)$. Then, (12.4.15) implies

$$\frac{\partial^2 S}{\partial y \partial x}\big(x(y), y\big)x'(y) + \frac{\partial^2 S}{\partial y^2}\big(x(y), y\big) = 0 \tag{12.6.31}$$

and hence $x'(0) = 0$. Differentiating (12.6.31) once again, we obtain

$$\frac{\partial^2 S}{\partial y \partial x}(0, 0)x''(0) + \frac{\partial^3 S}{\partial y^3}(0, 0) = 0. \tag{12.6.32}$$

Due to the assumption (12.6.27), S can be chosen so that $\frac{\partial^3 S}{\partial y^3}(0, 0) \neq 0$. The proof of this statement is left to the reader, see Exercise 12.6.4. Then, (12.6.32) implies that $x''(0) \neq 0$. Therefore, the Taylor expansion of $x(y)$ at $y = 0$ starts with the second order term. By replacing the coordinate x on Q by the coordinate $-x$ if necessary, we may assume that the corresponding coefficient is positive and hence that $x(y) \geq 0$ in some neighbourhood of $y = 0$. Then, by taking the positive square root of $x(y)$ for $y > 0$ and the negative square root of $x(y)$ for $y < 0$, we obtain a smooth function $y \mapsto \lambda(y)$ satisfying $x(y) = \lambda(y)^2$ and $\lambda'(0) > 0$. It follows that the mapping $(x, y) \mapsto (x, \lambda(y))$ is a fibre-preserving diffeomorphism of B which transforms S in such a way that B_S is given by $x = y^2$.

Now, consider the function $y \mapsto \frac{1}{2}(S(y^2, y) + S(y^2, -y))$. Since it is even, a lemma of Whitney's[17] yields that there exists a smooth function f on $\mathbb{R}$ such that

$$f(y^2) = \frac{1}{2}\big(S(y^2, y) + S(y^2, -y)\big).$$

Next, consider the function

$$\psi(y) := \frac{3}{4}\big(S(y^2, y) - S(y^2, -y)\big).$$

A brief computation shows that $\psi(0) = \psi'(0) = \psi''(0) = 0$ and $\psi'''(0) \neq 0$. Hence, the Taylor expansion of ψ at 0 starts with the third order term, so that by taking the third root of $\psi(y)$ we obtain a unique smooth function χ on a neighbourhood of zero in $\mathbb{R}$ such that $\chi(y)^3 = \psi(y)$. Since ψ is an odd function, so is χ. Hence, χ^2 is an even function, so that by the above lemma of Whitney's there exists a smooth function g on $\mathbb{R}$ such that $g(y^2) = \chi(y)^2$. Then,

$$g(y^2)\chi(y) = \frac{3}{4}\big(S(y^2, y) - S(y^2, -y)\big).$$

A brief calculation shows that the functions f, g and χ so constructed satisfy (12.6.30), indeed. Moreover, since $(\chi'(0))^3 = \frac{1}{6}\psi'''(0) \neq 0$ and $g'(0) = (\chi'(0))^2$, we have $\chi'(y) \neq 0$ and hence $g'(x) \neq 0$ in some neighbourhood of $x = 0$.

Finally, in the case $\dim Q > 1$, since $S''_{xy}(0, 0) \neq 0$, we can choose coordinates x^i in a neighbourhood of x_0 in Q such that

$$S''_{yx^1} \neq 0.$$

Then, we can carry out the argument for $\dim Q = 1$, thereby treating the variables $x^2, \dots, x^n$ as parameters. Due to the parameterized version of the lemma of Whitney's cited above, and since in each step of the above construction of the functions $f(x^1)$, $g(x^1)$ and $\chi(y)$, the smooth dependence on the parameters $x^2, \dots, x^n$ is preserved, these functions depend smoothly on the parameters and thus yield smooth functions on Q and $Q \times \mathbb{R}$, respectively. This completes the first step of the proof of Proposition 12.6.15.

[17] If h is an even smooth function on $\mathbb{R}$, there exists a unique smooth function $\tilde{h}$ on $\mathbb{R}$ such that $h(y) = \tilde{h}(y^2)$. If h depends smoothly on parameters, then so does $\tilde{h}$.

In the second step, we define a function $\tilde{S}$ in a neighbourhood of (x_0, y_0) in B by the right hand side of the first equation in (12.6.30). This function fulfils

$$\frac{\partial \tilde{S}}{\partial y}(x, y) = \big(g(x) - \chi(x, y)^2\big)\frac{\partial \chi}{\partial y}(x, y).$$

Since $g(x) = \chi(x, y)^2$ for all $(x, y) \in B_S$ and since in the coordinates x^i on Q used above, $\frac{\partial g}{\partial x^1} \neq 0$ and $\frac{\partial \chi}{\partial x^1} = 0$ everywhere, we conclude that $\frac{\partial \tilde{S}}{\partial y}(x, y) = 0$ iff $(x, y) \in B_S$. Thus, the singular subset of $\tilde{S}$ coincides with B_S. Moreover, choosing coordinates x^i on Q in a neighbourhood of x_0 such that $\frac{\partial^2 S}{\partial y \partial x^1} \neq 0$ and parameterizing B_S by the coordinates $y, x^2, \dots, x^n$, from (12.4.15) we conclude that both the singular points of S in B_S, defined by $\frac{\partial^2 S}{\partial y^2}(x, y) = 0$, and the singular points of $\tilde{S}$ in B_S, defined by $\frac{\partial^2 \tilde{S}}{\partial y^2}(x, y) = 0$, are characterized in these coordinates by the equation

$$\frac{\partial x^1}{\partial y}\big(y, x^2, \dots, n\big) = 0.$$

Thus, $\tilde{S}$ has the same singular subset as S and hence it generates the caustic $\Gamma(L)$ in a neighbourhood of x_0 in Q. Finally, we apply the fibre-preserving diffeomorphism $(x, y) \mapsto (x, \chi(x, y))$ transforming $\tilde{S}$ to the Morse family (12.6.28). This finishes the proof. □

For an application of this proposition in physics, we refer to Example 12.8.11.

Remark 12.6.16 (Caustics and Catastrophe Theory) We put the above proposition in the perspective of a typology of singularities of Lagrangian immersions. Recall that, by Theorem 12.5.5, germs of Lagrangian immersions at a point $\xi \in \mathrm{T}^* Q$ are in one-to-one correspondence with stable equivalence classes of Morse families at ξ. This correspondence can be carried over to the following situation.

On the one hand, two Lagrangian immersions $\iota : L \to \mathrm{T}^* Q$ and $\tilde{\iota} : \tilde{L} \to \mathrm{T}^* \tilde{Q}$ are said to be equivalent if there exist diffeomorphisms $\lambda : L \to \tilde{L}$ and $\psi : Q \to \tilde{Q}$ and a symplectomorphism $\varphi : \mathrm{T}^* Q \to \mathrm{T}^* \tilde{Q}$ such that

$$\tilde{\iota} \circ \lambda = \varphi \circ \iota, \qquad \psi \circ \pi_Q = \pi_{\tilde{Q}} \circ \varphi.$$

This induces an equivalence relation for germs of Lagrangian immersions in an obvious way. Caustics of equivalent Lagrangian immersions are mapped diffeomorphically onto one another. On the other hand, the concept of stable equivalence of Morse families at some point $\xi \in \mathrm{T}^* Q$ discussed in Sect. 12.5 can be generalized to stable equivalence of arbitrary Morse families by extending the operation of composition to include arbitrary fibre-preserving diffeomorphisms $\mathrm{T}^* Q \to \mathrm{T}^* \tilde{Q}$, projecting to diffeomorphisms $Q \to \tilde{Q}$ and allowing for arbitrary smooth functions on Q in the operation of addition. For reduced Morse families, the equivalence relation

Table 12.1 Normal forms for Morse families over Q with $\dim Q \le 4$

Name	Normal form
Fold	$y^3 + xy$
Cusp	$\pm y^4 + x_1 y^2 + x_2 y$
Swallowtail	$y^5 + x_1 y^3 + x_2 y^2 + x_3 y$
Butterfly	$\pm y^6 + x_1 y^4 + x_2 y^3 + x_3 y^2 + x_4 y$
Hyperbolic umbilic	$y_1^3 + y_2^3 + x_1 y_1 y_2 + x_2 y_1 + x_3 y_2$
Elliptic umbilic	$y_1 y_2^2 - y_1^3 + x_1 y_1^2 + y_2^2 + x_2 y_1 + x_3 y_2$
Parabolic umbilic	$y_1 y_2^2 \pm y_1^4 + x_1 y_1^2 + x_2 y_2^2 + x_3 y_1 + x_4 y_2$

reads

$$\tilde{S} = S \circ \varphi + \chi ,$$

with $\varphi : \mathrm{T}^* Q \to \mathrm{T}^* \tilde{Q}$ being a fibre-preserving diffeomorphism and $\chi : Q \to \mathbb{R}$ being a smooth function. Then, equivalence classes of germs of Lagrangian immersions are in bijective correspondence with stable equivalence classes (in the above generalized sense) of Morse families, see [14–16] and [19, §1.3]. Thus, the classification of Lagrangian immersions and their singularities (singular subsets) is reduced to the classification of Morse families, for which methods of general singularity theory can be applied, see [115]. In this context, the generating Morse family is referred to as an unfolding of the singularity. First, one analyzes the Taylor series to obtain normal forms similar to (12.6.28), but containing error terms which for $x = 0$ vanish to arbitrary order. Next, one shows that up to stable equivalence the error terms can be omitted. Thus, in particular, the normal form of Proposition 12.6.15 is stably equivalent, in the generalized sense, to the original Morse family. This way, one can classify, for example, the stable[18] Morse families over Q with $\dim Q \le 4$, see Table 12.1. This yields the famous Thom catastrophes. We see that the catastrophe described by Proposition 12.6.15 is a fold. The list in Table 12.1 is a classical result of singularity theory. This theory, which is sometimes also referred to as catastrophe theory, was developed by Whitney [313], Thom [287, 288], Mather [200–204], Boardman [49] and Arnold [14–16], see also [20, 24, 25, 110, 115]. Generally speaking, in this theory one studies the singularity structure of a smooth mapping $\psi : M \to Q$ between manifolds. In the first step, in complete analogy to the singular subset $\Sigma(L)$ of Π, one defines the singular set

$$\Sigma_i(\psi) := \left\{ m \in M : \dim \ker \psi'_m = i \right\}, \quad i = 0, \dots, \dim M,$$

of ψ. If $\Sigma_i(\psi)$ is a submanifold of M, one can define

$$\Sigma_{i,j}(\psi) := \Sigma_j(\psi_{\restriction \Sigma_i(\psi)}), \quad 0 \le j \le i.$$

[18] A Morse family over Q is stable if it is an inner point of its stable equivalence class with respect to a certain C^∞-topology on the space of all Morse families over Q.

Again, if this is a submanifold, one can go on with defining $\Sigma_{i,j,k}(\psi)$ and so on. These subsets are called the Thom-Boardman singularities.[19] As a result, one obtains a partition of M into a family of locally closed submanifolds with the property that the restriction of ψ to each component has maximal rank. In some special cases, the Thom-Boardman singularities yield a complete classification of generic mappings. The Thom catastrophes listed above are of this type. The fold, the cusp, the swallowtail and the butterfly correspond to, respectively, the Thom-Boardman singularities $\Sigma_{1,0}$, $\Sigma_{1,1,0}$, $\Sigma_{1,1,1,0}$ and $\Sigma_{1,1,1,1}$, and the umbilic catastrophes are of the type $\Sigma_{2,0}$. In the case under consideration, there are no further Thom-Boardman singularities, because all other singularities have a codimension greater than 4.

Exercises

12.6.1 Show that the figure eight immersion of Example 12.6.8/2 has vanishing Maslov index.

12.6.2 Show that the function $S(x, y_1, y_2) := -\frac{1}{3}y_1^3 - \frac{1}{3}y_2^3 + xy_1 + (1-x)y_2$ defines a Morse family. Determine the induced Lagrangian immersion, its singular subset and its caustic.

12.6.3 Find a system of generating Morse families for the unit circle in $\mathbb{R}^2$, cf. Example 12.6.13/1.

12.6.4 Complete the proof of Proposition 12.6.15 by showing that under the assumption (12.6.27), the generating family S can be chosen so that $\frac{\partial^3 S}{\partial y^3}(0,0) \neq 0$.
Hint. Study the kernel $\ker((\Pi')_{\restriction T\Sigma_1(L)})$ in an analogous way as in the proof of Lemma 12.5.1.

12.7 Geometric Asymptotics. The Eikonal Equation

In this section we apply the method of characteristics and the concept of Morse families to the equation of geometric optics, the so-called eikonal equation. We restrict our attention to the case $Q = \mathbb{R}^n$.

Geometric optics rests on the assumption that the wavelength λ of light is small compared with the typical length scale L of the optical system under consideration. Under this assumption, the wave character of light remains hidden and one may imagine light as a flow of particles (light rays).

For simplicity, we consider the scalar wave equation[20] on $Q \times \mathbb{R}$:

$$\left(\frac{n^2(\mathbf{x})}{c^2}\frac{\partial^2}{\partial t^2} - \Delta\right)u(\mathbf{x}, t) = 0, \tag{12.7.1}$$

[19] While one can show that for almost all functions f these subsets are submanifolds indeed [287, 288], the Thom-Boardman singularities can also be defined in the general case by using jet techniques as developed by Boardman.

[20] Thus, in particular, we ignore polarization phenomena.

where $n(\mathbf{x})$ is the local refractive index. Making the ansatz $u(t,\mathbf{x}) = u(\mathbf{x})e^{-i\omega t}$, we obtain the associated Helmholtz equation,

$$\left(\frac{1}{k^2}\Delta + n^2(\mathbf{x})\right)u(\mathbf{x}) = 0, \tag{12.7.2}$$

where $k = \frac{\omega}{c}$. For length scales r fulfilling $\lambda \ll r \ll L$, one can study this equation in the framework of short wave analysis. The starting point of this procedure is the ansatz

$$u(\mathbf{x},k) = a(\mathbf{x},k)e^{ikS(\mathbf{x})}. \tag{12.7.3}$$

The function S is called the eikonal function. It has the following physical interpretation.

(a) The equations $S(\mathbf{x}) = c$ describe surfaces of constant phase, called wave fronts.
(b) By expanding $S(\mathbf{x}) = S(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)\cdot\nabla S(\mathbf{x}_0) + \cdots$ at a given point $\mathbf{x}_0$, we see that, close to $\mathbf{x}_0$, (12.7.3) can be approximated by a plane wave with wave vector

$$\mathbf{k}(\mathbf{x}_0) = k\mathbf{n}(\mathbf{x}_0), \qquad \mathbf{n}(\mathbf{x}_0) := \nabla S(\mathbf{x}_0).$$

The vector-valued functions $\mathbf{k}$ and $\mathbf{n}$ are referred to as the local wave vector and the local refractive index vector, respectively. They are orthogonal to the wave fronts and we have $\mathbf{n}(\mathbf{x})^2 = n^2(\mathbf{x})$ for all $\mathbf{x}$.

Inserting the ansatz (12.7.3) into the Helmholtz equation, we obtain

$$\left(\frac{1}{k^2}\Delta a + \frac{2i}{k}\nabla a\cdot\nabla S + \frac{ia}{k}\Delta S - a(\nabla S)^2 + n^2 a\right)e^{ikS} = 0. \tag{12.7.4}$$

While the amplitude a, as a function of $\mathbf{x}$, varies on the length scale of the optical system, the eikonal function S varies on the length scale of the wave length $\lambda = \frac{2\pi}{k}$. Therefore, it makes sense to expand $a(\mathbf{x},k)$ in powers of $\frac{1}{k}$,

$$a(\mathbf{x},k) = a_0(\mathbf{x}) + \frac{1}{k}a_1(\mathbf{x}) + \frac{1}{k^2}a_2(\mathbf{x}) + \cdots . \tag{12.7.5}$$

Plugging in this expansion into (12.7.4) and comparing coefficients, we obtain the eikonal equation

$$(\nabla S)^2 = n^2 \tag{12.7.6}$$

in zeroth order of $\frac{1}{k}$ and the transport equation

$$\nabla S\cdot\nabla \ln a_0^2 + \Delta S = 0 \tag{12.7.7}$$

in first order of $\frac{1}{k}$. The eikonal equation is the Hamilton-Jacobi equation for the Hamiltonian function

$$H(\mathbf{x},\mathbf{p}) = \mathbf{p}^2 - n^2(\mathbf{x}) \tag{12.7.8}$$

on $\mathrm{T}^*Q = \mathbb{R}^n \times \mathbb{R}^n$, that is, for a particle with mass $\frac{1}{2}$ moving in the potential $n^2(\mathbf{x})$. It determines the eikonal function S. By plugging in the solution into the transport equation, the latter becomes an equation in the indeterminate a_0. Let us discuss the physical meaning of this equation. By a standard calculation, one can derive the following formulae for the energy density ρ and the energy current density[21] $\mathbf{P}$ of the scalar field u:

$$\rho(t, \mathbf{x}) = \frac{1}{2}\left(\frac{n^2}{c^2}\left(\frac{\partial u}{\partial t}(t, \mathbf{x})\right)^2 + \left(\nabla u(t, \mathbf{x})\right)^2\right),$$

$$\mathbf{P}(t, \mathbf{x}) = -\frac{\partial u}{\partial t}(t, \mathbf{x})\nabla u(t, \mathbf{x}).$$

Taking the time average, denoted by $\langle\cdot\rangle$, in leading order of $\frac{1}{k} \to 0$ we find

$$\langle\rho\rangle = \frac{1}{2}n^2k^2a_0^2, \qquad \langle\mathbf{P}\rangle = \frac{1}{2}ck^2a_0^2\nabla S. \tag{12.7.9}$$

Thus, the energy flows in the direction of the vector field ∇S, that is, in the direction of the local wave vector $\mathbf{k}$. Since by the eikonal equation, $\frac{\nabla S}{n}$ is a unit vector, we have

$$\|\langle\mathbf{P}\rangle\| = \frac{c}{n}\langle\rho\rangle.$$

This means that the velocity of the energy flow is given by the local phase velocity $\frac{c}{n}$. Moreover, taking the divergence of $\langle\mathbf{P}\rangle$ in (12.7.9), we find that the transport equation (12.7.7) describes energy conservation. From this discussion, we conclude that we may view light rays as flow lines of the vector field ∇S.[22] This interpretation constitutes the basis for the discussion below.

Remark 12.7.1 More generally, in the same spirit one can study the asymptotic behaviour for $k \to \infty$ of partial differential equations of the type

$$H\left(\mathbf{x}, -\frac{i}{k}\nabla\right)u(\mathbf{x}) = 0, \tag{12.7.10}$$

where

$$H : \mathrm{T}^*Q \cong \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

is a smooth function which is polynomial in the fibre variables (the momenta),

$$H(\mathbf{x}, \mathbf{p}) = h(\mathbf{x}) + \sum_r h_{i_1,\dots,i_r}(\mathbf{x})p_{i_1}\dots p_{i_r}.$$

[21] Built in analogy to the energy density and of the Poynting vector in Maxwell electrodynamics.

[22] More precisely, one may view light rays as wave packets, whose width in the direction transversal to the energy current vector is negligible. That such wave packets can be prepared follows from the uncertainty relation for the Fourier transform in the short wave approximation.

This function is referred to as the symbol of the differential operator given by (12.7.10). In the special case of the Helmholtz equation, the symbol is given by (12.7.8). Making the ansatz (12.7.3) and expanding the amplitude a according to (12.7.5), from (12.7.10) we obtain

$$\begin{aligned} 0 =& H\big(\mathbf{x}, \nabla S(\mathbf{x})\big) a_0(\mathbf{x}) \\ & - \frac{\mathrm{i}}{k}\left(\frac{a_0(\mathbf{x})}{2}\frac{\partial^2 H}{\partial p_i \partial p_j}\big(\mathbf{x}, \nabla S(\mathbf{x})\big)\frac{\partial^2 S}{\partial x^i \partial x^j}(\mathbf{x}) + \frac{\partial H}{\partial p_i}\big(\mathbf{x}, \nabla S(\mathbf{x})\big)\frac{\partial a_0}{\partial x^i}(\mathbf{x})\right) \\ & + \frac{1}{k} H\big(\mathbf{x}, \nabla S(\mathbf{x})\big) a_1(\mathbf{x}) + O\left(\frac{1}{k^2}\right). \end{aligned}$$

Comparison of coefficients yields the characteristic equation

$$H\big(\mathbf{x}, \nabla S(\mathbf{x})\big) = 0 \tag{12.7.11}$$

in zeroth order of $\frac{1}{k}$ and the transport equation

$$\left(\frac{\partial H}{\partial p_i}\big(\mathbf{x}, \nabla S(\mathbf{x})\big)\frac{\partial}{\partial x^i} + \frac{1}{2}\frac{\partial^2 H}{\partial p_i \partial p_j}\big(\mathbf{x}, \nabla S(\mathbf{x})\big)\frac{\partial^2 S}{\partial x^i \partial x^j}(\mathbf{x})\right) a_0(\mathbf{x}) = 0 \tag{12.7.12}$$

in first order of $\frac{1}{k}$. Equation (12.7.11) is the Hamilton-Jacobi equation for the Hamiltonian function H. In the special case of the Helmholtz equation, (12.7.11) reproduces the eikonal equation (12.7.6) and (12.7.12) reproduces the transport equation (12.7.7). Note that all of this carries over to an arbitrary Riemannian manifold (Q, g). For example, in this case the eikonal equation reads

$$\mathrm{g}(\nabla S, \nabla S) = n^2$$

and the corresponding Hamiltonian function $H : \mathrm{T}^* Q \to \mathbb{R}$ is given by

$$H(\xi) = \mathrm{g}^{-1}(\xi, \xi) - n^2\big(\pi(\xi)\big). \tag{12.7.13}$$

Now, let us study the eikonal equation (12.7.6). For the sake of clarity, we restrict our attention to the vacuum case $n = 1$,

$$(\nabla S)^2 = 1. \tag{12.7.14}$$

As noted above, this is the Hamilton-Jacobi equation for the Hamiltonian function

$$H(\mathbf{x}, \mathbf{p}) = \mathbf{p}^2 - 1,$$

which up to a constant energy shift models a free particle of mass $\frac{1}{2}$. To solve this equation, we apply the method of characteristics, cf. Sect. 12.2 and in particular Theorem 12.2.1. Let

$$\mathscr{C} = H^{-1}(0) \subset \mathrm{T}^* Q,$$

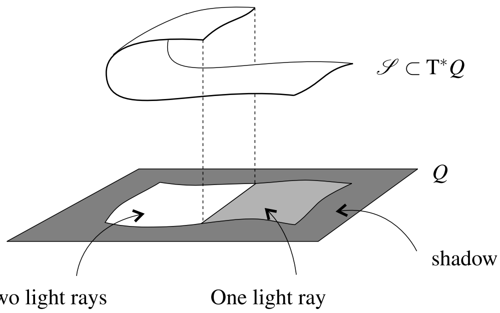


Fig. 12.3 Light rays as projections of characteristics to Q

let $D \subset Q$ be an oriented embedded submanifold of dimension $r < n$ and let S_0 be a smooth function on D. Assume that the canonical lift $\widehat{(D, S_0)}$ of (D, S_0), defined in Example 8.3.8/4, is transversal to $\mathscr{C}$ and that the intersection $\mathscr{S}_0 = \widehat{(D, S_0)} \cap \mathscr{C}$ is transversal in $\mathscr{C}$ to the integral curves of X_H, so that $\mathscr{S}_0$ is an admissible initial condition. The Hamiltonian vector field X_H is given by

$$X_H(\mathbf{x}, \mathbf{p}) = 2p_i \partial_{x^i}, \tag{12.7.15}$$

and the Hamilton equations read

$$\dot{\mathbf{p}} = 0, \qquad \dot{\mathbf{x}} = 2\mathbf{p}. \tag{12.7.16}$$

The integral curves, and thus the characteristics of $\mathscr{C}$, are given by

$$\mathbf{p}(t) = \mathbf{p}_0, \qquad \mathbf{x}(t) = \mathbf{x}_0 + 2\mathbf{p}_0 t, \tag{12.7.17}$$

where $(\mathbf{x}_0, \mathbf{p}_0) \in \mathscr{S}_0$. Thus, the method of characteristics yields the generalized solution $\iota : \mathscr{S} \to \mathrm{T}^*Q$, where

$$\mathscr{S} = \mathscr{S}_0 \times \mathbb{R}, \qquad \iota(\mathbf{x}_0, \mathbf{p}_0, t) = (\mathbf{x}_0 + t\mathbf{p}_0, \mathbf{p}_0). \tag{12.7.18}$$

The characteristics project to straight lines on Q. According to the above discussion, these straight lines can be interpreted as light rays. Thus, the generalized solution $(\mathscr{S}, \iota)$ has the following interpretation: D is a source emitting light rays with the prescribed phase S_0. The image $\Pi(\mathscr{S})$ is the region of Q which is illuminated, whereas its complement $Q \setminus \Pi(\mathscr{S})$ is the region which stays in the shadow.

The more times a fibre of T^*Q intersects $\mathscr{S}$, the more light rays run through its base point and, therefore, the brighter this point appears, see Fig. 12.3.

Remark 12.7.2 The observation that the projections of the characteristics to Q are straight lines generalizes to the case of an arbitrary Riemannian manifold (Q, g). Since the eikonal equation is the Hamilton-Jacobi equation of the Hamiltonian function (12.7.13), according to Example 9.2.1, the projections of the characteristics to Q are geodesics of the metric $n^{-2}\mathrm{g}$.

First, let us discuss in detail the case where the initial phase is $S_0 = 0$, that is, where D is a surface of constant phase. In this case, the canonical lift of (D, S_0) coincides with the conormal bundle $\widehat{D}$ of the submanifold D, cf. Example 8.3.8/3.

Lemma 12.7.3 *For every embedded submanifold D of Q of dimension $r < n$, the canonical lift $\widehat{D}$ is transversal to $\mathscr{C}$ and $\mathscr{S}_0 = \widehat{D} \cap \mathscr{C}$ is an admissible initial condition for $\mathscr{C}$.*

Proof In the following, let $a = 1, \ldots, r$ and $\alpha = r+1, \ldots, n$. Since D is embedded, we can find coordinates x^i on Q such that D is locally given by $x^\alpha = 0$. According to Example 8.3.8/3, in the bundle coordinates x^i and p_i induced on T^*Q, the canonical lift $\widehat{D}$ is given by

$$x^\alpha = 0, \qquad p_a = 0. \tag{12.7.19}$$

Correspondingly, in the fibre coordinates $\dot{x}^i$ and $\dot{p}_i$ induced on $\mathrm{T}(\mathrm{T}^*Q)$, for given $(\mathbf{x}, \mathbf{p}) \in D$, the tangent space $\mathrm{T}_{(\mathbf{x},\mathbf{p})}\widehat{D}$ is given by

$$\dot{x}^\alpha = 0, \qquad \dot{p}_a = 0. \tag{12.7.20}$$

Now, let $(\mathbf{x}, \mathbf{p}) \in \widehat{D} \cap \mathscr{C}$. In the coordinates x^i, the defining relation for $\mathscr{C}$ reads $\mathrm{g}^{ij}(\mathbf{x}) p_i p_j = 1$, with g^{ij} representing the Euclidean metric. Equation (12.7.19) implies $\mathrm{g}^{\alpha\beta}(\mathbf{x}) p_\alpha p_\beta = 1$. Thus, we can find α_0 such that $g^{\alpha_0\beta}(\mathbf{x}) p_\beta \neq 0$. Then, the tangent vector defined by $\dot{\mathbf{x}} = 0$ and $\dot{p}_i = \delta_{i\alpha_0}$ lies in $\mathrm{T}_{(\mathbf{x},\mathbf{p})}\widehat{D}$ and is transversal to $\mathrm{T}_{(\mathbf{x},\mathbf{p})}\mathscr{C}$. This shows that $\widehat{D}$ and $\mathscr{C}$ are transversal, so that $\mathscr{S}_0$ is an embedded submanifold of T^*Q and hence of $\mathscr{C}$. To prove that $\mathscr{S}_0$ is transversal in $\mathscr{C}$ to the integral curves of X_H, for dimensional reasons it suffices to show that X_H is nowhere tangent to $\mathscr{S}_0$. Thus, assume that $X_H(\mathbf{x}, \mathbf{p}) \in \mathrm{T}_{(\mathbf{x},\mathbf{p})}\widehat{D}$ for some $(\mathbf{x}, \mathbf{p}) \in \mathscr{S}_0$. Then, (12.7.15) and the first equation in (12.7.20) imply that $p_\alpha = 0$ for all α, in contradiction to $\mathrm{g}^{\alpha\beta}(\mathbf{x}) p_\alpha p_\beta = 1$. □

Example 12.7.4 Let $Q = \mathbb{R}^2$ and let D be the half circle

$$D = \left\{\mathbf{x} \in \mathbb{R}^2 : x_1^2 + x_2^2 = R^2, x_1 < 0\right\}.$$

Parameterizing D by $y \mapsto (R\cos(y), R\sin(y))$ with $\frac{\pi}{2} < y < \frac{3\pi}{2}$ and using the parameter y as a global coordinate on D, we find

$$\widehat{D} = \left\{\big(R\cos(y), R\sin(y), p, p\tan(y)\big) : \frac{\pi}{2} < y < \frac{3\pi}{2}, p \in \mathbb{R}\right\}$$

for the canonical lift and $\mathscr{S}_0 = \mathscr{S}_0^+ \cup \mathscr{S}_0^-$ with

$$\mathscr{S}_0^\pm = \left\{\big(R\cos(y), R\sin(y), \pm\cos(y), \pm\sin(y)\big) : \frac{\pi}{2} < y < \frac{3\pi}{2}\right\}$$

for the initial condition $\mathscr{S}_0$. Hence, by (12.7.18), $\mathscr{S}$ decomposes into the connected components $\mathscr{S}^\pm = \mathscr{S}_0^\pm \times \mathbb{R}$ and

$$\iota^\pm(y,t) = \big((R \pm 2t)\cos(y), (R \pm 2t)\sin(y), \pm\cos(y), \pm\sin(y)\big).$$

We see that all the light rays run through the origin, whereas through the other points, there run exactly two light rays in opposite directions, one from $\mathscr{S}_0^+$ and one from $\mathscr{S}_0^-$. The equation for the singular subset $\Sigma(\mathscr{S})$ is

$$\det \frac{\partial(\Pi_1, \Pi_2)}{\partial(y,t)} = 2(R \pm 2t) = 0,$$

where $\Pi = (\Pi_1, \Pi_2) \colon \mathscr{S} \to Q$, cf. (12.5.3). Therefore,

$$\Sigma(\mathscr{S}) = \left\{\big(0, 0, -\cos(y), -\sin(y)\big) \in \mathrm{T}^*Q : \frac{\pi}{2} < y < \frac{3\pi}{2}\right\},$$
$$\Gamma(\mathscr{S}) = \big\{(0,0) \in Q\big\}.$$

Thus, the caustic degenerates to a single focal point.

Example 12.7.5 Let $Q = \mathbb{R}^2$ and $D = \{\mathbf{x} \in \mathbb{R}^2 \colon (x_2)^2 = x_1\}$. Let us use $y = x_2$ as a coordinate on D. Then, the canonical lift $\widehat{D}$ is given by

$$\widehat{D} = \big\{\big(y^2, y, p, -2py\big) : y, p \in \mathbb{R}\big\}$$

and the initial condition $\mathscr{S}_0$ consists of the connected components

$$\mathscr{S}_0^\pm = \left\{\left(y^2, y, \frac{\pm 1}{\sqrt{1+4y^2}}, \frac{\mp 2y}{\sqrt{1+4y^2}}\right) : y \in \mathbb{R}\right\}.$$

Consequently, $\mathscr{S}$ consists of the connected components $\mathscr{S}^\pm = \mathscr{S}_0^\pm \times \mathbb{R}$ and ι is given by

$$\iota^\pm(y,t) = \left(y^2 \pm \frac{t}{\sqrt{1+4y^2}}, y \mp \frac{2ty}{\sqrt{1+4y^2}}, \frac{\pm 1}{\sqrt{1+4y^2}}, \frac{\mp 2y}{\sqrt{1+4y^2}}\right). \quad (12.7.21)$$

It is easy to see that ι is injective. Hence, $(\mathscr{S}, \iota)$ is in fact a geometric solution. The singular subset $\Sigma(\mathscr{S})$ and the caustic $\Gamma(\mathscr{S})$ will be discussed later in Example 12.7.8, where we solve the same initial value problem by means of a Morse family.

Next, we find a Morse family generating the generalized solution given by (12.7.18). This provides an alternative method to solve (12.7.14). We still limit our attention to the case of a constant initial phase $S_0 = 0$. In what follows, points on D will be denoted by $\hat{\mathbf{x}}$. We choose coordinates $y^1, \dots, y^r$ on D and define local vector fields on D by

$$\mathbf{e}_\alpha(\hat{\mathbf{x}}) = \frac{\partial \hat{\mathbf{x}}}{\partial y^\alpha}. \tag{12.7.22}$$

These vector fields form a local frame in $\mathrm{T}D$, viewed as a subset of $\mathrm{T}Q = \mathbb{R}^{2n}$.

Proposition 12.7.6 *Let D be an embedded submanifold of $Q = \mathbb{R}^n$ of dimension $r < n$ and let $B := \{(\mathbf{x}, \hat{\mathbf{x}}) \in Q \times D : \mathbf{x} \neq \hat{\mathbf{x}}\}$. The distance function*

$$S: B \to \mathbb{R}, \qquad S(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\|, \tag{12.7.23}$$

defines a Morse family along the submersion $B \to Q$ induced from the natural projection. This Morse family generates a generalized solution of the eikonal equation (12.7.14) *for the initial condition $S_0 = 0$ on D, given by*

$$B_S = \big\{(\mathbf{x}, \hat{\mathbf{x}}) \in B : (\mathbf{x} - \hat{\mathbf{x}}) \perp \mathrm{T}_{\hat{\mathbf{x}}} D\big\}, \qquad \Lambda_S(\mathbf{x}, \hat{\mathbf{x}}) = \left(\mathbf{x}, \frac{\mathbf{x} - \hat{\mathbf{x}}}{\|\mathbf{x} - \hat{\mathbf{x}}\|}\right). \tag{12.7.24}$$

Proof To see that S defines a Morse family, we have to show that the matrix (S''_{xy}, S''_{yy}) has rank r on the fibre-critical submanifold B_S. For $(\mathbf{x}, \hat{\mathbf{x}}) \in B$, we calculate

$$\frac{\partial S}{\partial x^i} = \frac{x_i - \hat{x}_i}{\|\mathbf{x} - \hat{\mathbf{x}}\|} =: p_i, \tag{12.7.25}$$

$$\frac{\partial S}{\partial y^\alpha} = -\mathbf{p} \cdot \mathbf{e}_\alpha, \tag{12.7.26}$$

$$\frac{\partial^2 S}{\partial x^i \partial y^\alpha} = \frac{(\mathbf{p} \cdot \mathbf{e}_\alpha) p_i - \mathbf{e}_{\alpha i}}{\|\mathbf{x} - \hat{\mathbf{x}}\|}, \tag{12.7.27}$$

$$\frac{\partial^2 S}{\partial y^\alpha \partial y^\beta} = \frac{\mathbf{e}_\alpha \cdot \mathbf{e}_\beta - (\mathbf{p} \cdot \mathbf{e}_\alpha)(\mathbf{p} \cdot \mathbf{e}_\beta)}{\|\mathbf{x} - \hat{\mathbf{x}}\|} - \mathbf{p} \cdot \frac{\partial \mathbf{e}_\beta}{\partial y^\alpha} \tag{12.7.28}$$

and read off that on B_S we have $\mathbf{p} \cdot \mathbf{e}_\alpha = 0$ and hence

$$\frac{\partial^2 S}{\partial x^i \partial y^\alpha}(\mathbf{x}, \hat{\mathbf{x}}) = -\frac{\mathbf{e}_{\alpha i}(\hat{\mathbf{x}})}{\|\mathbf{x} - \hat{\mathbf{x}}\|}.$$

Since the vectors $\mathbf{e}_\alpha(\hat{\mathbf{x}})$ are linearly independent, this matrix has rank r. Hence, S is a Morse family, indeed. Moreover, from (12.7.25) and (12.7.26) we read off (12.7.24). Since Λ_S takes values in $\mathscr{C}$, (B_S, Λ_S) is a generalized solution of (12.7.14). □

Remark 12.7.7

1. We determine the singular subset and the caustic of (B_S, Λ_S). By (12.6.14) and (12.7.28), the singular subset $\Sigma(B_S)$ consists of the points $(\mathbf{x}, \hat{\mathbf{x}}) \in B_S$ fulfilling

$$\det\left(\mathbf{e}_\alpha(\hat{\mathbf{x}}) \cdot \mathbf{e}_\beta(\hat{\mathbf{x}}) - (\mathbf{x} - \hat{\mathbf{x}}) \cdot \frac{\partial \mathbf{e}_\beta}{\partial y^\alpha}(\hat{\mathbf{x}})\right) = 0. \tag{12.7.29}$$

Hence, the caustic $\Gamma(B_S)$ consists of the points $\mathbf{x} \in Q$ for which there exists $\hat{\mathbf{x}} \in D$ such that $(\mathbf{x}, \hat{\mathbf{x}})$ belongs to B_S and fulfils (12.7.29).

Equation (12.7.29) can be formulated intrinsically in terms of the metric $\mathsf{h}_{\alpha\beta}$ on D and the exterior curvature $\mathbf{k}_{\alpha\beta}$ of D, both induced by the Euclidean metric on Q. For the necessary notions from Riemannian geometry, we refer to the standard literature, e.g. [6]. Let $\Gamma^{\gamma}_{\alpha\beta}$ denote the Christoffel symbols of the Levi-Civita connection on D associated with $\mathsf{h}_{\alpha\beta}$. Then, one has

$$\mathbf{e}_\alpha \cdot \mathbf{e}_\beta = \mathsf{h}_{\alpha\beta}, \qquad \frac{\partial \mathbf{e}_\beta}{\partial y^\alpha} = \mathbf{k}_{\alpha\beta} + \Gamma^{\gamma}_{\alpha\beta}\mathbf{e}_\gamma$$

and hence (12.7.29) becomes

$$\det\big(\mathsf{h}_{\alpha\beta}(\hat{\mathbf{x}}) - (\mathbf{x} - \hat{\mathbf{x}}) \cdot \mathbf{k}_{\alpha\beta}(\hat{\mathbf{x}})\big) = 0. \tag{12.7.30}$$

2. In the special case of an oriented surface $D \subset Q = \mathbb{R}^3$, we have

$$\mathbf{k}_{\alpha\beta} = k_{\alpha\beta}\mathbf{n}_D,$$

with $\mathbf{n}_D$ denoting the unit vector field orthogonal to D. Here, (12.7.30) yields

$$\det\big(\|\mathbf{x} - \hat{\mathbf{x}}\| k_{\alpha\beta}(\hat{\mathbf{x}}) - \mathsf{h}_{\alpha\beta}(\hat{\mathbf{x}})\big) = 0. \tag{12.7.31}$$

This is the characteristic equation for the principal radii of curvature of the initial surface D. Therefore, in this case the caustic coincides with the set of centres of curvature of D. Thus, no caustic occurs iff D is a plane.

3. Obviously, $S(\mathbf{x}, \hat{\mathbf{x}}) = -\|\mathbf{x} - \hat{\mathbf{x}}\|$ is a generating family, too. It has the same domain as the Morse family (12.7.23) and describes incoming rays. It is possible to describe incoming and outgoing rays by the help of a single Morse family: let

$$B = \big\{(\mathbf{x}, \hat{\mathbf{x}}, \mathbf{a}) \in Q \times D \times \mathrm{S}^{n-1} : \mathbf{x} \neq \hat{\mathbf{x}}\big\}$$

and define

$$S : B \to \mathbb{R}, \qquad S(\mathbf{x}, \hat{\mathbf{x}}, \mathbf{a}) = (\mathbf{x} - \hat{\mathbf{x}}) \cdot \mathbf{a}. \tag{12.7.32}$$

We leave it to the reader to check that this is a Morse family, indeed (Exercise 12.7.7). Let u^α and ϑ^i be coordinates on D and S^{n-1}, respectively. We calculate

$$\frac{\partial S}{\partial u^\alpha}(\mathbf{x}, \hat{\mathbf{x}}, \mathbf{a}) = -\mathbf{e}_\alpha(\hat{\mathbf{x}}) \cdot \mathbf{a}, \qquad \frac{\partial S}{\partial \vartheta^i}(\mathbf{x}, \hat{\mathbf{x}}, \mathbf{a}) = (\mathbf{x} - \hat{\mathbf{x}}) \cdot \frac{\partial \mathbf{a}}{\partial \vartheta^i}.$$

Thus, B_S is defined by the relations $\mathbf{a} \perp \mathrm{T}_{\hat{\mathbf{x}}}D$ and $(\mathbf{x} - \hat{\mathbf{x}}) \perp \mathrm{T}_{\mathbf{a}}\mathrm{S}^{n-1}$. The latter one requires $\mathbf{a}$ to be parallel or antiparallel to $\mathbf{x} - \hat{\mathbf{x}}$. Hence,

$$B_S = \big\{(\mathbf{x}, \hat{\mathbf{x}}, \mathbf{a}) \in B : (\mathbf{x} - \hat{\mathbf{x}}) \perp \mathrm{T}_{\hat{\mathbf{x}}}D \text{ and } \mathbf{x} - \hat{\mathbf{x}} = \pm\|\mathbf{x} - \hat{\mathbf{x}}\|\mathbf{a}\big\}.$$

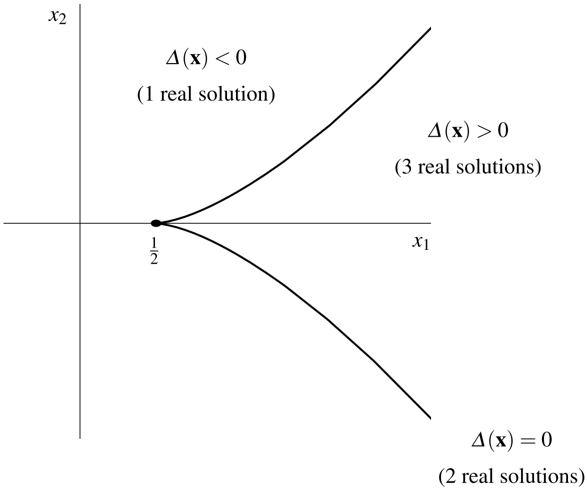


Fig. 12.4 Values of the discriminant $\Delta(\mathbf{x})$

Since $\nabla S = \mathbf{a}$, the induced Lagrangian immersion is given by

$$\Lambda_S(\mathbf{x}, \hat{\mathbf{x}}, \mathbf{a}) = (\mathbf{x}, \mathbf{a}) \equiv \left(\mathbf{x}, \pm\frac{\mathbf{x}-\hat{\mathbf{x}}}{\|\mathbf{x}-\hat{\mathbf{x}}\|}\right),$$

where the sign is positive if $\mathbf{x} - \hat{\mathbf{x}}$ and $\mathbf{a}$ are parallel and negative if they are antiparallel. This reproduces the Lagrangian immersions of the generating families $S(\mathbf{x}, \hat{\mathbf{x}}) = \pm\|\mathbf{x} - \hat{\mathbf{x}}\|$, because the latter are obtained from (12.7.32) by restriction to

$$\mathbf{a} = \pm\frac{\mathbf{x}-\hat{\mathbf{x}}}{\|\mathbf{x}-\hat{\mathbf{x}}\|}.$$

Let us add that if D is given as the zero level set of a smooth function $F : Q \to \mathbb{R}$, the Morse family (12.7.32) is equivalent to the function

$$S(\mathbf{x}, \hat{\mathbf{x}}, \mathbf{a}, \lambda) = (\mathbf{x} - \hat{\mathbf{x}}) \cdot \mathbf{a} + \lambda F(\hat{\mathbf{x}}) \qquad (12.7.33)$$

(Exercise 12.7.7).

To illustrate the solution method provided by Proposition 12.7.6, we take up Example 12.7.5.

Example 12.7.8 Let $Q = \mathbb{R}^2$ and $D = \{\mathbf{x} \in \mathbb{R}^2 : (x_2)^2 = x_1\}$ and let us choose $y = x_2$ as a coordinate on D. Then,

$$\hat{\mathbf{x}}(y) = (y^2, y), \qquad \mathbf{e}(y) = (2y, 1).$$

The Morse family of Proposition 12.7.6 is given by

$$S(\mathbf{x}, y) = \|\mathbf{x} - \hat{\mathbf{x}}(y)\| = \sqrt{(x_1 - y^2)^2 + (x_2 - y)^2}, \qquad (12.7.34)$$

where we have identified $B = \{(\mathbf{x}, y) \in \mathbb{R}^3 : \mathbf{x} \neq (y^2, y)\}$. According to (12.7.24), the defining equation for the fibre-critical submanifold B_S reads

$$y^3 - y\left(x_1 - \frac{1}{2}\right) - \frac{1}{2}x_2 = 0 \tag{12.7.35}$$

and the induced Lagrangian immersion is given by

$$\Lambda_S(\mathbf{x}, y) = \left(x_1, x_2, \frac{x_1 - y^2}{\sqrt{(x_1 - y^2)^2 + (x_2 - y)^2}}, \frac{x_2 - y}{\sqrt{(x_1 - y^2)^2 + (x_2 - y)^2}}\right).$$

Equation (12.7.35) is solved by

$$x_1 = y^2 + \frac{t}{\sqrt{1 + 4y^2}}, \qquad x_2 = y - \frac{2yt}{\sqrt{1 + 4y^2}} \tag{12.7.36}$$

with $(y, t) \in \mathbb{R}^2$ such that $t \neq 0$. This yields a global parameterization of B_S:

$$B_S = \left\{\left(y^2 + \frac{t}{\sqrt{1 + 4y^2}}, y - \frac{2yt}{\sqrt{1 + 4y^2}}, y\right) : (y, t) \in \mathbb{R}^2, t \neq 0\right\}. \tag{12.7.37}$$

In terms of this parameterization, Λ_S is given by

$$\Lambda_S(y,t) = \left(y^2 + \frac{t}{\sqrt{1 + 4y^2}}, y - \frac{2yt}{\sqrt{1 + 4y^2}}, \frac{\operatorname{sign}(t)}{\sqrt{1 + 4y^2}}, \frac{-2y \operatorname{sign}(t)}{\sqrt{1 + 4y^2}}\right). \tag{12.7.38}$$

It is easy to see that Λ_S is injective, so that (B_S, Λ_S) is a Lagrangian submanifold and hence a geometric solution whose image is $\Lambda_S(B_S)$. Next, we discuss the singular subset $\Sigma(B_S)$ and the caustic $\Gamma(B_S)$. According to (12.7.29), the defining equation for $\Sigma(B_S)$ is

$$1 + 4y^2 - \frac{2t}{\sqrt{1 + 4y^2}} = 0.$$

In particular, on $\Sigma(B_S)$ we have $t > 0$. Solving for t and plugging this into (12.7.37), we obtain

$$\Sigma(B_S) = \left\{\left(\frac{1}{2} + 3y^2, -4y^3, y\right) \in B_S : y \in \mathbb{R}\right\},$$

$$\Gamma(B_S) = \left\{\left(\frac{1}{2} + 3y^2, -4y^3\right) \in Q : y \in \mathbb{R}\right\}.$$

We read off that $\Sigma(B_S)$ is a submanifold of B_S of dimension one, whereas $\Gamma(B_S)$ is a subset containing a singular point, see Fig. 12.5. Finally, we count the number of intersection points of (B_S, Λ_S) with the fibre of T^*Q over a point $\mathbf{x} \in Q$. This can be determined from (12.7.35) by viewing the left hand side as a polynomial in

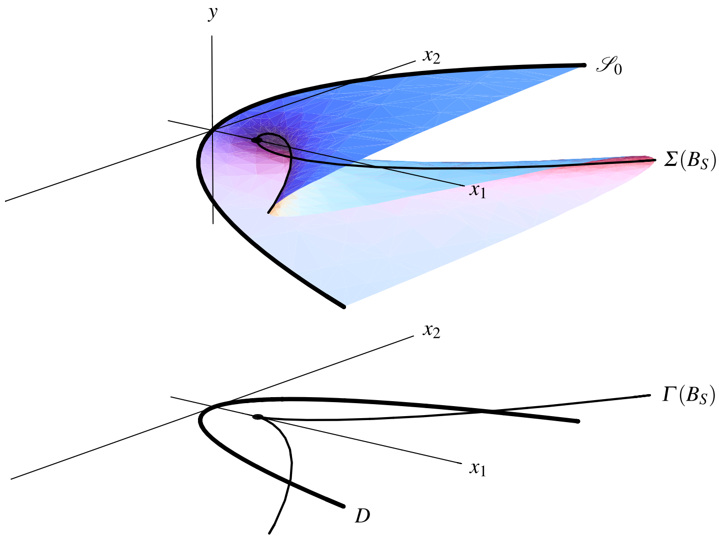


Fig. 12.5 The Lagrangian submanifold (B_S, Λ_S) of Example 12.7.8 for $t > 0$. Note that $\mathscr{S}_0$ does not belong to (B_S, Λ_S)

the variable y and counting the real zeros. Their number can be read off from the values of the discriminant

$$\Delta(\mathbf{x}) = 4\left(x_1 - \frac{1}{2}\right)^3 - \frac{27}{4}x_2^2,$$

depicted in Fig. 12.4. For $\Delta(\mathbf{x}) > 0$ this polynomial has three real solutions and for $\Delta(\mathbf{x}) < 0$ it has one real solution and this solution has multiplicity one. For $\Delta(\mathbf{x}) = 0$, the polynomial has two real solutions, one of them with multiplicity two, or one real solution with multiplicity three. By expressing $\Gamma(B_S)$ in terms of $\mathbf{x}$, we see that the zero set of $\Delta(\mathbf{x})$ coincides with the caustic and that the point $\mathbf{x}$ where the polynomial has one real solution with multiplicity three coincides with the singular point. Correspondingly, the fibre over $\mathbf{x}$ intersects $\Lambda_S(B_S)$ three times transversally for $\Delta(\mathbf{x}) > 0$, once transversally for $\Delta(\mathbf{x}) > 0$, once non-transversally over the singular point, and once transversally and once non-transversally over the remaining points of the caustic, see Fig. 12.5. It is clear that both for $\Delta(\mathbf{x}) < 0$ and for $\Delta(\mathbf{x}) > 0$ one can obtain analytic solutions from the geometric solution (B_S, Λ_S) by solving (12.7.35) for y.

Remark 12.7.9

1. The fact that, here, the singular subset $\Sigma(B_S)$ is in fact a submanifold is no accident. By Proposition 12.6.3, the complement $\Sigma(B_S) \setminus \Sigma_1(B_S)$ consists of submanifolds of codimension greater than 2. Hence, for dimensional reasons, $\Sigma(B_S) = \Sigma_1(B_S)$. From the point of view of catastrophe theory, the point

$(\frac{1}{2}, 0, 0) \in \Pi^{-1}((\frac{1}{2}, 0))$ is distinguished. The reader easily checks that this point is the only one where the third partial derivative S'''_{yyy} vanishes. This means that apart from this point one can find a normal form of fold type for S, cf. Proposition 12.6.15. If one includes this point, the normal form is of the cusp type

$$S(\mathbf{x}, y) = y^4 + x_1 y^2 + x_2 y,$$

which in the Thom-Boardman classification scheme is labelled by $\Sigma_{1,1,0}$, cf. Remark 12.6.16.

2. The parameter t in the parameterization (12.7.37) of B_S is closely related to the flow parameter of the generalized solution $\iota : \mathscr{S} \to \mathrm{T}^*Q$ obtained by the method of characteristics in Example 12.7.5. Comparison of the immersions (12.7.21) and (12.7.38) shows that for $t > 0$, we have $\Lambda_S(y, t) = \iota^+(y, t)$, whereas for $t < 0$, we have $\Lambda_S(y, t) = \iota^-(y, -t)$. Hence, in the dynamical interpretation of the method of characteristics as a Hamiltonian flow with the parameter t representing time, the solution given by (B_S, Λ_S) corresponds to integral curves emanating from $\mathscr{S}_0^+$ and from $\mathscr{S}_0^-$, in both cases evolving for time $t > 0$. In particular, the parameter t in (12.7.38) can be interpreted as the time in case $t > 0$ and as the negative of the time in case $t < 0$. This makes transparent what is meant by saying that the Morse family (12.7.23) models outgoing light rays. We leave it to the reader to carry out the analogous analysis for the Morse family $S(\mathbf{x}, \hat{\mathbf{x}}) = -\|\mathbf{x} - \hat{\mathbf{x}}\|$. As a result, the Lagrangian immersions of both families together make up the generalized solution $(\mathscr{S}, \iota)$ obtained by the method of characteristics.

The simpler example 12.7.4 can be analyzed in the same spirit. We leave this as an exercise to the reader (Exercise 12.7.8). Instead, we now turn to the case where the initial phase S_0 is arbitrary but the initial submanifold D has codimension one in Q. As before, we will first discuss the method of characteristics. Under the assumption of codimension one, we can make the ansatz that the initial condition $\mathscr{S}_0$ is the image of D under a smooth mapping $\alpha : D \to \mathrm{T}^*Q$. It is evident that this mapping must satisfy the conditions

1. $\pi \circ \alpha = \mathrm{id}_D$,
2. $\langle \alpha(\mathbf{x}), X \rangle = \langle \mathrm{d}S_0, X \rangle$ for all $\mathbf{x} \in D$, $X \in \mathrm{T}_{\mathbf{x}}D$,
3. $H \circ \alpha = 0$,
4. α is transversal to the integral curves of X_H.

Example 12.7.10 Let $Q = \mathbb{R}^2$ and $D = \{\mathbf{x} \in \mathbb{R}^2 : x_1 = 0\}$ and let $S_0 : D \to \mathbb{R}$ be a smooth function. We use $y = x_2$ as a coordinate on D. To find the mapping $\alpha : D \to \mathrm{T}^*Q$, we make the ansatz

$$\alpha(y) = \big(0, y, g(y), S_0'(y)\big),$$

which fulfils conditions 1 and 2 by construction. Condition 3 yields

$$S_0'(y)^2 + g(y)^2 = 1$$

for all $y \in \mathbb{R}$. Therefore, S_0 must satisfy $|S_0'(y)| \leq 1$ for all y. Under this assumption, $g(y) = \pm\sqrt{1-(S_0'(y))^2}$. Finally, to meet condition 4, we must require $g(y) \neq 0$ and hence $|S_0'(y)| < 1$ for all y. Thus, $\mathscr{S}_0$ consists of the connected components

$$\mathscr{S}_0^{\pm} = \left\{\left(0, y, \pm\sqrt{1-\left(S_0'(y)\right)^2}, S_0'(y)\right)\right\}, \tag{12.7.39}$$

$\mathscr{S}$ consists of the connected components $\mathscr{S}^{\pm} = \mathscr{S}_0^{\pm} \times \mathbb{R}$ and ι is given by

$$\iota^{\pm}(y,t) = \left(\pm 2t\sqrt{1-S_0'(y)^2}, y + 2S_0'(y)t, \pm\sqrt{1-S_0'(y)^2}, S_0'(y)\right). \tag{12.7.40}$$

The defining equation $\det \frac{\partial(\Pi_1,\Pi_2)}{\partial(y,t)} = 0$ for the singular subset $\Sigma(\mathscr{S})$ is equivalent to

$$1 + 2S_0''(y)t - S_0'^2(y) = 0. \tag{12.7.41}$$

We discuss two special cases in detail.

1. Let $S_0(y) = ay$ with $|a| < 1$. Here,

$$\iota^{\pm}(y,t) = \left(\pm 2\sqrt{1-a^2}t, y + 2at, \pm\sqrt{1-a^2}, a\right).$$

   The signs correspond to light travelling in the positive or negative x_1-direction. The images $\iota^{\pm}(\mathscr{S}^{\pm})$ are 2-dimensional hyperplanes in $\mathrm{T}^*Q = \mathbb{R}^4$. Since $S_0'' = 0$, Eq. (12.7.41) for the singular subset implies $S_0'^2(y) = 1$, which contradicts the transversality requirement $|S_0'(y)| < 1$. Hence, in this case there is no caustic.
2. Let $S_0''(y) \neq 0$ for all y and restrict attention to $\mathscr{S}^+$. In this case, the caustic is given by

$$\Gamma(\mathscr{S}) = \left\{\left(-\frac{(1-S_0'^2)^{\frac{3}{2}}}{S_0''}, y - \frac{S_0'(1-S_0'^2)}{S_0''}\right) \in Q : y \in \mathbb{R}\right\}. \tag{12.7.42}$$

   We show that one can choose S_0 so that the caustic degenerates to a single focal point on the x_1-axis,

$$\Gamma(\mathscr{S}) = \left\{(f,0)\right\}$$

   for some $f \in \mathbb{R}$. According to (12.7.42), this leads to the differential equation

$$\frac{y}{f} = -\frac{S_0'}{\sqrt{1-S_0'^2}},$$

   which up to an irrelevant additive constant has the solution

$$S_0(y) = \pm\sqrt{f^2 + y^2}.$$

For the positive sign, f must be negative. In this case, S_0 models an ideal concave lense. For the negative sign, f must be positive and S_0 models an ideal convex lense.

Next, we determine the eikonal function S. From (12.7.40) we read off that in terms of the coordinates y, t, the differential $\mathrm{d}S$ is given by

$$\begin{aligned} \mathrm{d}S(y,t) &= p_1(y,t)\mathrm{d}x_1(y,t) + p_2(y,t)\mathrm{d}x_2(y,t) \\ &= \sqrt{1 - S_0'(y)^2}\,\mathrm{d}\Big(2t\sqrt{1 - S_0'(y)^2}\Big) + S_0'(y)\mathrm{d}\big(y + 2S_0'(y)t\big) \\ &= 2\mathrm{d}t + S_0'(y)\mathrm{d}y. \end{aligned}$$

The initial condition $S_{\restriction D} = S_0$ yields $S(y,0) = S_0(y)$ and hence

$$S(y,t) = 2t + S_0(y). \tag{12.7.43}$$

From this, we can obtain S as a function on Q by solving the equations

$$x_1 = \pm 2\sqrt{1 - \big(S_0'(y)\big)^2}\,t, \qquad x_2 = y + 2S_0'(y)t$$

for y and t and plugging in the solutions into (12.7.43). For the two specific initial condition from above, this yields

$$S(\mathbf{x}) = ax_2 \pm \sqrt{1-a^2}\,x_1$$

for the linearly increasing initial phase $S_0(y) = ay$ and

$$S(\mathbf{x}) = \sqrt{(x_1 + f)^2 + x_2^2},$$

for the model of an ideal concave lense $S_0(y) = \sqrt{f^2 + y^2}$. In the first case, the wave fronts are planes and in the second example they are spheres centred at the focal point.

To conclude this section, we show that in the situation where D has codimension one, the generalized solution can be generated by a Morse family in much the same way as in the case of a constant initial phase $S_0 = 0$. As before, we denote the elements of D by $\hat{\mathbf{x}}$, choose coordinates y^α on D and define the vectors $\mathbf{e}_\alpha(\hat{\mathbf{x}})$ by (12.7.22).

Proposition 12.7.11 *Let D be an embedded submanifold of Q of codimension one and let S_0 be a smooth function on D. Assume that*

$$\frac{\partial S_0}{\partial y^\alpha}\frac{\partial S_0}{\partial y^\beta}\mathsf{h}^{\alpha\beta} < 1, \tag{12.7.44}$$

where $\mathsf{h}_{\alpha\beta} = \mathbf{e}_\alpha \cdot \mathbf{e}_\beta$ *is the induced metric on* D.[23] *Let* $B = \{(\mathbf{x}, \hat{\mathbf{x}}) \in Q \times D : \mathbf{x} \neq \hat{\mathbf{x}}\}$. *Then*,

$$S : B \to \mathbb{R}, \qquad S(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\| + S_0(\hat{\mathbf{x}}), \tag{12.7.45}$$

is a Morse family along the submersion $\pi : B \to Q$ *induced from the natural projection. This Morse family generates a generalized solution of the eikonal equation* (12.7.14) *for the initial condition* S_0 *on* D, *given by*

$$B_S = \left\{(\mathbf{x}, \hat{\mathbf{x}}) \in B : \frac{\mathbf{x} - \hat{\mathbf{x}}}{\|\mathbf{x} - \hat{\mathbf{x}}\|} \cdot \mathbf{e}_\alpha(\hat{\mathbf{x}}) = \frac{\partial S_0}{\partial y^\alpha}(\hat{\mathbf{x}})\right\}, \qquad \Lambda_S(\mathbf{x}, \hat{\mathbf{x}}) = \left(\mathbf{x}, \frac{\mathbf{x} - \hat{\mathbf{x}}}{\|\mathbf{x} - \hat{\mathbf{x}}\|}\right).$$

Proof We proceed as in the proof of Proposition 12.7.6. The relevant partial derivatives are

$$\frac{\partial S}{\partial x^i} = \frac{x_i - \hat{x}_i}{\|\mathbf{x} - \hat{\mathbf{x}}\|} =: p_i, \qquad \frac{\partial S}{\partial y^\alpha} = -\mathbf{p} \cdot \mathbf{e}_\alpha + \frac{\partial S_0}{\partial y^\alpha}, \qquad \frac{\partial^2 S}{\partial x^i \partial y^\alpha} = \frac{(\mathbf{p} \cdot \mathbf{e}_\alpha) p_i - \mathbf{e}_{\alpha i}}{\|\mathbf{x} - \hat{\mathbf{x}}\|}.$$

This yields the asserted formula for B_S. Moreover, it follows that the left hand side of (12.7.44), taken at $\hat{\mathbf{x}}$, coincides with the absolute square of the projection of $\mathbf{p}(\mathbf{x}, \hat{\mathbf{x}})$ to $\mathrm{T}_{\hat{\mathbf{x}}} D$. Since $\mathbf{p}(\mathbf{x}, \hat{\mathbf{x}})^2 = 1$, this implies that on B_S, we have $\mathbf{p}(\mathbf{x}, \hat{\mathbf{x}}) \notin \mathrm{T}_{\hat{\mathbf{x}}} D$. It follows that the vector fields $\mathbf{e}_\alpha - (\mathbf{p} \cdot \mathbf{e}_\alpha)\mathbf{p}$ are pointwise linearly independent on B_S: assume

$$\sum_\alpha \lambda_\alpha \big(\mathbf{e}_\alpha - (\mathbf{p} \cdot \mathbf{e}_\alpha)\mathbf{p}\big) = \left(\sum_\alpha \lambda_\alpha \mathbf{e}_\alpha\right) - \mathbf{p} \cdot \left(\sum_\alpha \lambda_\alpha \mathbf{e}_\alpha\right)\mathbf{p} = 0.$$

Since $\sum_\alpha \lambda_\alpha \mathbf{e}_\alpha$ is tangent to D but $\mathbf{p}$ is not, this equation is only fulfilled if all λ_α vanish. We conclude that the $n \times (n-1)$-matrix S''_{xy} has rank $n - 1$ on B_S and, therefore, that S is a Morse family, indeed. That Λ_S takes values in $\mathscr{C}$ is obvious. □

Remark 12.7.12 The discussion of alternative Morse families in Remark 12.7.7/3 carries over to the present case. In particular, the function

$$S(\mathbf{x}, \hat{\mathbf{x}}) = -\|\mathbf{x} - \hat{\mathbf{x}}\| + S_0(\hat{\mathbf{x}})$$

is a generating Morse family, too, and to (12.7.32) and (12.7.33) there correspond the extensions

$$S(\mathbf{x}, \hat{\mathbf{x}}, \mathbf{a}) = (\mathbf{x} - \hat{\mathbf{x}}) \cdot \mathbf{a} + S_0(\hat{\mathbf{x}}), \qquad S(\mathbf{x}, \hat{\mathbf{x}}, \mathbf{a}, \lambda) = (\mathbf{x} - \hat{\mathbf{x}}) \cdot \mathbf{a} + \lambda F(\hat{\mathbf{x}}) + S_0(\hat{\mathbf{x}}),$$

respectively.

We refer to the book of Benenti [38] for a lot of additional material. There, the reader can find a systematic treatment of optical systems including sources, mirrors

[23] This is the abstract counterpart of the condition $|S_0'(y)| < 1$ found in Example 12.7.10.

and lenses in the language of Morse families. For instance, for a system consisting of a source U and a mirror V, the generating family is

$$S(\mathbf{x}, \mathbf{u}, \mathbf{v}) = \|\mathbf{x} - \mathbf{v}\| + \|\mathbf{v} - \mathbf{u}\|,$$

with $\mathbf{u} \in U$ and $\mathbf{v} \in V$. Benenti uses the language of symplectic relations, which is well adapted to this sort of problems. To every component of the system there corresponds a symplectic relation and the description of the system as a whole is obtained by taking the composition of these relations. To every relation, there corresponds a generating Morse family and the Morse family of the full system is the sum of them.

Exercises

12.7.1 Work out Examples 12.7.4 and 12.7.5.

12.7.2 Analyze the coorientation of the singular subset $\Sigma(B_S)$ in Example 12.7.8 by studying the sign of S''_{yy} along a curve transversal to $\Sigma(B_S)$.

12.7.3 Show that, in the situation of Proposition 12.7.6, the function

$$S : Q \times D \to \mathbb{R}, \qquad S(\mathbf{x}, \hat{\mathbf{x}}) = (\mathbf{x} - \hat{\mathbf{x}})^2,$$

is a Morse family. Determine the induced Lagrangian immersion, its singular subset and the caustic. Is the induced Lagrangian immersion a generalized (geometric) solution of the eikonal equation?

12.7.4 Show that the function

$$W : Q \times \mathrm{S}^{n-1} \to \mathbb{R}, \qquad W(\mathbf{x}, \mathbf{a}) = \mathbf{a} \cdot \mathbf{x}$$

is a Morse family generating a generalized solution of the eikonal equation (12.7.14). What initial condition does this function describe?

12.7.5 Let $\mathbf{x}_0 \in Q$. Show that the function

$$S\colon Q \times \mathrm{S}^{n-1} \to \mathbb{R}, \qquad (\mathbf{x}, \mathbf{a}) \mapsto (\mathbf{x} - \mathbf{x}_0) \cdot \mathbf{a}$$

is a Morse family generating a generalized solution of the eikonal equation (12.7.14) for the point source $D = \{\mathbf{x}_0\}$. What is the difference between S and the Morse family (12.7.23) in this case?

12.7.6 For the case $Q = \mathbb{R}^3$ and D having codimension 1, verify explicitly that Eq. (12.7.29) for the singular subset $\Sigma(B_S)$ of the Morse family (12.7.23) is equivalent to $\det \Pi' = 0$ by showing that for $(\mathbf{x}, \hat{\mathbf{x}}) \in B_S$,

$$\det \Pi'(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\| \det\big(S''_{yy}(\mathbf{x}, \hat{\mathbf{x}})\big). \tag{12.7.46}$$

Hint. Parameterize B_S by $\hat{\mathbf{x}} \in D$ and $t \in \mathbb{R} \setminus \{0\}$ as follows:

$$\mathbf{x}(\hat{\mathbf{x}}, t) = \hat{\mathbf{x}} + t\frac{\mathbf{e}_1(\hat{\mathbf{x}}) \times \mathbf{e}_2(\hat{\mathbf{x}})}{\|\mathbf{e}_1(\hat{\mathbf{x}}) \times \mathbf{e}_2(\hat{\mathbf{x}})\|}.$$

12.7.7 Show that (12.7.32) is a Morse family and that this Morse family is equivalent to the one defined by (12.7.33).

12.7.8 Work out the construction of a generalized solution of the eikonal equation for Example 12.7.4 by the Morse family method of Proposition 12.7.6. Show that the analytic solution (apart from the caustic) is given by $S(\mathbf{x}) = |\|\mathbf{x}\| - R|$.
Hint. Parameterize B_S by y and the time parameter of the solution obtained by the method of characteristics.

12.8 Geometric Asymptotics. Beyond Lowest Order

In this section we continue the discussion of the short wave asymptotics of partial differential equations of the type

$$H\left(\mathbf{x}, -\frac{\mathrm{i}}{k}\nabla\right)u(\mathbf{x}) = 0 \tag{12.8.1}$$

on $Q = \mathbb{R}^n$, defined by a Hamiltonian function H on T^*Q which restricts to a polynomial on each fibre. In the previous section we have discussed the characteristic equation

$$H\big(\mathbf{x}, \nabla S(\mathbf{x})\big) = 0. \tag{12.8.2}$$

Here, we study the corresponding transport equation

$$\left(\frac{\partial H}{\partial p_i}\big(\mathbf{x}, \nabla S(\mathbf{x})\big)\frac{\partial}{\partial x^i} + \frac{1}{2}\frac{\partial^2 H}{\partial p_i \partial p_j}\big(\mathbf{x}, \nabla S(\mathbf{x})\big)\frac{\partial^2 S}{\partial x^i \partial x^j}(\mathbf{x})\right)a_0(\mathbf{x}) = 0 \tag{12.8.3}$$

and construct first order asymptotic solutions of (12.8.1) from solutions of (12.8.2) and (12.8.3). The leading example will be the Helmholtz equation (12.7.2), with its transport equation given by (12.7.7).[24] Let $\mathscr{S}_0$ be an admissible initial condition for the characteristic equation, given by the canonical lift of a submanifold D of Q of codimension 1 with an initial phase S_0 on D. Let $\alpha : D \to \mathrm{T}^*Q$ be the corresponding lifting mapping. As explained in Sect. 12.7, by the method of characteristics, $\mathscr{S}_0$ generates a generalized solution $\iota : \mathscr{S} \to \mathrm{T}^*Q$ of the characteristic equation (12.8.2), where $\mathscr{S}$ coincides with the intersection of $\mathscr{S}_0 \times \mathbb{R}$ with the domain of the Hamiltonian vector field X_H and where ι is induced by the flow Φ of X_H,

$$\iota : \mathscr{S} \subset \mathscr{S}_0 \times \mathbb{R} \to \mathrm{T}^*Q, \qquad \iota(\xi_0, t) = \Phi_t(\xi_0). \tag{12.8.4}$$

For a moment, let us assume that we are away from the caustic of $(\mathscr{S}, \iota)$, and let us restrict $\mathscr{S}$ in such a way that it coincides with the image of $\mathrm{d}S$ for an analytic solution S. Then, $\Pi = \pi_Q \circ \iota$ is a diffeomorphism onto some open subset

[24] For a quite exhaustive discussion of this equation in theoretical optics, we refer to the book of Römer [253].

of Q and (12.8.3) can be rewritten as follows. Since X_H is tangent to $(\mathscr{S}, \iota)$, by Proposition 2.7.16, it restricts to a vector field $\tilde{X}_H$ on $\mathscr{S}$. Define

$$Y_H := \Pi_* \tilde{X}_H \tag{12.8.5}$$

and let ϕ denote the flow of Y_H. Obviously,

$$Y_H(\mathbf{x}) = \frac{\partial H}{\partial p_i}\big(\mathbf{x}, \nabla S(\mathbf{x})\big)\partial_{x^i}. \tag{12.8.6}$$

By (12.8.6), the first term in the transport equation (12.8.3) can be written in terms of Y_H as

$$\frac{\partial H}{\partial p_i}\big(\mathbf{x}, \nabla S(\mathbf{x})\big)\frac{\partial a_0}{\partial x^i}(\mathbf{x}) = (Y_H a_0)(\mathbf{x}) = \nabla a_0(\mathbf{x}) \cdot Y_H(\mathbf{x}).$$

By computing $(\nabla \cdot Y_H)(\mathbf{x})$ we find that the second term in (12.8.3) has the form

$$\frac{\partial^2 H}{\partial p_i \partial p_j}\big(\mathbf{x}, \nabla S(\mathbf{x})\big)\frac{\partial^2 S}{\partial x^i \partial x^j}(\mathbf{x}) = (\nabla \cdot Y_H)(\mathbf{x}) - \frac{\partial^2 H}{\partial x^i \partial p_i}\big(\mathbf{x}, \nabla S(\mathbf{x})\big). \tag{12.8.7}$$

As a result, the transport equation (12.8.3) can be rewritten as

$$(\nabla a_0)(\mathbf{x}) \cdot Y_H(\mathbf{x}) + \frac{1}{2}a_0(\mathbf{x})(\nabla \cdot Y_H)(\mathbf{x}) = \frac{1}{2}a_0(\mathbf{x})\frac{\partial^2 H}{\partial x^i \partial p_i}\big(\mathbf{x}, \nabla S(\mathbf{x})\big). \tag{12.8.8}$$

Next, we are going to interpret this equation geometrically in terms of half densities on $\mathscr{S}$. This will allow us to solve the transport equation globally. In the same spirit as for the characteristic equation, such a global solution will be referred to as a generalized solution of the transport equation. For simplicity, we limit our attention to the case where

$$\frac{\partial^2 H}{\partial x^i \partial p_i} = 0. \tag{12.8.9}$$

This includes the Helmholtz equation and, more generally, the case where the operator H in (12.8.1) is Hermitian. Under this assumption, the transport equation is equivalent to

$$\nabla \cdot \big(a_0^2 Y_H\big) = 0. \tag{12.8.10}$$

Let $\mathsf{v}_n = \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$ be the canonical volume form on $Q = \mathbb{R}^n$. Then, $\Pi^* \mathsf{v}_n$ defines a natural volume form on the submanifold $\Sigma_0(\mathscr{S})$, cf. Formula (12.6.1). By (4.1.24), for every vector field X and every smooth function f on Q, one has

$$\mathcal{L}_X \mathsf{v}_n = (\nabla \cdot X)\mathsf{v}_n, \qquad \mathrm{d}f \wedge (X \lrcorner \mathsf{v}_n) = X(f)\mathsf{v}_n. \tag{12.8.11}$$

Using these two relations and the identity (4.1.27), we find

$$\big(\nabla \cdot \big(a_0^2 Y_H\big)\big)\mathsf{v}_n = \mathcal{L}_{a_0^2 Y_H}\mathsf{v}_n = \mathcal{L}_{Y_H}\big(a_0^2 \mathsf{v}_n\big),$$

and by Proposition 3.3.3/5, we obtain

$$\mathcal{L}_{Y_H}\big(a_0^2 \mathsf{v}_n\big) = \big(\Pi^{-1}\big)^*\big(\mathcal{L}_{\tilde{X}_H}\Pi^*\big(a_0^2 \mathsf{v}_n\big)\big).$$

Thus, the transport equation (12.8.10) is equivalent to

$$\mathcal{L}_{\tilde{X}_H}\Pi^*\big(a_0^2 \mathsf{v}_n\big) = 0. \tag{12.8.12}$$

This is an equation for a_0^2, formulated in terms of the n-form $\Pi^*(a_0^2 \mathsf{v}_n)$ on $\mathscr{S}$. To obtain an equation for a_0, we have to pass to half-densities on $\mathscr{S}$. For an exhaustive discussion of this concept we refer to the literature, e.g. [119].[25]

Lemma 12.8.1 *We have*

$$\mathcal{L}_{\tilde{X}_H}\Pi^*\big(a_0|\mathsf{v}_n|^{\frac{1}{2}}\big) = \Pi^*\bigg(\nabla a_0 \cdot Y_H + \frac{1}{2}a_0 \nabla \cdot Y_H\bigg)\Pi^*|\mathsf{v}_n|^{\frac{1}{2}}. \tag{12.8.13}$$

Proof Using Proposition 3.3.3/5 and the first relation in (12.8.11), we find

$$\mathcal{L}_{\tilde{X}_H}\big(\Pi^* \mathsf{v}_n\big) = \Pi^*(\mathcal{L}_{Y_H}\mathsf{v}_n) = \Pi^*(\nabla \cdot Y_H)\Pi^* \mathsf{v}_n.$$

This formula carries over to the density $|\mathsf{v}_n|$. On the other hand, by the derivation property of the Lie derivative,

$$\mathcal{L}_{\tilde{X}_H}\big(\Pi^*|\mathsf{v}_n|\big) = 2\Pi^*|\mathsf{v}_n|^{\frac{1}{2}}\mathcal{L}_{\tilde{X}_H}\big(\Pi^*|\mathsf{v}_n|^{\frac{1}{2}}\big)$$

and thus

$$\mathcal{L}_{\tilde{X}_H}\big(\Pi^*|\mathsf{v}_n|^{\frac{1}{2}}\big) = \frac{1}{2}\Pi^*(\nabla \cdot Y_H)\Pi^*|\mathsf{v}_n|^{\frac{1}{2}}.$$

Now, the assertion follows by applying once again the derivation property of the Lie derivative. □

As a consequence of the lemma, the transport equation can be written in the form

$$\mathcal{L}_{\tilde{X}_H}\Pi^*\big(a_0|\mathsf{v}_n|^{\frac{1}{2}}\big) = 0. \tag{12.8.14}$$

---

[25] In brief, for $s \in \mathbb{R}$, an s-density on a real vector space W of dimension n is a mapping $\nu : W^n \to \mathbb{R}$ satisfying $\nu(Aw_1, \ldots, Aw_n) = |\det A|^s \nu(w_1, \ldots, w_n)$ for all endomorphisms A of W. The s-densities on W form a vector space of dimension 1. By taking the s-densities on the tangent spaces at every point of a manifold M one obtains the real line bundle $|\Lambda|^s M$ of pointwise s-densities on M. Sections in this bundle are called s-densities on M. Every n-form v on M defines an s-density $|\mathsf{v}|^s$ by $|\mathsf{v}|^s(X_1, \ldots, X_n) = |\mathsf{v}(X_1, \ldots, X_n)|^s$. In particular, if M is orientable, $|\Lambda|^s M$ is trivial for every s. The product of an s_1-density ν_1 and an s_2-density ν_2 on M is defined by $(\nu_1\nu_2)(X_1, \ldots, X_n) := \nu_1(X_1, \ldots, X_n)\nu_2(X_1, \ldots, X_n)$. It yields an $(s_1 + s_2)$-density. The calculus of differential forms, notably the pull-back and, based on that, the Lie derivative, extends in an obvious way to s-densities.

In this form, it generalizes to arbitrary half-densities $\hat{a}_0$ on $\mathscr{S}$:

$$\mathcal{L}_{\tilde{X}_H} \hat{a}_0 = 0. \tag{12.8.15}$$

The solutions of this equation will be referred to as generalized solutions of the transport equation (12.8.3). They can be obtained by means of the method of characteristics as follows. Let $a_0^{(0)}$ be a given initial condition for a_0 on D. In terms of the representation of $(\mathscr{S}, \iota)$ by (12.8.4), the solution of (12.8.15) generated by $a_0^{(0)}$ is given by

$$\hat{a}_0(\xi_0, t) = a_0^{(0)}\big(\Pi(\xi_0)\big)\Big(\Phi_{-t}^* \circ \Pi^*\big(|\mathsf{v}_n|^{\frac{1}{2}}\big)\Big)(\xi_0, t), \tag{12.8.16}$$

where $\xi_0 \in \mathscr{S}_0$ and Φ denotes the flow of $\tilde{X}_H$. By restricting a generalized solution $\hat{a}_0$ to an open subset of $\mathscr{S}$ which under Π is mapped diffeomorphically to an open subset of Q, one can construct an analytic solution a_0 of (12.8.3) on the latter via the relation

$$\hat{a}_0 = \Pi^*\big(a_0 |\mathsf{v}_n|^{\frac{1}{2}}\big). \tag{12.8.17}$$

By 3.2.13/2, the flows Φ of $\tilde{X}_H$ and ϕ of Y_H are related by

$$\Pi \circ \Phi_t = \phi_t \circ \Pi. \tag{12.8.18}$$

Using this, from (12.8.16) we obtain

$$\hat{a}_0\big(\alpha(\mathbf{x}_0), t\big) = a_0^{(0)}(\mathbf{x}_0)\big(\big|\det(\phi_t)'_{\mathbf{x}_0}\big|\big)^{-\frac{1}{2}}\big(\Pi^*|\mathsf{v}_n|^{\frac{1}{2}}\big)\big(\alpha(\mathbf{x}_0), t\big), \tag{12.8.19}$$

where $\mathbf{x}_0 \in D$ and $\alpha : D \to \mathscr{S}_0 \subset \mathrm{T}^*Q$ denotes the mapping which yields the canonical lift of the pair (D, S_0). The partial derivatives in $(\phi_t)'_{\mathbf{x}_0}$ are taken with respect to the standard coordinates on $Q = \mathbb{R}^n$ (Exercise 12.8.2). Thus, denoting

$$\mathscr{J}(t, \mathbf{x}) := \det(\phi_t)'_{\mathbf{x}},$$

we read off that the analytic solution induced by $\hat{a}_0$ is

$$a_0(\mathbf{x}) = \frac{a_0^{(0)}(\mathbf{x}_0)}{\sqrt{|\mathscr{J}(t, \mathbf{x}_0)|}}, \tag{12.8.20}$$

where $t \in \mathbb{R}$ and $\mathbf{x}_0 \in D$ are determined by $\mathbf{x}$ through the relation $\mathbf{x} = \phi_t(\mathbf{x}_0)$. Note that this solution is limited to a region around D where t and $\mathbf{x}_0$ are uniquely determined by this relation. From (12.8.20) we obtain a short wave asymptotic solution of (12.8.1) up to first order:

$$u(\mathbf{x}) = \frac{a_0^{(0)}(\mathbf{x}_0)}{\sqrt{|\mathscr{J}(t, \mathbf{x}_0)|}} \mathrm{e}^{\mathrm{i}kS(\mathbf{x})}. \tag{12.8.21}$$

Remark 12.8.2

1. In the general case, the local solution of Eq. (12.8.8) is given by

$$a_0(\mathbf{x}) = \frac{1}{\sqrt{|\mathscr{J}(t,\mathbf{x}_0)|}} a_0^{(0)}(\mathbf{x}_0) \exp\left\{\frac{1}{2}\int_0^t \mathrm{d}t' \frac{\partial^2 H}{\partial x^i \partial p_i}\big(\phi_{t'}(\mathbf{x}_0), \nabla S\big(\phi_{t'}(\mathbf{x}_0)\big)\big)\right\},$$

   see Exercise 12.8.1.
2. We study the behaviour of the local solution (12.8.20) near the caustic. By (12.8.18) and $\Pi \circ \alpha = \mathrm{id}_D$, we have

$$(\phi_t)'_{\mathbf{x}_0} = \Pi'_{\Phi_t(\alpha(\mathbf{x}_0))} \circ (\Phi_t)'_{\alpha(\mathbf{x}_0)} \circ \alpha'_{\mathbf{x}_0}. \tag{12.8.22}$$

   The right hand side makes sense for all t such that $(t, \alpha(\mathbf{x}_0))$ is in the domain of Φ and we may use it to extend the mappings $t \mapsto (\phi_t)'_{\mathbf{x}_0}$ and $t \mapsto \mathscr{J}(t,\mathbf{x}_0)$ accordingly. For the extension, it may happen that $\phi_t(\mathbf{x}_0)$ belongs to the caustic $\Gamma(\mathscr{S})$. Since, then, $\Phi_t(\alpha(\mathbf{x}_0))$ belongs to the singular subset $\Sigma(\mathscr{S})$, the mappings $\Pi'_{\Phi_t(\alpha(\mathbf{x}_0))}$ and hence $(\phi_t)'_{\mathbf{x}_0}$ are not bijective, so that $\mathscr{J}(t,\mathbf{x}_0) = 0$. Thus, if $\phi_t(\mathbf{x}_0)$ approaches the caustic with t running and $\mathbf{x}_0$ fixed, $\mathscr{J}(t,\mathbf{x}_0)$ necessarily tends to zero and hence the solution (12.8.20) diverges.
3. In local coordinates ξ^i on $\mathscr{S}$, we have

$$\Pi^*|\mathsf{v}_n|^{\frac{1}{2}} = \sqrt{|\det \Pi'|}\,\big|\mathrm{d}\xi^1 \wedge \cdots \wedge \mathrm{d}\xi^n\big|^{\frac{1}{2}}. \tag{12.8.23}$$

   By (12.8.22), for every $\mathbf{x}_0 \in D$,

$$\big|\mathscr{J}(t,\mathbf{x}_0)\big|^{-\frac{1}{2}}\big|\det\big(\Pi'_{\Phi_t(\alpha(\mathbf{x}_0))}\big)\big|^{\frac{1}{2}} = \big|\det\big((\Phi_t)'_{\alpha(\mathbf{x}_0)}\big)\big|^{-\frac{1}{2}}\big|\det\big(\Pi'_{\alpha(\mathbf{x}_0)}\big)\big|^{\frac{1}{2}}.$$

   Thus, by (12.8.23), the solution (12.8.17) takes the form

$$\Pi^*\big(a_0|\mathsf{v}_n|^{\frac{1}{2}}\big)(\xi) = a_0^{(0)}(\mathbf{x}_0)\big|\det\big((\Phi_t)'_{\alpha(\mathbf{x}_0)}\big)\big|^{-\frac{1}{2}}\big|\det\big(\Pi'_{\alpha(\mathbf{x}_0)}\big)\big|^{\frac{1}{2}}|\mathrm{d}\xi^1 \wedge \cdots \wedge \mathrm{d}\xi^n|^{\frac{1}{2}}, \tag{12.8.24}$$

   where $\xi = \Phi_t(\alpha(\mathbf{x}_0))$. Since $\det((\Phi_t)'_{\alpha(\mathbf{x}_0)})$ and $\det(\Pi'_{\alpha(\mathbf{x}_0)})$ are regular, the right hand side of this formula makes sense for all t. Therefore, for given initial data, the generalized solution $\hat{a}_0$ can also be constructed explicitly as a continuation of an analytic solution a_0 through the caustic as follows. Choose a covering of $\mathscr{S}$ by local coordinates ξ^i and a subordinate partition of unity and use (12.8.24) to successively propagate $\Pi^*(a_0|\mathsf{v}_n|^{\frac{1}{2}})$ to all of $\mathscr{S}$.

Now, given the data $(\mathscr{S}, \iota)$ and $\hat{a}_0$, we will construct a global solution on Q of the differential equation (12.8.1) up to first order in $\frac{1}{k}$. For that purpose, let $\{U_i\}$ be a covering of $\mathscr{S}$ by contractible subsets such that all intersections $U_i \cap U_j$ are contractible and let $\{(B_i, \pi_i, S_i)\}$ be a system of generating Morse families such that the immersions $\Lambda_{S_i} : B_{S_i} \to \mathrm{T}^*Q$ and $\iota_{\restriction U_i} : U_i \to \mathrm{T}^*Q$ are equivalent. We may assume that these families are reduced and that the B_i are open subsets of $Q \times \mathbb{R}^{r_i}$ with π_i being induced from the natural projection. In what follows, for simplicity, we identify B_{S_i} with U_i. Now, let $\mathbf{x}_0 \in Q$ be an arbitrary regular point. Then, for every $\xi_0 \in \Pi^{-1}(\mathbf{x}_0)$, we have a solution of (12.8.1) up to first order in $\frac{1}{k}$ in some neighbourhood W of $\mathbf{x}_0$,

$$u(\mathbf{x}) = a_0(\mathbf{x})\mathrm{e}^{\mathrm{i}kS(\mathbf{x})},$$

where a_0 is determined by (12.8.17) and S is an analytic solution of the characteristic equation, locally generating $\mathscr{S}$. On W, the global solution we are looking for should be given by a sum of contributions of this type. However, if we want to put together such terms we face the problem that these contributions are only determined up to a constant phase.[26] To fix the relative phases we proceed as follows.

(a) To fix the relative phases of contributions coming from a given element of the covering, we will use the method of stationary phase.
(b) To combine the contributions of different elements U_i and U_j we must require that these contributions coincide if they come from the intersection $U_i \cap U_j$. This leads to an additional topological condition on $(\mathscr{S}, \iota)$ and k, known as the Bohr-Sommerfeld quantization condition.

To accomplish step (a), let us choose an element U of the above covering and let (B, π, S) be the corresponding Morse family. First, we show that, given the canonical[27] volume forms v_r and v_n on $\mathbb{R}^r$ and $Q = \mathbb{R}^n$ respectively, S induces a natural volume form v_{B_S} on the fibre-critical submanifold B_S. Let $\rho : B \to \mathbb{R}^r$ be the restriction of the canonical projection.

Lemma 12.8.3 *There exists a unique volume form v_{B_S} on B_S such that*

$$\pi^*\mathsf{v}_n \wedge \rho^*\mathsf{v}_r = \mathsf{v}_{B_S} \wedge \left(S'_y\right)^*\mathsf{v}_r \tag{12.8.25}$$

on $B_S \subset B$. This volume form satisfies

$$\det\left(S''_{yy}\right)\mathsf{v}_{B_S} = \Pi^*\mathsf{v}_n. \tag{12.8.26}$$

Proof Denote $F := S'_y : B \to \mathbb{R}^r$. Let $(\mathbf{x}, \mathbf{y}) \in B_S$. Choose a basis $\{X_1, \dots, X_n\}$ in $\mathrm{T}_{(\mathbf{x},\mathbf{y})}B_S$ and vectors $Y_1, \dots, Y_r$ complementing this basis to a basis in $\mathrm{T}_{(\mathbf{x},\mathbf{y})}B$. On the one hand, we have $B_S = F^{-1}(0)$ and hence $\ker F'_{(\mathbf{x},\mathbf{y})} = \mathrm{T}_{(\mathbf{x},\mathbf{y})}B_S$. On the other hand, since S is a Morse family, $F'_{(\mathbf{x},\mathbf{y})} = (S''_{xy}(\mathbf{x}, \mathbf{y}), S''_{yy}(\mathbf{x}, \mathbf{y}))$ must have rank r. It follows that $F^*\mathsf{v}_r(Y_1, \dots, Y_r) \neq 0$, so that we can define

$$\mathsf{v}_{B_S}(X_1, \dots, X_n) := \frac{(\pi^*\mathsf{v}_n \wedge \rho^*\mathsf{v}_r)(X_1, \dots, X_n, Y_1, \dots, Y_r)}{F^*\mathsf{v}_r(Y_1, \dots, Y_r)}.$$

[26] Because the analytic solutions are determined up to an additive constant.

[27] With respect to the Euclidean metric.

This defines an n-form on B_S, indeed, because a change in the choice of the Y's would produce the same determinant factor in the numerator and the denominator. Since the right hand side is nonzero, v_{B_S} is a volume form. It satisfies (12.8.25), because $F'_{(\mathbf{x},\mathbf{y})}X_j = 0$ for all $j = 1, \ldots, n$. Finally, the relation (12.8.26) follows from $F^*\mathsf{v}_r = \det(S''_{yy})\rho^*\mathsf{v}_r$. □

By means of v_{B_S}, to $\hat{a}_0$ we can assign a function $a\colon B_S \equiv U \to \mathbb{R}$, defined by

$$a|\mathsf{v}_{B_S}|^{\frac{1}{2}} = \hat{a}_0. \tag{12.8.27}$$

We extend a smoothly to a function on the whole of B with compact support in the y-variables and, instead of the ansatz (12.7.3), we now consider the oscillatory integral

$$\left(\frac{k}{2\pi}\right)^{\frac{r}{2}} \int \mathrm{e}^{\mathrm{i}kS(\mathbf{x},\mathbf{y})} a(\mathbf{x},\mathbf{y})\mathrm{d}^r y. \tag{12.8.28}$$

We assume that a admits an asymptotic expansion of the form (12.7.5). Then, the integral (12.8.28) is absolutely convergent and depends smoothly on $\mathbf{x}$ and k. For the analysis of the large k behaviour of integrals of this type one uses the method of stationary phase. For increasing k the function $y \mapsto \mathrm{e}^{\mathrm{i}kS(\mathbf{x},\mathbf{y})}$ oscillates more and more quickly. Thus, only contributions from neighbourhoods of stationary points of S should count, whereas the other contributions should give zero in the average. This intuition is correct, as the following classical result shows.

Theorem 12.8.4 (Stationary Phase Method) *Let f and φ be smooth functions on $\mathbb{R}^r$. Assume that f has compact support and that φ has a finite number of stationary points $\mathbf{y}_1, \ldots, \mathbf{y}_p$ in the support of f, all of which are non-degenerate. Then,*

$$\int \mathrm{d}^r y f(\mathbf{y})\mathrm{e}^{\mathrm{i}k\varphi(\mathbf{y})} = \left(\frac{2\pi}{k}\right)^{\frac{r}{2}} \sum_{A=1}^{p} \frac{\mathrm{e}^{\mathrm{i}\frac{\pi}{4}\operatorname{sign}\varphi''(\mathbf{y}_A)}}{\sqrt{|\det\varphi''(\mathbf{y}_A)|}} f(\mathbf{y}_A)\mathrm{e}^{\mathrm{i}k\varphi(\mathbf{y}_A)} + O\big(k^{-\frac{r}{2}-1}\big). \tag{12.8.29}$$

Proof Denote $I(k) := \int \mathrm{d}^r y f(\mathbf{y})\mathrm{e}^{\mathrm{i}k\varphi(\mathbf{y})}$. The proof is in three steps.

1. First, we show that if φ does not have stationary points in the support of f, then

$$\lim_{k\to\infty} k^N I(k) = 0 \tag{12.8.30}$$

for all $N > 0$. Indeed, in this case, the differential operator

$$D = \frac{1}{\mathrm{i}k}\frac{1}{\|\nabla\varphi\|^2}\partial_j\varphi\frac{\partial}{\partial y^j}$$

is defined everywhere. Since $D\mathrm{e}^{\mathrm{i}k\varphi} = \mathrm{e}^{\mathrm{i}k\varphi}$, for any positive integer N we have

$$I(k) = \int \mathrm{d}^r y f(\mathbf{y}) D^{N+1} \mathrm{e}^{\mathrm{i}k\varphi(\mathbf{y})} = \int \mathrm{d}^r y \big(D^{\dagger(N+1)} f\big)(\mathbf{y}) \mathrm{e}^{\mathrm{i}k\varphi(\mathbf{y})} = O\big(k^{-N-1}\big),$$

where we have used that the Hermitian conjugate $D^\dagger$ of D is given by

$$D^\dagger f = -\frac{1}{\mathrm{i}k} \partial_j \left(\frac{\partial_j \varphi}{\|\nabla\varphi\|^2} f \right).$$

2. Next, we show that for every smooth function $g\colon \mathbb{R} \to \mathbb{R}$ with compact support and every $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) \geq 0$, we have

$$\int_{-\infty}^{\infty} \mathrm{d}t g(t) \mathrm{e}^{-\frac{1}{2}\lambda t^2} = \sqrt{\frac{2\pi}{\lambda}} g(0) \big(1 + O\big(\lambda^{-1}\big)\big), \tag{12.8.31}$$

where the root is chosen so that $\mathrm{Re}(\sqrt{\lambda}) \geq 0$. Indeed, the Gauss integral formula yields

$$\int_{-\infty}^{\infty} \mathrm{d}t g(t) \mathrm{e}^{-\frac{1}{2}\lambda t^2} = \sqrt{\frac{2\pi}{\lambda}} g(0) + \int_{-\infty}^{\infty} \mathrm{d}t \big(g(t) - g(0)\big) \mathrm{e}^{-\frac{1}{2}\lambda t^2}$$

for every real $\lambda > 0$. The terms on the right hand side can be separately continued analytically to $\mathrm{Re}(\lambda) > 0$ and by continuity to $\mathrm{Re}(\lambda) \geq 0$. Then, the square root in the first term satisfies $\mathrm{Re}(\lambda) \geq 0$. Rewriting the second term as

$$\int_{-\infty}^{\infty} \mathrm{d}t \big(g(t) - g(0)\big) \mathrm{e}^{-\frac{1}{2}\lambda t^2} = \frac{1}{\lambda} \int_{-\infty}^{\infty} \mathrm{d}t \left(\frac{g(t) - g(0)}{t} \right)' \mathrm{e}^{-\frac{1}{2}\lambda t^2},$$

we obtain the assertion.

3. By point 1, we can write the integral $I(k)$ as a sum of integrals over sufficiently small domains U_A, each containing exactly one stationary point $\mathbf{y}_A$ of φ. The contributions omitted this way vanish faster than any power of $\frac{1}{k}$. Since $\mathbf{y}_A$ is assumed to be non-degenerate, $\varphi''(\mathbf{y}_A)$ is invertible. By the Morse Lemma 8.9.4, there exist local coordinates v^i on U_A such that $v^i(\mathbf{y}_A) = 0$ and

$$\varphi\big(y(\mathbf{v})\big) = \varphi(\mathbf{y}_A) + \frac{1}{2} \mathsf{Q}(\mathbf{v}), \qquad \mathsf{Q}(\mathbf{v}) = -v_1^2 - \cdots - v_{i_0}^2 + v_{i_0+1}^2 + \cdots + v_r^2,$$

where i_0 is the Morse index of φ at $\mathbf{y}_A$. Thus, the contribution from U_A is given by

$$I_{U_A}(k) = \mathrm{e}^{\mathrm{i}k\varphi(\mathbf{y}_A)} \int_{U_A} \mathrm{d}^r v \big|J(\mathbf{v})\big| f\big(y(\mathbf{v})\big) \mathrm{e}^{\frac{\mathrm{i}k}{2}\mathsf{Q}(\mathbf{v})},$$

with J denoting the determinant of the Jacobi matrix of the coordinate transformation. Since $\mathbf{v} = 0$ is the only stationary point of Q, by point 1, we may extend this integral to $\mathbb{R}^r$ by choosing an arbitrary extension g of the integrand $|J(\mathbf{v})| f(y(\mathbf{v}))$

to $\mathbb{R}^r$. Since f has compact support, g may be chosen to have compact support, too. Thus, $I_{U_A}(k)$ reads

$$\mathrm{e}^{\mathrm{i}k\varphi(\mathbf{y}_A)}\int \mathrm{d}v_1 \mathrm{e}^{\mathrm{i}\frac{k}{2}v_1^2}\cdots\int \mathrm{d}v_{i_0}\mathrm{e}^{\mathrm{i}\frac{k}{2}v_{i_0}^2}\int \mathrm{d}v_{i_0+1}\mathrm{e}^{-\mathrm{i}\frac{k}{2}v_{i_0+1}^2}\cdots\int \mathrm{d}v_r \mathrm{e}^{-\mathrm{i}\frac{k}{2}v_r^2} g(\mathbf{v}).$$

According to (12.8.31), the integral over v_r yields

$$\sqrt{\frac{2\pi}{-\mathrm{i}k}}\, g(v_1,\ldots,v_{r-1},0)\big(1+O\big(k^{-1}\big)\big).$$

Iterating this argument, we obtain a factor $\sqrt{\frac{2\pi}{-\mathrm{i}k}} = \sqrt{\frac{2\pi}{k}}\mathrm{e}^{\mathrm{i}\frac{\pi}{4}}$ for each integration over v_l with $l > i_0$ and a factor $\sqrt{\frac{2\pi}{\mathrm{i}k}} = \sqrt{\frac{2\pi}{k}}\mathrm{e}^{-\mathrm{i}\frac{\pi}{4}}$ for each integration over v_l with $l \leq i_0$. Thus, we end up with

$$I_{U_A}(k) = \left(\frac{2\pi}{k}\right)^{\frac{r}{2}} \sum_{A=1}^{p} \mathrm{e}^{\mathrm{i}\frac{\pi}{4}\,\mathrm{sign}\,\varphi''(\mathbf{y}_A)}\big|J(0)\big| f(\mathbf{y}_A)\mathrm{e}^{\mathrm{i}k\varphi(\mathbf{y}_A)} + O\big(k^{-\frac{r}{2}-1}\big).$$

It remains to determine $|J(0)|$. For that purpose, we calculate

$$\begin{aligned}\frac{1}{2}\frac{\partial^2 \mathsf{Q}}{\partial v^m \partial v^l}(\mathbf{v}) &= \frac{\partial^2\varphi}{\partial v^m \partial v^l}\big(y(\mathbf{v})\big)\\ &= \frac{\partial^2\varphi}{\partial y^j \partial y^k}\big(y(\mathbf{v})\big)\frac{\partial y^j}{\partial v^m}(\mathbf{v})\frac{\partial y^k}{\partial v^l}(\mathbf{v}) + \frac{\partial\varphi}{\partial y^j}\big(y(\mathbf{v})\big)\frac{\partial^2 y^j}{\partial v^m \mathrm{d}v^l}(\mathbf{v})\end{aligned}$$

and take the absolute value of the determinant at $\mathbf{v} = 0$. Since

$$\frac{1}{2}\frac{\partial^2 \mathsf{Q}}{\partial v^m \partial v^l} = \pm\delta_{ml},$$

this yields

$$\big|J(0)\big| = \big|\det\varphi''(\mathbf{y}_A)\big|^{-\frac{1}{2}}$$

and thus completes the proof of the theorem. □

Now, we apply Theorem 12.8.4 to the integral (12.8.28). Every regular point of $\Pi(U)$ possesses an open neighbourhood W such that Π restricts to a diffeomorphism $\Pi_A : W_A \to W$ on each connected component W_A of $\Pi^{-1}(W)\cap U$. For every $\mathbf{x}\in W$, the stationary points of the mapping $\mathbf{y}\mapsto S(\mathbf{x},\mathbf{y})$ coincide with the y-variables of the points $\Pi_A^{-1}(\mathbf{x})$. Denote

$$\sigma := \mathrm{sign}\big(S''_{yy}\big).$$

Corollary 12.8.5 *For every regular point* $\mathbf{x} \in \Pi(U) \subset Q$, *the integral* (12.8.28) *admits the asymptotic expansion* $u(\mathbf{x}) + O(k^{-1})$, *where*

$$u(\mathbf{x}) = \sum_A a_{0A}(\mathbf{x}) \mathrm{e}^{\mathrm{i}\frac{\pi}{4}\sigma_A(\mathbf{x})} \mathrm{e}^{\mathrm{i}kS_A(\mathbf{x})} \tag{12.8.32}$$

with

$$a_{0A}(\mathbf{x}) = \frac{a(\Pi_A^{-1}(\mathbf{x}))}{\sqrt{|\det S''_{yy}(\Pi_A^{-1}(\mathbf{x}))|}},$$

$$S_A(\mathbf{x}) = S\big(\Pi_A^{-1}(\mathbf{x})\big),$$

$$\sigma_A(\mathbf{x}) = \sigma\big(\Pi_A^{-1}(\mathbf{x})\big).$$

The functions a_{0A} *are analytic solutions of the transport equation and the function* u *is an asymptotic solution of* (12.8.1) *to first order.*

Proof We only have to check that every a_{0A} is an analytic solution of the transport equation. By (12.8.26) and (12.8.27), on W_A,

$$\hat{a}_0 = \big(a|\mathsf{v}_{B_S}|^{\frac{1}{2}}\big) = \frac{a}{\sqrt{|\det S''_{yy}|}} \Pi^*\big(|\mathsf{v}_n|^{\frac{1}{2}}\big) = \Pi^*\big(a_{0A}|\mathsf{v}_n|^{\frac{1}{2}}\big).$$

Since $\hat{a}_0$ is a generalized solution of the transport equation, the assertion follows. □

Remark 12.8.6

1. Since σ jumps at $\Sigma(B_S)$ by a multiple of 2, the relative phases between the summands in (12.8.32) are multiples of $\frac{\pi}{2}$. In particular, if ξ_A and $\xi_{A'}$ can be joined by a curve in Σ which has a single crossing with $\Sigma(B_S)$ and if this crossing is simple and transversal, then we know from the proof of Proposition 12.6.11 that σ jumps by ± 2 at this crossing. Hence, the phase shift between the contributions of ξ_A and $\xi_{A'}$ is $\pm\frac{\pi}{2}$.
2. That a_{0A} is a solution of the transport equation can also be confirmed by a direct computation, see Exercise 12.8.3.

With Corollary 12.8.5 we have accomplished step (a) of our programme for fixing the relative phases, outlined on page 714: the constant phase factors $\mathrm{e}^{\frac{\mathrm{i}\pi}{4}\sigma_A}$ yield the desired relative phases between contributions to the asymptotic solution of (12.8.1) stemming from one and the same element of the chosen covering of $(\mathscr{S}, \iota)$.

To accomplish step (b) we must analyze the condition that the contributions of different elements U_i and U_j coincide if they come from the intersection $U_i \cap U_j$. For that purpose, we first note that the solution u given by (12.8.32) is obtained from the half density

$$\frac{a}{\sqrt{|\det S''_{yy}|}} \mathrm{e}^{\frac{\mathrm{i}\pi}{4}\sigma} \mathrm{e}^{\mathrm{i}kS} \Pi^*\big(|\mathsf{v}_n|^{\frac{1}{2}}\big) = \mathrm{e}^{\mathrm{i}(kS+\frac{\pi}{4}\sigma)} \hat{a}_0$$

on $U \cap \Sigma_0(\mathscr{S})$, where by an abuse of notation the restriction of the Morse family S to U is denoted by the same symbol.[28] Thus, from the covering $\{U_j\}$ and the corresponding system $\{(B_j, \pi_j, S_j)\}$ of generating Morse families, we obtain a system $\{\hat{u}_j\}$ of local half densities on $\Sigma_0(\mathscr{S})$ given by

$$\hat{u}_j = \mathrm{e}^{\mathrm{i}(kS_j + \frac{\pi}{4}\sigma_j)} \hat{a}_0. \tag{12.8.33}$$

The local half densities $\hat{u}_i$ and $\hat{u}_j$ coincide on $U_i \cap U_j \cap \Sigma_0(\mathscr{S})$ and hence combine to a half density $\hat{u}$ on $\Sigma_0(\mathscr{S})$ iff their phases coincide modulo 2π at every regular point of $U_i \cap U_j$:

$$kS_i + \frac{\pi}{4}\sigma_i = kS_j + \frac{\pi}{4}\sigma_j \mod 2\pi. \tag{12.8.34}$$

This is the consistency condition announced in step (b) above. It is known as the Bohr-Sommerfeld quantization condition. A topological interpretation of this condition will be given below. The half-density $\hat{u}$ may be referred to as a generalized first order solution of (12.8.1). By means of a partition of unity $\{\chi_j\}$ subordinate to the covering $\{U_j\}$, it is explicitly given by $\hat{u} = \sum_j \chi_j \hat{u}_j$, and projection to Q yields

$$u(\mathbf{x}) = \sum_j \sum_{\xi \in \Pi^{-1}(\mathbf{x}) \cap U_j} \chi_j(\xi) \frac{a_j(\xi)}{\sqrt{|\det(S_j)''_{yy}(\xi)|}} \mathrm{e}^{\mathrm{i}(kS_j(\xi) + \frac{\pi}{4}\sigma_j(\xi))}. \tag{12.8.35}$$

This is the desired global first order solution $u : Q \setminus \Gamma(\mathscr{S}) \to \mathbb{R}$ of (12.8.1) associated with the generalized solutions $(\mathscr{S}, \iota)$ and $\hat{a}_0$ of the characteristic equation and the transport equation, respectively. By the Bohr-Sommerfeld quantization condition, u does not depend on the choice of the partition of unity. It is, therefore, uniquely determined by $(\mathscr{S}, \iota)$ and $\hat{a}_0$, indeed.

The solution (12.8.35) can be rewritten as a so-called oscillatory half density on $Q \setminus \Gamma(\mathscr{S})$ as follows. For every point of $Q \setminus \Gamma(\mathscr{S})$, there exists an open neighbourhood W such that Π restricts to a diffeomorphism $\Pi_A : W_A \to W$ on each connected component W_A of $\Pi^{-1}(W)$. Then, on W, we have

$$u|\mathsf{v}_n|^{\frac{1}{2}} = K^{\mathscr{S}} \hat{u} := \sum_A \big(\Pi_A^{-1}\big)^* \hat{u}.$$

The mapping $K^{\mathscr{S}}$ is known as Maslov's canonical operator. Let us derive an explicit formula for $K^{\mathscr{S}}$ in terms of the canonical 1-form θ on T^*Q and the Maslov intersection index of curves in $\mathscr{S}$. For a detailed presentation we refer to [198].

[28] For convenience, in the remainder, we stick to this simplified notation.

Proposition 12.8.7 *Let* $\mathbf{x} \in Q \setminus \Gamma(\mathscr{S})$. *Let* $\xi_0 \in \mathscr{S}$ *be a regular point, let* γ_A *be curves*[29] *from* ξ_0 *to* $\Pi_A^{-1}(\mathbf{x})$ *and let* $S(\xi_0)$ *be an arbitrarily chosen phase. Then,*

$$\big(K^{\mathscr{S}}\hat{u}\big)(\mathbf{x}) = \sum_A \exp\left\{\mathrm{i}\left(k\int_{\gamma_A}\theta - \frac{\pi}{2}\mathrm{ind}_{\mathscr{S}}(\gamma_A) + kS(\xi_0)\right)\right\}\big(a_{0A}|\mathsf{v}_n|^{\frac{1}{2}}\big)(\mathbf{x}),$$

where $\mathrm{ind}_{\mathscr{S}}(\gamma_A)$ *is the intersection index of* γ_A *with* $\Sigma(\mathscr{S})$ *and* $a_{0A}|\mathsf{v}_n|^{\frac{1}{2}} = (\Pi_A^{-1})^*\hat{a}_0$.

Proof It suffices to verify the asserted formula for every A separately. Thus, let $\gamma : [0, 1] \to \mathscr{S}$ be a curve from ξ_0 to one of the ξ_A. Choose j such that $\xi_A \in U_j$. Then,

$$\big((\Pi_A^{-1})^*\hat{u}\big)(\mathbf{x}) = \big((\Pi_A^{-1})^*\hat{u}_j\big)(\mathbf{x}) = \mathrm{e}^{\mathrm{i}(kS_j(\xi_A)+\frac{\pi}{4}\sigma_j(\xi_A))}\big(a_{0A}|\mathsf{v}_n|^{\frac{1}{2}}\big)(\mathbf{x}). \quad (12.8.36)$$

Choosing $0 = t_0 < t_1 < \cdots < t_l < t_{l+1} = 1$ and $j_0, \dots, j_l = j$ such that $\gamma([t_i, t_{i+1}]) \subset U_{j_i}$ for all $i = 0, \dots, l$, we can rewrite $S_j(\xi_A)$ and $\sigma_j(\xi_A)$ as

$$\begin{aligned} S_j(\xi_A) &= \sum_{i=1}^{l}\int_{\gamma([t_i,t_{i+1}])} \mathrm{d}S_{j_i} + \sum_{i=1}^{l}\big(S_{j_i}\big(\gamma(t_i)\big) - S_{j_{i-1}}\big(\gamma(t_i)\big)\big) + S_{j_0}(\xi_0),\\ \sigma_j(\xi_A) &= \sum_{i=0}^{r}\big(\sigma_{j_i}\big(\gamma(t_{i+1})\big) - \sigma_{j_i}\big(\gamma(t_i)\big)\big)\\ &\quad + \sum_{i=1}^{r}\big(\sigma_{j_i}\big(\gamma(t_i)\big) - \sigma_{j_{i-1}}\big(\gamma(t_i)\big)\big) + \sigma_{j_0}(\xi_0). \end{aligned}$$

On the one hand, by (12.4.10), under the identification of U_{j_i} with $B_{S_{j_i}}$, we have

$$\mathrm{d}S_{j_i} = \iota^*\theta \quad (12.8.37)$$

for all i. Hence, the first term in the equation for $S_j(\xi_A)$ yields

$$\sum_{i=1}^{l}\int_{\gamma([t_i,t_{i+1}])} \mathrm{d}S_{j_i} = \sum_{i=1}^{l}\int_{\gamma([t_i,t_{i+1}])} \iota^*\theta = \int_\gamma \iota^*\theta = \int_{\iota\circ\gamma}\theta \equiv \int_\gamma \theta.$$

On the other hand, by Proposition 12.6.11 and (12.6.16), the first term in the equation for $\sigma_j(\xi_A)$ yields

$$\sum_{i=0}^{r}\big(\sigma_{j_i}\big(\gamma(t_{i+1})\big) - \sigma_{j_i}\big(\gamma(t_i)\big)\big) = -2\mathrm{ind}_{\mathscr{S}}(\gamma).$$

[29] Which can always be chosen to intersect the singular subset $\Sigma(\mathscr{S})$ transversally.

The asserted formula now follows by inserting all of this into (12.8.36) and taking into account the Bohr-Sommerfeld quantization condition (12.8.34). □

Remark 12.8.8

1. We give a topological interpretation of the Bohr-Sommerfeld condition. For that purpose, we observe that the functions $s_{ij} := S_i - S_j$ on $U_i \cap U_j$ are constant, because (12.8.37) implies that on $U_i \cap U_j$ we have

$$\mathrm{d}S_i = \iota^*\theta = \mathrm{d}S_j. \tag{12.8.38}$$

Moreover, from Lemma 12.6.10 we know that the integer-valued functions

$$c_{ij} = \operatorname{index}\big((S_i)''_{yy}\big) - \operatorname{index}\big((S_j)''_{yy}\big)$$

on $U_i \cap U_j$ are constant, too, and hence smooth. As noted in Sect. 12.6, the family $\{c_{ij}\}$ defines an element $\hat{\mu}_{\mathscr{S}}$ of the first integer-valued Čech cohomology $H^1_{\check{c}}(\mathscr{S}, \mathbb{Z})$ of $\mathscr{S}$. The image of this element under the canonical homomorphism

$$h : H^1_{\check{c}}(\mathscr{S}, \mathbb{Z}) \to H^1(\mathscr{S}, \mathbb{R})$$

to the first de Rham cohomology $H^1(\mathscr{S}, \mathbb{R})$ coincides with the Maslov class $\mu_{\mathscr{S}}$ of $(\mathscr{S}, \iota)$. Analogously, the family $\{s_{ij}\}$ defines a 1-cocycle and thus an element of the first real-valued Čech cohomology of $\mathscr{S}$. The image $\alpha_{\mathscr{S}}$ of this class under h is usually called the Liouville class of $(\mathscr{S}, \iota)$. By (12.8.38), it satisfies $\alpha_{\mathscr{S}} = [\iota^*\theta]$. By Formula (12.6.16) we have

$$c_{ij} = \frac{1}{2}\big((r_i - r_j) - (\sigma_i - \sigma_j)\big), \tag{12.8.39}$$

where r_i and r_j are the fibre dimensions of B_i and B_j, respectively. Thus, the Bohr-Sommerfeld condition takes the form

$$\frac{k}{2\pi} s_{ij} - \frac{1}{4} c_{ij} + \frac{1}{8}(r_i - r_j) \in H^1(\mathscr{S}, \mathbb{Z}).$$

Since the third term vanishes under h, the Bohr-Sommerfeld quantization condition (12.8.34) can be rewritten in terms of the Maslov and the Liouville classes as

$$\frac{k}{2\pi}\alpha_{\mathscr{S}} - \frac{1}{4}\mu_{\mathscr{S}} \in h\big(H^1_{\check{c}}(\mathscr{S}, \mathbb{Z})\big). \tag{12.8.40}$$

The cohomology class

$$\varphi_{\mathscr{S},k} := \frac{k}{2\pi}\alpha_{\mathscr{S}} - \frac{1}{4}\mu_{\mathscr{S}}$$

is called the phase class associated with $(\mathscr{S}, \iota)$ and k.

2. Let us analyze which geometric object the local half-densities $\hat{u}_j$ combine to in the general case. According to Theorem 2.2.11, the families of transition mappings $\{e^{iks_{ij}}\}$, $\{e^{-i\frac{\pi}{2}c_{ij}}\}$ and $\{e^{i(ks_{ij}-\frac{\pi}{2}c_{ij})}\}$ define complex line bundles over $\mathscr{S}$. The first one will be denoted by $\mathscr{E}_{\mathscr{S},k}$.[30] The second one is the Maslov line bundle $\mathscr{M}_{\mathscr{S}}$ constructed in Sect. 12.6, cf. Remark 12.6.14. The third one is the tensor product

$$\Phi_{\mathscr{S},k} := \mathscr{E}_{\mathscr{S},k} \otimes \mathscr{M}_{\mathscr{S}},$$

called the phase bundle associated with $(\mathscr{S}, \iota)$ and k. Note that the bundle $\mathscr{E}_{\mathscr{S},k}$ is trivial (like $\mathscr{M}_{\mathscr{S}}$), because the family $\{e^{ikS_i}\}$ defines a global non-vanishing section in this bundle, cf. Example 2.3.3/2. The family $\{e^{i\frac{\pi}{4}\sigma_i}\}$ defines a section in the complex line bundle generated by the family of transition mappings $\{e^{i\frac{\pi}{4}\sigma_{ij}}\}$. Due to (12.8.39),

$$e^{-i\frac{\pi}{2}c_{ij}} = e^{i\frac{\pi}{4}r_j} e^{i\frac{\pi}{4}\sigma_{ij}} e^{-i\frac{\pi}{4}r_i},$$

so that this bundle is naturally isomorphic to the Maslov line bundle $\mathscr{M}_{\mathscr{S}}$, cf. Remark 2.2.12/1. Via this natural isomorphism, the family $\{e^{i\frac{\pi}{4}\sigma_i}\}$ defines a section in $\mathscr{M}_{\mathscr{S}}$. We conclude that, in the general case, the local half-densities $\hat{u}_j$ on $\Sigma_0(\mathscr{S})$ combine to a section in the restriction of the line bundle $\Phi_{\mathscr{S},k} \otimes |\Lambda|^{\frac{1}{2}}\mathscr{S}$ to $\Sigma_0(\mathscr{S})$.

By means of the phase bundle $\Phi_{\mathscr{S},k}$, the Bohr-Sommerfeld quantization condition can also be interpreted geometrically as follows. Recall from Sect. 12.6 that the Maslov line bundle is associated with a principal $\mathbb{Z}_4$-bundle over $\mathscr{S}$. Since the latter has discrete structure group, it carries a unique connection, given by the unique lift of curves in $\mathscr{S}$ to curves in this bundle. This connection induces a natural connection in the Maslov line bundle $\mathscr{M}_{\mathscr{S}}$ and hence in the phase bundle $\Phi_{\mathscr{S},k}$. Let $\Gamma_{\mathrm{par}}(\Phi_{\mathscr{S},k})$ be the space of the sections of $\Phi_{\mathscr{S},k}$ which are parallel with respect to this connection. One can show that parallel sections exist iff the phase class is integer-valued, that is, iff the Bohr-Sommerfeld quantization condition holds. Moreover, if parallel sections exist, they are unique up to a constant factor and hence given by a constant multiple of the family $\{e^{i(kS_j+\frac{\pi}{4}\sigma_j)}\}$.

Let us add that it is common to pass to a k-independent universal object by building the $\mathbb{C}$-module

$$\Psi_{\mathscr{S}} := \Gamma\big(|\Lambda|^{\frac{1}{2}}\mathscr{S}\big) \otimes_{\mathbb{C}} \left(\prod_{k>0} \Gamma_{\mathrm{par}}(\Phi_{\mathscr{S},k})\right),$$

called the symbol space of $(\mathscr{S}, \iota)$. Obviously, if the Bohr-Sommerfeld quantization condition holds, the half-density $\hat{u}$ is a symbol (an element of $\Psi_{\mathscr{S}}$). Thus, the above construction assigns a symbol, $\hat{u}$, to every generalized solution $(\mathscr{S}, \iota)$ of the characteristic equation, every generalized solution $\hat{a}_0$ of the

[30] $\mathscr{E}_{\mathscr{S},k}$ may be identified with the pull-back under ι of the so-called prequantum line bundle over $(\mathrm{T}^*Q, \mathrm{d}\theta)$, see Sect. 4.1 and Appendix D of [36] for a detailed description.

transport equation and every $k > 0$ such that the Bohr-Sommerfeld quantization condition (12.8.34) is satisfied.

Example 12.8.9 (Schrödinger equation) We consider the time-independent Schrödinger equation of a particle with unit mass in one space dimension under the influence of an external potential V,

$$\left(-\frac{\hbar^2}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + (V - E)\right)\psi = 0. \tag{12.8.41}$$

This equation is of the form (12.8.1), where the symbol H is given by the classical Hamiltonian function of the system,

$$H(x, p) = \frac{p^2}{2} + V - E,$$

and $\frac{1}{k}$ is given by Planck's constant $\hbar$. Thus, the short wave asymptotics of the solutions ψ corresponds to the semiclassical, or WKB,[31] approximation $\hbar \to 0$. Let $\mathscr{C}_E$ be a regular compact connected component of the energy surface $H^{-1}(0)$. Then, $\mathscr{C}_E$ is a smooth closed curve in $\mathrm{T}^*\mathbb{R} \cong \mathbb{R}^2$. For dimensional reasons, $\mathscr{C}_E$ is a Lagrangian submanifold. Assume that V is such that $\mathscr{C}_E$ has two intersection points x_+ and x_- with the x-axis. Then, $E = V(x_\pm)$ and the singular subset $\Sigma(\mathscr{C}_E)$ consists of x_+ and x_-. From Example 12.6.8/1 we know that the Maslov index $\mu_{\mathscr{C}_E}$ of the fundamental cycle of $\mathscr{C}_E$ is equal to 2. Consequently, by integrating the phase class $\varphi_{\mathscr{S},\frac{1}{\hbar}}$ over $\mathscr{C}_E$, we find that the Bohr-Sommerfeld condition (12.8.40) reads

$$\frac{1}{2\pi\hbar}\int_{\mathscr{C}_E}\theta - \frac{1}{2} \in \mathbb{Z}.$$

Obviously, $\int_{\mathscr{C}_E}\theta = A(E)$, with $A(E)$ being the area enclosed by the curve $\mathscr{C}_E$. As a result, the Bohr-Sommerfeld quantization condition for the Schrödinger equation in one dimension takes the form

$$A(E) = 2\pi\hbar\left(n + \frac{1}{2}\right), \quad n = 0, 1, \ldots. \tag{12.8.42}$$

For example, for the harmonic oscillator with potential $V(x) = \frac{\omega^2}{2}x^2$, the curve $\mathscr{C}_E$ is an ellipse with semiaxes $\sqrt{2E}$ in momentum direction and $\frac{\sqrt{2E}}{\omega}$ in position direction. Hence, $A(E) = 2\pi\frac{E}{\omega}$, so that the Bohr-Sommerfeld quantization condition becomes

$$E = \hbar\left(n + \frac{1}{2}\right).$$

[31] Named after Wentzel, Brillouin and Kramers.

This is the exact formula for the quantum energy levels. The quantization condition (12.8.42) generalizes in an obvious way to an arbitrary integrable system.

Next, we solve the Schrödinger equation (12.8.41) in the WKB approximation, that is, we determine the short wave asymptotic solutions of this equation. The characteristic equation (12.8.2) reads

$$\frac{1}{2}\left(\frac{\mathrm{d}S}{\mathrm{d}x}\right)^2 + V = E \tag{12.8.43}$$

and the transport equation in the form (12.8.8) is given by

$$\sqrt{2(E-V)}\frac{\mathrm{d}}{\mathrm{d}x}a_0^2 + a_0^2\frac{\mathrm{d}}{\mathrm{d}x}\sqrt{2(E-V)} = 0. \tag{12.8.44}$$

We restrict our attention to the interior of the classically allowed region, where $E > V$. Here, the solutions of (12.8.43) and (12.8.44) are given by

$$S(x) = \pm\int\sqrt{2\big(E-V(x)\big)}\mathrm{d}x, \qquad a_0(x) = c\big(2\big(E-V(x)\big)\big)^{-\frac{1}{4}},$$

respectively. To determine the corresponding asymptotic solution of the Schrödinger equation (12.8.41), we apply Proposition 12.8.7. Let us choose ξ_0 in the region of $\mathscr{C}_E$ where $p > 0$ and let us put $S(\xi_0) = 0$. For every x in the interior of the classically allowed region, $\Pi^{-1}(x)$ consists of two points, ξ_+ with $p > 0$ and ξ_- with $p < 0$. Connecting ξ_0 with ξ_- by a clockwise oriented curve, we obtain

$$\psi(x) = \frac{c}{\sqrt[4]{2(E-V(x))}}\left\{\mathrm{e}^{\frac{\mathrm{i}}{\hbar}\int\sqrt{2(E-V(x))}\mathrm{d}x} + \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\int\sqrt{2(E-V(x))}\mathrm{d}x + \frac{\mathrm{i}\pi}{2}}\right\} + O(\hbar),$$

because such a curve has Maslov index -1. This formula yields the WKB approximation for the solutions of the Schrödinger equation (12.8.41) in the classically allowed region. Up to a constant phase factor, it can be rewritten as a real-valued function:

$$\psi(x) \sim \frac{1}{\sqrt[4]{2(E-V(x))}}\cos\left(\frac{1}{\hbar}\int\sqrt{2\big(E-V(x)\big)}\mathrm{d}x - \frac{\pi}{4}\right).$$

We encourage the reader to compare the derivation given here with the usual derivation of this formula in the standard text books of quantum mechanics, see e.g. [101].

For an exhaustive discussion of the n-dimensional Schrödinger equation we refer to [198] and for the study of general spectrum conditions in the semiclassical approximation we recommend [78]. The above discussed structures also yield a general geometric framework for WKB quantization. Regarding this aspect, we make the following remark. For details, the reader may consult [36].

Remark 12.8.10 As mentioned above, the line bundle $\mathscr{E}_{\mathscr{S},k}$ can be identified with the pull-back to the Lagrangian immersion $(\mathscr{S},\iota)$ of the prequantum line bundle

$\mathscr{E}_{\mathrm{T}^*Q,k}$ over T^*Q and the phase bundle can be identified with

$$\Phi_{\mathscr{S},k} = \iota^* \mathscr{E}_{\mathrm{T}^*Q,k} \otimes \mathscr{M}_{\mathscr{S}}.$$

One says that a Lagrangian immersion $(\mathscr{S}, \iota)$ of T^*Q is quantizable if it satisfies the Bohr-Sommerfeld quantization condition with $k = \frac{1}{\hbar}$. By Remark 12.8.8/1, this is equivalent to the requirements that the phase class $\varphi_{\mathscr{S},k}$ be integer-valued, or that the phase bundle admit a global parallel section $\hat{u}$. The pair consisting of $(\mathscr{S}, \iota)$ and $\hat{u}$ is called a semiclassical state and Maslov's canonical operator is called a semiclassical quantization mapping.

The method of stationary phase is restricted to points apart from the caustic, because Corollary 12.8.5 applies to such points only. Over focal points, every generating Morse family necessarily has degenerate critical points, so that this method cannot be applied. To deal with this situation in a systematic way, one has to use distribution-valued half-densities, see [115], [141] and [36]. However, if we have a normal form for the generating families in the neighbourhood of the caustic at our disposal, we can nevertheless get insight into the behaviour of asymptotic solutions near the caustic. We discuss this for the simplest type, the fold. For an exhaustive treatment we refer again to [115]. From Proposition 12.6.15 we read off that the oscillatory integral (12.8.28) has the form

$$\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k(f(\mathbf{x})+g(\mathbf{x})y-\frac{1}{3}y^3)} a(\mathbf{x}, y)\mathrm{d}y, \qquad (12.8.45)$$

where $\mathbf{x} \in Q = \mathbb{R}^n$. The Malgrange Preparation Theorem[32] entails that a can be written in the form

$$\begin{aligned} a(\mathbf{x}, y) &= b_0(\mathbf{x}) + b_1(\mathbf{x})y + h(\mathbf{x}, y)\frac{\partial S}{\partial y}(\mathbf{x}, y) \\ &= b_0(\mathbf{x}) + b_1(\mathbf{x})y + h(\mathbf{x}, y)\big(g(\mathbf{x}) - y^2\big), \end{aligned}$$

where b_0, b_1 and h are smooth functions. Inserting this into (12.8.45) and performing partial integration, we obtain

$$\mathrm{e}^{\mathrm{i}kf(\mathbf{x})}\left\{a_0(\mathbf{x}, k)\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k(g(\mathbf{x})y-\frac{1}{3}y^3)}\mathrm{d}y + a_1(\mathbf{x}, k)\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}k(g(\mathbf{x})y-\frac{1}{3}y^3)}y\mathrm{d}y\right\}, \qquad (12.8.46)$$

with a_0 and a_1 denoting certain asymptotic series. Using the Airy functions

$$A(\tau) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}(\tau y-\frac{1}{3}y^3)}\mathrm{d}y, \quad \tau \in \mathbb{R},$$

[32] See e.g. [110, §IV.2].

we can rewrite this formula as

$$e^{ikf(\mathbf{x})}\left\{\frac{a_0(\mathbf{x},k)}{k^{\frac{1}{3}}}A\big(k^{\frac{2}{3}}g(\mathbf{x})\big)+\frac{a_1(\mathbf{x},k)}{ik^{\frac{2}{3}}}A'\big(k^{\frac{2}{3}}g(\mathbf{x})\big)\right\}. \tag{12.8.47}$$

Thus, we obtain the following asymptotic expansion of the oscillatory integral (12.8.45):

$$e^{ikf(\mathbf{x})}\left\{\frac{b_0(\mathbf{x})}{k^{\frac{1}{3}}}A\big(k^{\frac{2}{3}}g(\mathbf{x})\big)+\frac{b_1(\mathbf{x})}{ik^{\frac{2}{3}}}A'\big(k^{\frac{2}{3}}g(\mathbf{x})\big)\right\}+O\left(\frac{1}{k}\right). \tag{12.8.48}$$

Example 12.8.11 (Helmholtz equation) Let us return to the Helmholtz equation (12.7.2). We will use Formula (12.8.48) to study the qualitative behaviour of light rays in the neighbourhood of the caustic.[33] Inserting the normal form (12.6.28) for S into the eikonal equation, one obtains

$$(\nabla f)^2+y^2(\nabla g)^2+2y\nabla f\cdot\nabla g=1.$$

According to (12.6.29), on B_S we have $g(\mathbf{x})=y^2$. Hence, for every $\mathbf{x}$, the variable y can take the values $\pm\sqrt{g(\mathbf{x})}$, and thus this equation breaks into the two equations

$$(\nabla f)^2+g(\nabla g)^2=1, \qquad 2\nabla f\cdot\nabla g=0. \tag{12.8.49}$$

Inserting the asymptotic expansion (12.8.48) of the oscillatory integral (12.8.45) into the Helmholtz equation and comparing coefficients, one obtains the following form of the transport equation, see Exercise 12.8.4:

$$2\nabla f\cdot\nabla b_0+\Delta f b_0+2g\nabla g\cdot\nabla b_1+g\Delta g b_1+(\nabla g)^2 b_1=0, \tag{12.8.50}$$

$$2\nabla g\cdot\nabla b_0+\Delta g b_0+2\nabla f\cdot\nabla b_1+\Delta f b_1=0. \tag{12.8.51}$$

Thus, Formula (12.8.48) yields a solution u of the Helmholtz equation up to first order in $\frac{1}{k}$ provided the functions f, g, b_0 and b_1 solve Eqs. (12.8.49), (12.8.50) and (12.8.51). Equations (12.8.49) form a system of nonlinear partial differential equations, which is elliptic for $g<0$, hyperbolic for $g>0$ and parabolic for $g=0$. According to (12.6.29), these cases correspond, respectively, to points in the shadow, points in the illuminated region and points on the caustic. Using asymptotic formulae for the Airy functions for large positive arguments [115] one finds that for large k and $g>0$ one has

$$u\sim\frac{k^{-\frac{1}{3}}e^{ikf}}{\sqrt{\pi}(k^{\frac{2}{3}}g)^{\frac{1}{4}}}\left\{b_0\cos\left(\frac{2kg^{\frac{3}{2}}}{3}-\frac{\pi}{4}\right)-b_1g^{\frac{1}{2}}\sin\left(\frac{2kg^{\frac{3}{2}}}{3}-\frac{\pi}{4}\right)\right\}. \tag{12.8.52}$$

This formula yields a model for the intensity of light in the illuminated region. For large k and $g<0$ one finds an exponentially decaying solution, which of course

[33] This goes back to Ludwig [187], see also [86, 157].

corresponds to the shaded region. However, close to the caustic, where $k^{\frac{2}{3}}g$ is small even for large k, a detailed analysis yields

$$u \sim \frac{b_0}{k^{\frac{1}{3}}} A\big(k^{\frac{2}{3}} g\big) \sim \frac{b_0}{k^{\frac{1}{3}}} A(0). \tag{12.8.53}$$

Comparing (12.8.52) with (12.8.53), we see that the value of u on the caustic is of a similar magnitude as the value of u in the illuminated region, multiplied by $k^{\frac{1}{6}}$. Let us conclude. While the ordinary short wave asymptotics given by (12.8.21) yields a divergent intensity of light on the caustic, the finer analysis sketched above provides an estimate for the intensity near the caustic. In particular, the intensity on the caustic is large but finite.

Exercises

12.8.1 Prove that the local solution of Eq. (12.8.8) is given by the formula in Remark 12.8.2/1.
Hint. Show that $\frac{1}{\mathscr{J}} \frac{\partial \mathscr{J}}{\partial t}(t, \mathbf{x}) = (\nabla \cdot Y_H)(\phi_t(\mathbf{x}))$.

12.8.2 Confirm Formula (12.8.19).

12.8.3 Prove by direct inspection that the function a_{0A} given in Corollary (12.8.5) is a solution of the transport equation.
Hint. A guide to the proof can be found in [187].

12.8.4 Use the identities

$$A''(\tau) + \tau A(\tau) = 0, \qquad A'''(\tau) + \tau A'(\tau) + A(\tau) = 0$$

for the Airy functions to verify Formulae (12.8.50) and (12.8.51).

References

1. Abraham, R., Marsden, J.E.: Foundations of Mechanics. Benjamin-Cummings, Reading (1978)
2. Abraham, R., Robbin, J.: Transversal Mappings and Flows. Benjamin, New York (1967). With an appendix by Al Kelley
3. Abramowitz, M., Stegun, I.A.: Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. National Bureau of Standards Applied Mathematics Series, vol. 55 (1964)
4. Adams, J.F.: Vector fields on spheres. Ann. Math. **75**, 603–632 (1962)
5. Adler, M.: On a trace functional for formal pseudo-differential operators and symplectic structure of the Korteweg-de Vries equation. Invent. Math. **50**, 219–249 (1979)
6. Agricola, I., Friedrich, T.: Globale Analysis. Vieweg, Wiesbaden (2001), in German
7. Andoyer, H.: Cours de Mécanique Céleste. Gauthier-Villars, Paris (1923)
8. Anné, C.: A topological definition of the Maslov bundle. Math. J. Univ. Frontera **8**, 1–15 (2006)
9. Anosov, D.V., Arnold, V.I. (eds.): Dynamical Systems I. Encyclopaedia Math. Sci., vol. 1. Springer, Berlin (1988)
10. Arms, J.M., Cushman, R., Gotay, M.J.: A universal reduction procedure for Hamiltonian group actions. In: Ratiu, T.S. (ed.) The Geometry of Hamiltonian Systems, pp. 33–51. Springer, Berlin (1991)
11. Arnold, V.I.: Proof of A.N. Kolmogorov's theorem on the preservation of quasi-periodic motions under small perturbations of the Hamiltonian. Russ. Math. Surv. **18**, 9–36 (1963)
12. Arnold, V.I.: Small divisor problems in classical and celestial mechanics. Russ. Math. Surv. **18**, 85–192 (1963)
13. Arnold, V.I.: On a characteristic class entering in quantization conditions. Funct. Anal. Appl. **1**, 1–13 (1967)
14. Arnold, V.I.: Singularities of smooth mappings. Russ. Math. Surv. **23**, 1–43 (1968)
15. Arnold, V.I.: Normal forms of functions near degenerate critical points, Weyl groups A_k, D_k, E_k, and Lagrangian singularities. Funct. Anal. Appl. **6**, 254–272 (1972)
16. Arnold, V.I.: Critical points of smooth functions and their normal forms. Russ. Math. Surv. **30**, 1–75 (1975)
17. Arnold, V.I.: Geometrical Methods in the Theory of Ordinary Differential Equations. Grundlehren Math. Wiss., vol. 250. Springer, Berlin (1988)
18. Arnold, V.I.: Mathematische Methoden der klassischen Mechanik. Birkhäuser, Basel (1988)
19. Arnold, V.I.: Singularities of Caustics and Wave Fronts. Kluwer Academic, Norwell (1990)
20. Arnold, V.I.: Catastrophe Theory. Springer, Berlin (1992)
21. Arnold, V.I.: Ordinary Differential Equations. Springer, Berlin (1992)
22. Arnold, V.I.: Symplectic geometry and topology. J. Math. Phys. **41**(6), 3307–3343 (2000)

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*, Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7,

23. Arnold, V.I., Avez, A.: Ergodic Problems of Classical Mechanics. Benjamin, Elmsford (1968)
24. Arnold, V.I., Givental, A.B.: Symplectic geometry. In: Arnold, V.I., Novikov, S.P. (eds.) Dynamical Systems IV. Symplectic Geometry and Its Applications. Springer, Berlin (2001)
25. Arnold, V.I., Goryunov, V.V., Lyashko, O.V., Vasil'ev, V.A.: Singularity Theory, vol. I. Springer, Berlin (1998)
26. Arnold, V.I., Kozlov, V.V., Neishtadt, A.I.: Mathematical aspects of classical and celestial mechanics. In: Arnold, V.I. (ed.) Dynamical Systems III. Springer, Berlin (1988)
27. Atiyah, M.F., et al.: Responses to theoretical mathematics: toward a cultural synthesis of mathematics and theoretical physics, by A. Jaffe and F. Quinn. Bull. Am. Math. Soc. **30**(2), 178–207 (1994).
28. Audin, M., Cannas da Silva, A., Lerman, E.: Symplectic Geometry of Integrable Systems. Birkhäuser, Basel (2003)
29. Baez, J.: The octonions. Bull. Am. Math. Soc. **39**, 145–205 (2001)
30. Banyaga, A.: Sur la structure du groupe des difféomorphismes qui préservent une forme symplectique. Comment. Math. Helv. **53**, 174–227 (1978)
31. Banyaga, A.: The Structure of Classical Diffeomorphism Groups. Mathematics and Its Applications, vol. 400. Kluwer Academic, Norwell (1997)
32. Banyaga, A., Hurtubise, D.E.: A proof of the Morse-Bott lemma. Expo. Math. **22**, 365–373 (2004)
33. Bates, L., Lerman, E.: Proper group actions and symplectic stratified spaces. Pac. J. Math. **181**, 201–229 (1997)
34. Bates, L., Śniatycki, J.: On action-angle variables. Arch. Ration. Mech. Anal. **120**, 337–343 (1992)
35. Bates, L., Śniatycki, J.: Nonholonomic reduction. Rep. Math. Phys. **32**(1), 99–115 (1993)
36. Bates, S., Weinstein, A.: Lectures on the Geometry of Quantization. Berkeley Mathematics Lecture Notes, vol. 8. AMS, Providence (1995)
37. Benenti, S.: The category of symplectic reductions. In: Benenti, S., Francaviglia, M., Lichnerowicz, A. (eds.) Proceedings of the IUTAM-ISIMM Symposium on Modern Developments in Analytical Mechanics, pp. 39–91. Acta Academiae Scientiarum Taurinensis, Torino (1983)
38. Benenti, S.: Hamiltonian Optics and Generating Families. Napoli Series on Physics and Astrophysics, vol. 7. Bibliopolis, Naples (2003)
39. Benenti, S.: A 'user-friendly' approach to the dynamical equations of non-holonomic systems. Contribution to the Vadim Kuznetsow Memorial issue, *Integrable Systems and Related Topics*. SIGMA Symmetry Integrability Geom. Methods Appl. **3**, paper 036 (2007)
40. Benenti, S., Tulczyjew, W.M.: The geometrical meaning and globalization of the Hamilton-Jacobi method. In: Proc. Differential Geometrical Methods in Mathematical Physics, Aix-en-Provence, Salamanca, 1979. Lecture Notes in Mathematics, vol. 836. Springer, Berlin (1980)
41. Benenti, S., Tulczyjew, W.M.: Geometry of static mechanical systems with constraints. Ann. Mat. Pura Appl. **124**, 139–160 (1980)
42. Benenti, S., Tulczyjew, W.M.: Remarques sur les réductions symplectiques. C. R. Acad. Sci. Paris **294**, 561–564 (1982)
43. Benettin, G.: The elements of Hamiltonian perturbation theory. In: Benest, D., Froeschle, C., Lega, E. (eds.) Hamiltonian Systems and Fourier Analysis, pp. 1–98. Cambridge Scientific, Cambridge (2004)
44. Bhatia, N.P., Szegő, G.P.: Stability Theory of Dynamical Systems. Grundlehren Math. Wiss., vol. 161. Springer, Berlin (1970)
45. Birkhoff, G.D.: Proof of Poincaré's geometric theorem. Trans. Am. Math. Soc. **14**, 14–22 (1913)
46. Birkhoff, G.D.: An extension of Poincaré's last geometric theorem. Acta Math. **47**, 297–311 (1925)

47. Birkhoff, G.D.: Dynamical Systems. American Mathematical Society Colloquium Publications, vol. IX (1927)
48. Birkhoff, G.D., Lewis, D.C.: On the periodic motions near a given periodic motion of a dynamical system. Ann. Mat. Pura Appl. **12**, 117–133 (1933)
49. Boardman, J.M.: Singularities of differentiable maps. Publ. Math. Inst. Hautes Études Sci. **33**, 21–57 (1967)
50. Borel, A.: Essays in the History of Lie Groups and Algebraic Groups. History of Mathematics, vol. 21. American Mathematical Society and London Mathematical Society, Providence (2001)
51. Bott, R.: Morse theory indomitable. Publ. Math. Inst. Hautes Études Sci. **68**, 99–114 (1988)
52. Bott, R., Tu, L.W.: Differential Forms in Algebraic Topology. Springer, Berlin (1982)
53. Bourbaki, N.: Topology, vol. I. Springer, Berlin (1998)
54. Bredon, G.E.: Introduction to Compact Transformation Groups. Pure and Applied Mathematics, vol. 46. Academic Press, San Diego (1972)
55. Bredon, G.E.: Topology and Geometry. Graduate Texts in Mathematics, vol. 139. Springer, Berlin (1997)
56. Byun, Y.: Maslov indices in practice. Trends Math. **5**(1), 1–9 (2002). Information Center for Mathematical Sciences
57. Cannas da Silva, A.: Lectures on Symplectic Geometry. Lecture Notes in Mathematics, vol. 1764. Springer, Berlin (2008)
58. Cappell, S.E., Lee, R., Miller, E.Y.: On the Maslov index. Commun. Pure Appl. Math. **47**, 121–186 (1994)
59. Carr, J.: Applications of Centre Manifold Theory. Applied Mathematical Sciences, vol. 35. Springer, Berlin (1981)
60. Cartan, É.: Leçons sur les Invariants Intégraux. Hermann, Paris (1922)
61. Charzynski, Sz., Kijowski, J., Rudolph, G., Schmidt, M.: On the stratified classical configuration space of lattice QCD. J. Geom. Phys. **55**, 137–178 (2005)
62. Charzynski, Sz., Rudolph, G., Schmidt, M.: On the topology of the reduced classical configuration space of lattice QCD. J. Geom. Phys. **58**, 1607–1623 (2008)
63. Chierchia, L.: Kolmogorov-Arnold-Moser (KAM) theory. In: Meyers, R.A. (editor-in-chief) Encyclopedia of Complexity and Systems Science, pp. 5064–5091. Springer, Berlin (2009)
64. Choquet-Bruhat, Y., De Witt-Morette, C., Dillard-Bleick, M.: Analysis, Manifolds and Physics. Part I: Basics. North-Holland, Amsterdam (1982)
65. Chossat, P., Lauterbach, R.: Methods in Equivariant Bifurcations and Dynamical Systems. Advanced Series in Nonlinear Dynamics, vol. 15. World Scientific, Singapore (2000)
66. Conley, C., Zehnder, E.: The Birkhoff-Lewis fixed point theorem and a conjecture of V.I. Arnold. Invent. Math. **73**, 33–49 (1983)
67. Cordani, B.: The Kepler Problem. Birkhäuser, Basel (2003)
68. Coulson, C.A., Joseph, A.: A constant of the motion for the two-centre Kepler problem. Int. J. Quant. Chem. **1**(4), 337–347 (1967)
69. Cushman, R.H., Bates, L.M.: Global Aspects of Classical Integrable Systems. Birkhäuser, Basel (1997)
70. Cushman, R.H., Kemppainen, D., Śniatycki, J., Bates, L.M.: Geometry of nonholonomic constraints. Rep. Math. Phys. **36**(2/3), 275–286 (1995)
71. Dazord, P., Delzant, Th.: Le problème général des variables actions-angles. J. Differ. Geom. **26**, 223–251 (1987)
72. de Gosson, M.: Symplectic Geometry and Quantum Mechanics. Operator Theory Advances and Applications, vol. 166. Birkhäuser, Basel (2006)
73. Dieudonné, J.: Grundzüge der modernen Analysis, Bände I und III. Deutscher Verlag der Wissenschaften, Berlin (1971)
74. Dittmann, J., Rudolph, G.: Canonical realizations of Lie algebras associated with foliated coadjoint orbits. Ann. Inst. Henri Poincaré **43**(3), 251–267 (1985)

75. Douady, R.: Une démonstration directe de l'équivalence des théorèmes de tores invariants pour difféomorphismes et champs des vecteurs. C. R. Acad. Sci. Paris Sér. I Math. **295**, 201–204 (1982)
76. Dubrovin, B.A., Novikov, S.P., Fomenko, A.T.: Modern Geometry—Methods and Applications. Springer, Berlin (1992)
77. Duflo, M., Vergne, M.: Une propriété de la représentation coadjointe d'une algèbre de Lie. C. R. Acad. Sci. Paris Sér. A-B **268**, A583–A585 (1969)
78. Duistermaat, J.J.: Oscillatory integrals, Lagrange immersions and unfolding of singularities. Commun. Pure Appl. Math. **27**, 207–281 (1974)
79. Duistermaat, J.J.: On the Morse index in variational calculus. Adv. Math. **21**, 173–195 (1976)
80. Duistermaat, J.J.: On global action angle coordinates. Commun. Pure Appl. Math. **33**, 686–706 (1980)
81. Duistermaat, J.J.: Dynamical systems with symmetry. In: van der Ban, E., Kolk, J.A.C. (eds.) Utrecht Spring Summer School on Lie Groups. Springer, Berlin (to appear). Available at http://www.projects.science.uu.nl/Duistermaat/www/homepageHD/index.html
82. Ebin, D.G., Marsden, J.E.: Groups of diffeomorphisms and the motion of an incompressible fluid. Ann. Math. **92**, 102–163 (1970)
83. Efstathiou, K., Joyeux, M., Sadovskii, D.A.: Global bending quantum number and the absence of monodromy in the HCN-CNH molecule. Phys. Rev. A **69**, 032504 (2004)
84. Ekeland, I., Hofer, H.: Symplectic topology and Hamiltonian dynamics. Math. Z. **200**, 355–378 (1990)
85. Eliashberg, Y.: Estimates on the number of fixed points of area preserving transformations. Syktyvkar University preprint (1979)
86. Erdélyi, A.: Asymptotic Expansions. Dover, New York (1956)
87. Faddeev, L.D.: Modern mathematical physics: what it should be. math-ph/0002018v2
88. Fadell, E., Rabinowitz, P.: Generalized cohomological index theories for the group actions with an application to bifurcation questions for Hamiltonian systems. Invent. Math. **45**, 139–174 (1978)
89. Fasso, F.: The Euler-Poinsot top: a non-commutatively integrable system without global action-angle coordinates. Z. Angew. Math. Phys. **47**, 953–976 (1996)
90. Field, M.J.: Dynamics and Symmetry. ICP Advanced Texts in Mathematics, vol. 3. Imperial College Press, London (2007)
91. Fiorani, E., Sardanashvily, G.: Global action-angle coordinates for completely integrable systems with noncompact invariant submanifolds. J. Math. Phys. **48**, 032901 (2007)
92. Fischer, E., Rudolph, G., Schmidt, M.: A lattice gauge model of singular Marsden-Weinstein reduction. Part I. Kinematics. J. Geom. Phys. **57**, 1193–1213 (2007)
93. Floer, A.: The unregularized gradient flow of the symplectic action. Commun. Pure Appl. Math. **41**, 775–813 (1988)
94. Floer, A.: Symplectic fixed points and holomorphic spheres. Commun. Math. Phys. **120**, 513–547 (1989)
95. Florentino, C.A.A., Lawton, S.: The topology of moduli spaces of free group representations. Math. Ann. **345**, 453–489 (2009)
96. Fomenko, A.T., Fuchs, D.B., Gutenmacher, V.L.: Homotopic Topology. Akadémiai Kiadó, Budapest (1986)
97. Frankel, Th.: Critical submanifolds of the classical groups and Stiefel manifolds. In: Cairns, S.S. (ed.) Differential and Combinatorial Topology, pp. 37–53. Princeton University Press, Princeton (1965)
98. Freed, D.S., Uhlenbeck, K.K.: Instantons and Four-Manifolds. Springer, Berlin (1984)
99. Freifeld, C.: One-parameter subgroups do not fill a neighbourhood of the identity in an infinite-dimensional Lie (pseudo-) group. In: DeWitt, C.M., Wheeler, J.A. (eds.) Battelle Rencontres. 1967 Lectures in Mathematics and Physics, pp. 538–543. Benjamin, Elmsford (1968)
100. Fukaya, K., Ono, K.: Arnold conjecture and Gromov-Witten invariants. Topology **38**, 933–1048 (1999)

101. Galindo, A., Pascual, P.: Quantum Mechanics, vol. II. Springer, Berlin (1991)
102. Geiges, H.: Contact geometry. In: Dillen, F.J.E., Verstraelen, L.C.A. (eds.) Handbook of Differential Geometry, vol. II, pp. 315–382. Elsevier, Berlin (2005)
103. Geiges, H.: An Introduction to Contact Topology. Cambridge Studies in Advanced Mathematics, vol. 109. Cambridge University Press, Cambridge (2010)
104. Ginzburg, V.L.: Some remarks on symplectic actions of compact groups. Math. Z. **210**, 625–640 (1992)
105. Ginzburg, V.L.: An embedding $S^{2n-1} \to \mathbb{R}^{2n}$, $2n - 1 \geq 7$, whose Hamiltonian flow has no periodic trajectories. Int. Math. Res. Not. **2**, 83–98 (1995)
106. Ginzburg, V.L.: The Weinstein conjecture and the theorems of nearby and almost existence. In: Marsden, J.E., Ratiu, T.S. (eds.) The Breadth of Symplectic and Poisson Geometry. Festschrift in Honor of Alan Weinstein, pp. 139–172. Birkhäuser, Basel (2005)
107. Gleason, A.M.: Spaces with a compact Lie group of transformations. Proc. Am. Math. Soc. **1**, 35–43 (1950)
108. Goldstein, H.: Prehistory of the Runge-Lenz vector. Am. J. Phys. **43**, 737–738 (1975)
109. Goldstein, H.: More on the prehistory of the Runge-Lenz vector. Am. J. Phys. **44**, 1123–1124 (1976)
110. Golubitski, M., Guillemin, V.: Stable Mappings and Their Singularities. Graduate Texts in Mathematics, vol. 14. Springer, Berlin (1974)
111. Gotay, M.J., Isenberg, J.A.: The symplectization of science. Gaz. Math. **54**, 59–79 (1992)
112. Greenberg, J.T.: Mathematical physics in eighteenth-century France. Isis **77**(1), 59–78 (1986). Published by the University of Chicago Press on behalf of the History of Science Society
113. Grondin, L., Sadovskii, D.A., Zhilinskii, B.I.: Monodromy as topological obstruction to global action-angle variables in systems with coupled angular momenta and rearrangement of bands in quantum spectra. Phys. Rev. A **65**, 012105 (2002)
114. Guckenheimer, J., Holmes, P.: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Springer, Berlin (1986)
115. Guillemin, V., Sternberg, S.: Geometric Asymptotics. Math. Surveys Monogr., vol. 14. AMS, Providence (1977)
116. Guillemin, V., Sternberg, S.: Symplectic Techniques in Physics. Cambridge University Press, Cambridge (1984)
117. Guillemin, V., Sternberg, S.: A normal form for the momentum map. In: Sternberg, S. (ed.) Differential Geometric Methods in Mathematical Physics. Mathematical Physics Studies, vol. 6. Reidel, Dordrecht (1984)
118. Guillemin, V., Sternberg, S.: Variations on a Theme by Kepler. AMS Colloquium Publications, vol. 42 (1990)
119. Guillemin, V., Sternberg, S.: Semi-classical analysis. Lecture Notes (2010). Available at Victor Guillemins homepage, http://math.mit.edu/~vwg
120. Hamilton, W.R.: On a general method in dynamics. Philos. Trans. R. Soc., part II, 247–308 (1834)
121. Hamilton, W.R.: On the application to dynamics of a general mathematical method previously applied to optics. British Association Report, Edinburgh, pp. 513–518 (1834)
122. Hamilton, W.R.: Second essay on a general method in dynamics. Philos. Trans. R. Soc., part I, 95–144 (1835)
123. Hartman, P.: Ordinary Differential Equations. Wiley, New York (1964)
124. Hatcher, A.: Algebraic Topology. Cambridge University Press, Cambridge (2002)
125. Hatcher, A.: Vector Bundles & K-Theory (2009). Available at Allen Hatcher's homepage, http://www.math.cornell.edu/~hatcher/
126. Hawkins, Th.: Emergence of the Theory of Lie Groups. Sources and Studies in the History of Mathematics and Physical Sciences. Springer, Berlin (2000)
127. Helgason, S.: Differential Geometry, Lie Groups, and Symmetric Spaces. Academic Press, San Diego (1978)

128. Hilbert, D.: Mathematical problems. Lecture delivered before the International Congress of Mathematicians at Paris in 1900. Bull. Am. Math. Soc. **8**, 437–479 (1900). The original article in German appeared in Gött. Nachr. **3**, 253–297 (1900)
129. Hilgert, J., Neeb, K.-H.: Lie-Gruppen und Lie-Algebren. Vieweg, Wiesbaden (1991)
130. Hirsch, M.W.: Differential Topology. Springer, Berlin (1976)
131. Hirsch, M.W., Pugh, C.C.: Stable manifolds and hyperbolic sets. Bull. Am. Math. Soc. **75**, 149–152 (1969)
132. Hirsch, M.W., Pugh, C.C., Shub, M.: Invariant Manifolds. Lecture Notes in Mathematics, vol. 583. Springer, Berlin (1977)
133. Hirzebruch, F.: Topological Methods in Algebraic Geometry. Springer, Berlin (1978)
134. Hofer, H.: On the topological properties of symplectic maps. Proc. R. Soc. Edinb. **115**, 25–38 (1990)
135. Hofer, H.: Pseudoholomorphic curves in symplectizations with applications to the Weinstein conjecture in dimension three. Invent. Math. **114**, 515–563 (1993)
136. Hofer, H.: Dynamics, topology and holomorphic curves. In: Proceedings of the International Congress of Mathematicians, Berlin, 1998. Doc. Math., extra vol. I pp. 255–280 (1998)
137. Hofer, H., Salamon, D.: Floer homology and Novikov rings. In: The Floer Memorial Volume. Progress in Mathematics, vol. 133, pp. 483–524. Birkhäuser, Basel (1995)
138. Hofer, H., Zehnder, E.: A new capacity for symplectic manifolds. In: Rabinowitz, P., Zehnder, E. (eds.) Analysis et Cetera, pp. 405–428. Academic Press, San Diego (1990)
139. Hofer, H., Zehnder, E.: Symplectic Invariants and Hamiltonian Dynamics. Birkhäuser Advanced Texts. Birkhäuser, Basel (1994)
140. Hoppe, J.: Lectures on Integrable Systems. Lecture Notes in Physics. Springer, Berlin (1992)
141. Hörmander, L.: Fourier integral operators I. Acta Math. **127**, 79–183 (1971)
142. Huebschmann, J.: Kähler Spaces, Nilpotent Orbits, and Singular Reduction. Mem. Am. Math. Soc., vol. 172(814) (2004)
143. Huebschmann, J.: Singular Poisson-Kähler geometry of certain adjoint quotients. In: The Mathematical Legacy of C. Ehresmann, Bedlewo, 2005. Banach Center Publications, vol. 76, pp. 325–347 (2005)
144. Huebschmann, J., Rudolph, G., Schmidt, M.: A lattice gauge model for quantum mechanics on a stratified space. Commun. Math. Phys. **286**, 459–494 (2009)
145. Humphreys, J.E.: Introduction to Lie Algebras and Representation Theory. Graduate Texts in Mathematics. Springer, Berlin (1972)
146. Hutchings, M.: Lecture notes on Morse homology (with an eye towards Floer theory and pseudoholomorphic curves). UC Berkeley preprint (2002). Available at Michael Hutchings' homepage, http://math.berkeley.edu/~hutching/
147. Iglesias, D.I., Marrero, J.C., de Diego, D.M., Sosa, D.: A general framework for nonholonomic mechanics: nonholonomic systems on Lie algebroids. J. Math. Phys. **48**, 083513 (2007)
148. Irwin, M.C.: A classification of elementary cycles. Topology **9**, 35–47 (1970)
149. Jacobson, N.: Lie Algebras. Dover Publishers, New York (1979)
150. Jänich, K.: Differenzierbare G-Mannigfaltigkeiten. Lecture Notes in Mathematics, vol. 59. Springer, Berlin (1968)
151. Jaffe, A., Quinn, F.: Theoretical mathematics: toward a cultural synthesis of mathematics and theoretical physics. Bull. Am. Math. Soc. **29**(1), 1–13 (1993)
152. Janeczko, S.: Singularities of Lagrangian varieties and related topics. Acta Appl. Math. **25**, 21–57 (1991)
153. José, J.V., Saletan, E.J.: Classical Dynamics. A Contemporary Approach. Cambridge University Press, Cambridge (1998)
154. Jovanovic, B.: Symmetries and integrability. Publ. Inst. Math. (Belgr.) **84**(98), 1–36 (2008)
155. Kac, M., van Moerbeke, P.: A complete solution of the periodic Toda problem. Proc. Natl. Acad. Sci. USA **72**, 2879–2880 (1975)
156. Karoubi, M.: K-Theory: An Introduction. Grundlehren Math. Wiss., vol. 226. Springer, Berlin (1978)

157. Keller, J.B.: A geometrical theory of diffraction. In: Calculus of Variations and Its Applications. Proc. Sympos. Appl. Math., vol. 8, pp. 27–52. McGraw-Hill, New York (1958)
158. Kervaire, M.: A manifold which does not admit any differentiable structure. Comment. Math. Helv. **34**, 304–312 (1960)
159. Kervaire, M.A., Milnor, J.W.: Groups of homotopy spheres I. Ann. Math. **77**, 504–537 (1963)
160. Kijowski, J., Tulczyjew, W.M.: A Symplectic Framework for Field Theories. Lecture Notes in Physics, vol. 107. Springer, Berlin (1979)
161. Kijowski, J., Rudolph, G.: On the Gauß law and global charge for QCD. J. Math. Phys. **43**, 1796–1808 (2002)
162. Kijowski, J., Rudolph, G.: Charge superselection sectors for QCD on the lattice. J. Math. Phys. **46**, 032303 (2005)
163. Jarvis, P.D., Kijowski, J., Rudolph, G.: On the structure of the observable algebra of QCD on the lattice. J. Phys. A, Math. Gen. **38**, 5359–5377 (2005)
164. Rudolph, G., Schmidt, M.: On the algebra of quantum observables for a certain gauge model. J. Math. Phys. **50**, 052102 (2009)
165. Klingenberg, W.: Lectures on Closed Geodesics. Grundlehren Math. Wiss., vol. 230. Springer, Berlin (1978)
166. Kobayashi, S., Nomizu, K.: Foundations of Differential Geometry. Wiley-Interscience, New York (1963)
167. Kogut, J., Susskind, L.: Hamiltonian formulation of Wilson's lattice gauge theories. Phys. Rev. D **11**, 395–408 (1975)
168. Kolmogorov, A.N.: On conservation of conditionally periodic motions under small perturbations of the Hamiltonian. Dokl. Akad. Nauk SSSR **98**(4), 527–530 (1954), in Russian
169. Kolmogorov, A.N.: General theory of dynamical systems and classical mechanics. In: Proc. Int. Congr. Math., vol. 1, Amsterdam, 1954, pp. 315–333 (1957), in Russian. English translation in Abraham, R., Marsden, J.E.: Foundations of Mechanics, pp. 741–757. Benjamin-Cummings, Reading (1978)
170. Koon, W.S., Marsden, J.E.: The Hamiltonian and Lagrangian approaches to the dynamics of nonholonomic systems. Rep. Math. Phys. **40**(1), 21–62 (1997)
171. Kostant, B.: The solution to a generalized Toda lattice and representation theory. Adv. Math. **39**, 195–338 (1979)
172. Koszul, J.L.: Sur certains groupes de transformations de Lie. In: Géométrie Différentielle, Colloques Internationaux du CNRS, Strasbourg, 1953, pp. 137–141. CNRS, Paris (1953), in French
173. Krichever, I.M.: Algebraic curves and nonlinear difference equations. Usp. Mat. Nauk **33**(4), 215–216 (1978), in Russian
174. Krupa, M.: Bifurcation of relative equilibria. SIAM J. Math. Anal. **21**, 1453–1486 (1990)
175. Kummer, M.: On the regularization of the Kepler problem. Commun. Math. Phys. **84**, 133–152 (2002)
176. Kustaanheimo, P., Stiefel, E.: Perturbation theory of Kepler motion based on spinor regularization. J. Reine Angew. Math. **218**, 609–636 (1965)
177. Lagrange, J.L.: Mémoire sur la théorie des variations des éléments des planètes. Mém. Cl. Sci. Math. Phys. Inst. France, 1–72 (1808)
178. Lalonde, F., McDuff, D.: The geometry of symplectic energy. Ann. Math. **141**, 349–371 (1995)
179. Lalonde, F., McDuff, D., Polterovich, L.: On the flux conjectures. In: Geometry, Topology, and Dynamics. CRM Proc. Lecture Notes, vol. 15, pp. 69–85. AMS, Providence (1998)
180. Lang, S.: Differential and Riemannian Manifolds. Springer, Berlin (1995)
181. Libermann, P., Marle, C.-M.: Symplectic Geometry and Analytical Mechanics. Reidel, Dordrecht (1987)
182. Lie, S.: Theorie der Transformationsgruppen I. Teubner, Leipzig (1888). Written with the help of Friedrich Engel

183. Lie, S.: Theorie der Transformationsgruppen II. Teubner, Leipzig (1890). Written with the help of Friedrich Engel
184. Lie, S.: Theorie der Transformationsgruppen III. Teubner, Leipzig (1893). Written with the help of Friedrich Engel
185. Ligon, T., Schaaf, M.: On the global symmetry of the classical Kepler problem. Rep. Math. Phys. **9**, 281–300 (1976)
186. Liu, G., Tian, G.: Floer homology and Arnold conjecture. J. Differ. Geom. **49**, 1–74 (1998)
187. Ludwig, D.: Uniform asymptotic expansions at a caustic. Commun. Pure Appl. Math. **19**, 215–250 (1966)
188. Lyapunov, A.M.: Problème général de la stabilité du mouvement. Ann. Fac. Sci. Toulouse **2**, 203–474 (1907)
189. Markus, L., Meyer, K.R.: Generic Hamiltonian Systems are Neither Integrable nor Ergodic. Mem. Am. Math. Soc., vol. 144 (1974)
190. Marle, C.-M.: Normal forms generalizing action-angle coordinates for Hamiltonian actions of Lie groups. Lett. Math. Phys., **7**, 55–62 (1983)
191. Marle, C.-M.: Le voisinage d'une orbite d'une action hamiltonienne d'un groupe de Lie. In: Dazord, P., Desolneux-Moulis, N. (eds.) Séminaire Sud-Rhodanien de Geometrie II, pp. 19–35 (1984)
192. Marle, C.-M.: Modèle d'action hamiltonienne d'un groupe de Lie sur une variété symplectique. Rend. Semin. Mat. (Torino) **43**(2), 227–251 (1985)
193. Marle, C.-M.: Various approaches to conservative and nonconservative nonholonomic constraints. Rep. Math. Phys. **42**(1/2), 211–228 (1998)
194. Marsden, J.E., Ratiu, T.S.: Introduction to Mechanics and Symmetry. Texts in Applied Mathematics, vol. 17. Springer, Berlin (1999)
195. Marsden, J.E., Weinstein, A.: Reduction of symplectic manifolds with symmetries. Rep. Math. Phys. **5**, 121–130 (1974)
196. Marsden, J.E., Weinstein, A.: Some comments on the history, theory and applications of symplectic reduction. In: Landsman, N.P., Pflaum, M.J., Schlichenmaier, M. (eds.) Quantization of Singular Symplectic Quotients. Progress in Mathematics, vol. 198, pp. 1–19. Birkhäuser, Basel (2001)
197. Maslov, V.P.: Théorie des Perturbations et Méthodes Asymptotiques. Gauthier-Villars, Paris (1971)
198. Maslov, V.P., Fedoryuk, M.V.: Quasi-Classical Approximation for the Equations of Quantum Mechanics. Reidel, Dordrecht (1981)
199. Massey, W.S.: A Basic Course in Algebraic Topology. Graduate Texts in Mathematics, vol. 127. Springer, New York (1991)
200. Mather, J.N.: Stability of C^∞-mappings: I. The division theorem. Ann. Math. **87**, 89–104 (1968)
201. Mather, J.N.: Stability of C^∞-mappings: II. Infinitesimal stability implies stability. Ann. Math. **89**, 254–291 (1969)
202. Mather, J.N.: Stability of C^∞-mappings: III. Finitely determined map-germs. Publ. Math. Inst. Hautes Études Sci. **35**, 127–156 (1968)
203. Mather, J.N.: Stability of C^∞-mappings: IV. Classification of stable germs by R-relations. Publ. Math. Inst. Hautes Études Sci. **37**, 223–248 (1969)
204. Mather, J.N.: Stability of C^∞-mappings: V. Transversality. Adv. Math. **4**, 301–336 (1970)
205. McDuff, D.: Symplectic diffeomorphisms and the flux homomorphism. Invent. Math. **77**, 353–366 (1984)
206. McDuff, D., Salamon, D.: Introduction to Symplectic Topology. Clarendon Press, Oxford (1998)
207. Medved, M.: Fundamentals of Dynamical Systems and Bifurcation Theory. Hilger, Bristol (1992)
208. Meinrenken, E.: Symplectic geometry. Lecture Notes, University of Toronto (2000)
209. Meixner, J., Schaefke, W.: Mathieusche Funktionen und Sphäroidfunktionen. Grundlehren Math. Wiss., vol. 71. Springer, Berlin (1954), in German

210. Meyer, K.R.: Symmetries and integrals in mechanics. In: Peixoto, M.M. (ed.) Dynamical Systems, pp. 259–273. Academic Press, San Diego (1973)
211. Michor, P.W.: Topics in Differential Geometry. Graduate Studies in Mathematics. AMS, Providence (2008)
212. Milnor, J.: Morse Theory. Princeton University Press, Princeton (1963)
213. Milnor, J.: Remarks on infinite-dimensional Lie groups. In: DeWitt, B.S., Stora, R. (eds.) Relativité, Groupes et Topologie II. Les Houches, Session XL, 1983, pp. 1007–1057. North-Holland, Amsterdam (1984)
214. Mishchenko, A.S., Fomenko, A.T.: A generalized Liouville method for the integration of Hamiltonian systems. Funct. Anal. Appl. **12**, 113–121 (1978)
215. Montgomery, D., Yang, C.T.: The existence of a slice. Ann. of Math. (2) **65**, 108–116 (1957)
216. Morbidelli, A.: Modern Celestial Mechanics. Taylor & Francis, London (2002)
217. Moser, J.: On invariant curves of area preserving mappings of the annulus. Nachr. Akad. Wiss. Göttingen, Math. Phys. Kl. **II**, 286–294 (1962)
218. Moser, J.: On the volume elements on a manifold. Trans. Am. Math. Soc. **120**, 286–294 (1965)
219. Moser, J.: A rapidly convergent iteration method and nonlinear partial differential equations I. Ann. Sc. Norm. Super. Pisa, Cl. Sci. **20**(2), 265–315 (1966)
220. Moser, J.: A rapidly convergent iteration method and nonlinear partial differential equations II. Ann. Sc. Norm. Super. Pisa, Cl. Sci. **20**(3), 499–533 (1966)
221. Moser, J.: On the construction of almost periodic solutions for ordinary differential equations. In: Proceedings of the International Conference on Functional Analysis and Related Topics, Tokyo, pp. 60–67 (1969)
222. Moser, J.: Regularization of Kepler's problem and the averaging method on a manifold. Commun. Pure Appl. Math. **23**, 609–636 (1970)
223. Moser, J.: Three integrable Hamiltonian systems connected with isospectral deformations. Adv. Math. **16**, 197–220 (1975)
224. Moser, J.: Periodic orbits near an equilibrium and a theorem by Alan Weinstein. Commun. Pure Appl. Math. **29**, 727–747 (1976)
225. Moser, J.: Proof of a generalized form of a fixed point theorem due to G.D. Birkhoff. In: Geometry and Topology. Lecture Notes in Mathematics, vol. 597, pp. 464–494. Springer, Berlin (1977)
226. Moser, J.: Addendum to "Periodic orbits near an equilibrium and a theorem by Alan Weinstein". Commun. Pure Appl. Math. **31**, 529–530 (1978)
227. Mostow, G.D.: Equivariant embeddings in Euclidean space. Ann. of Math. (2) **65**, 432–446 (1957)
228. Nekhoroshev, L.: Action-angle coordinates and their generalizations. Trans. Mosc. Math. Soc. **26**, 180–198 (1972)
229. Newns, N., Walker, A.: Tangent planes to a differentiable manifold. J. Lond. Math. Soc. **31**, 400–407 (1956)
230. Omori, H.: Infinite Dimensional Lie Transformation Groups. Lecture Notes in Mathematics, vol. 427. Springer, Berlin (1974)
231. Ono, K.: Floer-Novikov cohomology and the flux conjecture. Geom. Funct. Anal. **16**, 981–1020 (2006)
232. Ortega, J.-P., Ratiu, T.S.: Momentum Maps and Hamiltonian Reduction. Progress in Mathematics, vol. 222. Birkhäuser, Basel (2004)
233. Palais, R.S.: The Classification of G-Spaces. Mem. Am. Math. Soc., vol. 36 (1960)
234. Palais, R.S.: On the existence of slices for actions of non-compact Lie groups. Ann. of Math. (2) **73**, 295–323 (1961)
235. Palais, R.S.: The Morse lemma for Banach spaces. Bull. Am. Math. Soc. **75**, 968–971 (1969)
236. Palis, J.: On the local structure of hyperbolic points in Banach spaces. An. Acad. Bras. Ciênc. **40**, 263–266 (1968)
237. Perelomov, A.M.: Integrable Systems of Classical Mechanics and Lie Algebras, vol. I. Birkhäuser, Basel (1990)

238. Pflaum, M.J.: Analytic and Geometric Study of Stratified Spaces. Lecture Notes in Mathematics, vol. 1768. Springer, Berlin (2001)
239. Poincaré, H.: Analysis situs. J. Éc. Polytech. **2**(1), 1–123 (1895)
240. Poincaré, H.: Méthodes Nouvelles de la Méchanique Céleste. Gauthier-Villars, Paris (1899)
241. Poincaré, H.: L'état actuel et l'avenir de la physique mathématique. Bull. Sci. Math. **28**, 302–324 (1904). English translation by George Bruce Halsted: The principles of mathematical physics. In: The Foundations of Science (The Value of Science), pp. 297–320. Science Press, New York (1913), Chaps. 7–9
242. Polterovich, L.: The Geometry of the Group of Symplectic Diffeomorphisms. Birkhäuser, Basel (2001)
243. Pontrjagin, L.S.: Topological Groups. Princeton University Press, Princeton (1939)
244. Pöschel, J.: Über invariante Tori in differenzierbaren Hamiltonschen Systemen. Bonner Math. Schr. **120**, 1–103 (1980)
245. Pöschel, J.: Integrability of Hamiltonian systems on Cantor sets. Commun. Pure Appl. Math. **35**, 653–695 (1982)
246. Pöschel, J.: A lecture on the classical KAM theorem. Proc. Symp. Pure Math. **69**, 707–732 (2001)
247. Procesi, C.: Lie Groups. An Approach Through Invariants and Representations. Springer, Berlin (2007)
248. Pugh, C.C.: On a theorem of P. Hartman. Am. J. Math. **91**, 363–367 (1969)
249. Rabinowitz, P.: Periodic solutions of Hamiltonian systems. Commun. Pure Appl. Math. **31**, 157–184 (1978)
250. Reinfelds, A.A.: A reduction theorem for closed trajectories. Differ. Equ. **11**, 1353–1358 (1975)
251. Riemann, B.: Über die Hypothesen, welche der Geometrie zu Grunde liegen. Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen **13**, 133–152 (1867), in German
252. Robbin, J., Salamon, D.: The Maslov index for paths. Topology **32**, 827–844 (1993)
253. Römer, H.: Theoretische Optik. VCH, Weinheim (1994)
254. Rudolph, G., Schmidt, M., Volobuev, I.P.: Classification of gauge orbit types for $\mathrm{SU}(n)$-gauge theories. Math. Phys. Anal. Geom. **5**, 201–241 (2002)
255. Rudolph, G., Schmidt, M., Volobuev, I.P.: Partial ordering of gauge orbit types for $\mathrm{SU}(n)$-gauge theories. J. Geom. Phys. **42**, 106–138 (2002)
256. Rudolph, G., Schmidt, M., Volobuev, I.P.: On the gauge orbit space stratification: a review. J. Phys. A, Math. Gen. **35**, R1–R50 (2002)
257. Ruelle, D.: Dynamical Systems and Bifurcation Theory. Academic Press, San Diego (1989)
258. Sadetov, S.T.: A proof of the Mishchenko-Fomenko conjecture. Dokl. Akad. Nauk SSSR **397**(6), 751–754 (2004), in Russian
259. Salamon, D.: The Kolmogorov-Arnold-Moser theorem. Preprint, ETH-Zürich (1986)
260. Salamon, D.: Lectures on Floer homology. In: Eliashberg, Y., Traynor, L. (eds.) Symplectic Geometry and Topology. IAS/Park City Mathematics Series, vol. 7, pp. 143–230 (1999)
261. Sardanashvily, G.: Superintegrable Hamiltonian systems with noncompact invariant submanifolds. Kepler systems. Int. J. Geom. Methods Mod. Phys. **6**(8), 1391–1414 (2009)
262. Schmid, R.: Die Symplektomorphismengruppe als Fréchet-Lie-Gruppe. Dissertation, Universität Zürich, Juris Druck und Verlag Zürich (1978), in German
263. Schmid, R.: Infinite dimensional Lie groups with applications to mathematical physics. J. Geom. Symmetry Phys. **1**, 54–120 (2004)
264. Scholz, E.: Geschichte des Mannigfaltigkeitsbegriffs von Riemann bis Poincaré. Birkhäuser, Basel (1980)
265. Scholz, E.: Riemanns frühe Notizen zum Mannigfaltigkeitsbegriff und zu den Grundlagen der Geometrie. Arch. Hist. Exact Sci. **27**, 213–282 (1982)
266. Scholz, E.: The concept of manifold, 1850–1950. In: James, I.M. (ed.) History of Topology, pp. 25–64. Elsevier, Amsterdam (1999)
267. Schubert, H.: Topologie. Teubner, Stuttgart (1975)

268. Schwarz, M.: Morse Homology. Progress in Mathematics, vol. 111. Birkhäuser, Basel (1993)
269. Schwarz, M.: Introduction to symplectic Floer homology. In: Thomas, C.B. (ed.) Contact and Symplectic Geometry. Publ. Newton Inst., vol. 8, pp. 151–170. Cambridge University Press, Cambridge (1996)
270. Sevryuk, M.B.: Some problems of KAM-theory: conditionally-periodic motion in typical systems. Russ. Math. Surv. **50**(2), 341–353 (1995)
271. Shoshitajshvili, A.M.: Bifurcations of topological type of a vector field near a singular point. Tr. Semin. Im. I.G. Petrovskogo **1**, 279–308 (1975), in Russian
272. Siegel, C.L.: Symplectic geometry. Am. J. Math. **65**, 1–86 (1943)
273. Siegel, C.L., Moser, J.K.: Lectures on Celestial Mechanics. Grundlehren Math. Wiss., vol. 187. Springer, Berlin (1971)
274. Sjamaar, R.: Singular orbit spaces in Riemannian and symplectic geometry. Ph.D. thesis, Rijksuniversiteit te Utrecht (1990)
275. Sjamaar, R., Lerman, E.: Stratified symplectic spaces and reduction. Ann. Math. **134**, 375–422 (1991)
276. Smale, S.: Generalized Poincaré's conjecture in dimension greater than four. Ann. Math. **74**, 391–406 (1961)
277. Smale, S.: Topology and mechanics I. Invent. Math. **10**, 305–331 (1970)
278. Smale, S.: Topology and mechanics II. Invent. Math. **11**, 45–64 (1970)
279. Steenrod, N.: The Topology of Fibre Bundles. Princeton University Press, Princeton (1951)
280. Stefan, P.: Accessible sets, orbits, and foliations with singularities. Proc. Lond. Math. Soc. **29**, 699–713 (1974)
281. Struik, D.J.: Outline of a history of differential geometry. Part I. Isis **19**(1), 92–120 (1933)
282. Struik, D.J.: Outline of a history of differential geometry. Part II. Isis **20**(1), 161–191 (1933). Published by the Chicago Press on behalf of The History of Science Society
283. Sussmann, H.J.: Orbits of families of vector fields and integrability of distributions. Trans. Am. Math. Soc. **180**, 171–188 (1973)
284. Symes, W.: Systems of Toda type, inverse spectral problems and representation theory. Invent. Math. **59**, 13–53 (1980)
285. Takens, F.: A note on the differentiability of centre manifolds. In: Dynamical Systems and Partial Differential Equations, pp. 101–104. Univ. Simon Bolívar, Caracas (1986)
286. Thirring, W.: Lehrbuch der Mathematischen Physik. Band 1: Klassische Dynamische Systeme. Springer, Berlin (1977), in German
287. Thom, R.: Les singularités des applications différentiables. Ann. Inst. Fourier (Grenoble) **6**, 43–87 (1955)
288. Thom, R.: Structural Stability and Morphogenesis. An Outline of a General Theory of Models. Benjamin, Elmsford (1976)
289. Thomas, T.: The Maslov index as a quadratic space. Math. Res. Lett. **13**, 985–999 (2006)
290. Torrence, R.J., Tulczyjew, W.M.: Gauge invariant canonical mechanics for charged particles. J. Math. Phys. **14**, 1725–1732 (1973)
291. Trautman, A.: Notes on groups and their representations. Warsaw (2000), in Polish
292. Tulczyjew, W.M.: Gauge independent formulation of dynamics of charged particles. Ann. Inst. H. Poincaré Sect. A **34**, 25–43 (1981)
293. Tulczyjew, W.M.: Lagrangian submanifolds, statics and dynamics of dynamical systems. In: Dynamical Systems and Microphysics, pp. 3–26. Academic Press, New York (1982)
294. Tulczyjew, W.M.: Control of static mechanical systems. In: Dynamical Systems and Microphysics, pp. 359–377. Academic Press, Orlando (1984)
295. Vaisman, I.: Lectures on the Geometry of Poisson Manifolds. Birkhäuser, Basel (1994)
296. van Strien, S.J.: Center manifolds are not C^∞. Math. Z. **166**, 143–145 (1979)
297. Varadarajan, V.S.: Lie Groups, Lie Algebras, and Their Representations. Springer, Berlin (1984)
298. Vershik, A.M., Gershkovich, V.Ya.: Nonholonomic dynamical systems. Geometry of distributions and variational problems. In: Arnold, V.I., Novikov, S. (eds.) Dynamical Systems VII. Springer, Berlin (1994)

299. Viterbo, C.: A proof of the Weinstein conjecture in $\mathbb{R}^{2n}$. Ann. Inst. Henri Poincaré, Anal. Non Linéaire **4**, 337–357 (1987)
300. Vivarelli, M.: The KS-transformation in hypercomplex form. Celest. Mech. **29**, 45–50 (1983)
301. Waldmann, S.: Poisson-Geometrie und Deformationsquantisierung. Springer, Berlin (2007), in German
302. Warner, F.W.: Foundations of Differentiable Manifolds and Lie Groups. Scott, Foresman, Glenview (1971). Graduate Texts in Mathematics, vol. 94, Springer (1983)
303. Weinstein, A.: Symplectic manifolds and their Lagrangian submanifolds. Adv. Math. **6**, 329–346 (1971)
304. Weinstein, A.: Normal modes for nonlinear Hamiltonian systems. Invent. Math. **20**, 47–57 (1973)
305. Weinstein, A.: Lectures on Symplectic Manifolds. CBMS Regional Conference Series in Mathematics, vol. 29. AMS, Providence (1977)
306. Weinstein, A.: Periodic orbits for convex Hamiltonian systems. Ann. Math. **108**, 507–518 (1978)
307. Weinstein, A.: On the hypothesis of Rabinowitz's periodic orbit theorems. J. Differ. Equ. **33**, 353–358 (1979)
308. Weinstein, A.: Symplectic geometry. Bull. Am. Math. Soc. **5**(1), 1–13 (1981)
309. Weinstein, A.: Sophus Lie and symplectic geometry. Expo. Math. **1**, 95–96 (1983)
310. Weinstein, A.: The local structure of Poisson manifolds. J. Differ. Geom. **18**, 523–557 (1983)
311. Weinstein, A.: Coisotropic calculus and Poisson groupoids. J. Math. Soc. Jpn. **40**, 705–727 (1988)
312. Weyl, H.: The Classical Groups. Their Invariants and Representations. Princeton University Press, Princeton (1939)
313. Whitney, H.: On singularities of mappings of euclidean spaces: I. Mappings of the plane into the plane. Ann. Math. **62**, 374–410 (1955)
314. Whitehead, J.H.C.: Certain equations in the algebra of a semi-simple infinitesimal group. Quart. J. Math. Oxford **8**, 220–237 (1932)
315. Woodhouse, N.M.J.: Geometric Quantization. Clarendon Press, Oxford (1992)
316. Wulff, C., Lamb, J.S.W., Melbourne, I.: Bifurcation from relative periodic solutions. Ergod. Theory Dyn. Syst. **21**, 605–635 (2001)
317. Zung, N.T.: Symplectic topology of integrable Hamiltonian systems I: Arnold-Liouville with singularities. Compos. Math. **101**, 179–215 (1996)
318. Zung, N.T.: Symplectic topology of integrable Hamiltonian systems II: topological classification. Compos. Math. **138**, 125–156 (2003)

Index

G. Rudolph, M. Schmidt, *Differential Geometry and Mathematical Physics*,
Theoretical and Mathematical Physics,
DOI 10.1007/978-94-007-5345-7,

M

R

S